

# G

## Game Theory, Introduction to

MARILDA SOTOMAYOR<sup>1,2</sup>

<sup>1</sup> Department of Economics, University of São Paulo,  
São Paulo, Brasil

<sup>2</sup> Department of Economics, Brown University,  
Providence, USA

Game theory is the study of decision problems which involves several individuals (the decision-makers or players) interacting rationally. The models of game theory are abstract representations of a number of real-life situations and have applications to economics, political sciences, computer sciences, evolutionary biology, social psychology, law, among others. These applications are also important for the development of the theory, since the questions that emerge may lead to new theoretic results.

This session is an attempt to provide the main features of Game Theory, covering most of the fundamental theoretical aspects under the cooperative, non-cooperative and “general” or “mixed” approaches.

The cooperative approach focuses on the possible outcomes of the decision-maker’s interaction by abstracting from the actions or decisions that may lead to these outcomes. Specifically, Cooperative game theory studies the interactions among coalitions of players. Its main question is: Given the sets of feasible payoffs for each coalition, what payoff will be awarded to each player? One can take a positive or normative approach to answering this question, and different solution concepts in the theory lead towards one or the other.

The first cooperative solution concept is the von Neumann–Morgenstern stable sets, treated in ► [Cooperative Games \(Von Neumann–Morgenstern Stable Sets\)](#) (Jun Wako and Shigeo Muto). However, the two best known solution concepts in cooperative game theory are perhaps the core and the Shapley value which are presented and discussed in ► [Cooperative Games](#) (Roberto Serrano).

The non-cooperative approach focuses on the actions

that the decision-makers can take. Historically, the first contribution to the non-cooperative Game Theory is due to Zermelo (1913), but the idea of a general theory of games was introduced by John von Neumann and Oskar Morgenstern in their famous book of 1944 entitled *Theory of Games and Economic Behavior*. These authors argued that most economic questions should be analyzed as games. They introduced the extensive-form and the strategic-form representations of a game, also known as Dynamic and Static games, respectively.

Dynamic games stress the sequentiality of the various decisions that agents can make. An essential component of a dynamic game is the description of who moves first, who moves second, etc. Static games, on the other hand, abstract from sequentiality of the possible moves and model interactions as simultaneous decisions. All extensive form games can be modeled as static games, and all strategic form games can be modeled as dynamic games. However some situations may be more conveniently modeled as one or the other kind of game. Dynamic games are examined in ► [Dynamic Games with an Application to Climate Change Models](#) (Prajit K. Dutta). The structure, as well as its principal results, is discussed in detail. The Chapter ends with an important application, the economics of climate change.

The main ideas and results related to Static games, as well as some interesting relationships that connect equilibrium concepts with the idea of rationality are reviewed in ► [Static Games](#) (Oscar Volij). In this chapter it is presented the general Theorem of existence of strategic equilibria, due to Nash (1950). This result extends to more general games the minimax Theorem, which was proved in von Neuman (1928) for two-player zero-sum games. In the literature there are two proofs published by Nash. One of them uses Brower’s fixed point theorem. The other one is a simpler proof, attributed to Gale by Nash that uses Kakutani’s fixed point theorem. Some version of the proof that uses Brower’s fixed point theorem, by Geanakoplos (2003), is presented in this chapter as well as some discussion on Correlated equilibrium and Bayesian games.

The Correlated equilibrium is a Game theoretic solution concept proposed by Aumann (1974, 1987) in order to capture the strategic correlation opportunities that the players face when they take into account the extraneous environment in which they interact. The Chap. ► [Correlated Equilibria and Communication in Games](#) (Françoise Forges) focuses on two possible extensions of the correlated equilibrium to Bayesian games: the strategic form correlated equilibrium and the communication equilibrium. The general framework of games with incomplete information is treated in ► [Bayesian Games: Games with Incomplete Information](#) (Shmuel Zamir) with special reference to “Bayesian games”.

Repeated games deals with situations in which a group of agents engage in a strategic interaction over and over. The Chap. ► [Repeated Games with Complete Information](#) (Olivier Gossner and Tristan Tomala) is devoted to Repeated games with complete information. In such games the data of the strategic interaction is fixed over time and is known by all the players. The Chap. ► [Repeated Games with Incomplete Information](#) (Jérôme Renault) discusses Repeated games with incomplete information, a situation where several players repeat the same stage game, the players having different knowledge of the stage game which is repeated.

Repeated games have many equilibria, including the repetition of stage game Nash equilibria. At the same time, particularly when monitoring is imperfect, certain plausible outcomes are not consistent with equilibrium. Reputation effects is the term used for the impact upon the set of equilibria (typically of a repeated game) of perturbing the game by introducing incomplete information of a particular kind. This issue is treated in ► [Reputation Effects](#) (George Mailath).

Games with two players are of particular significance. The first two-person game studied in the literature was the Zero sum two-person game, first analyzed by von Neumann and Morgenstern (1944). In such a game one player’s gain is the other player’s loss. Chess, Checkers, Rummy, Two finger Morra, and Tic-Tac-Toe are all examples of zero-sum two-person games. The theory for such games is surveyed in ► [Zero-Sum Two Person Games](#) (T.E.S. Raghavan). Recent results on stochastic zero-sum games are presented in ► [Stochastic Games](#) (Eilon Solan). Stochastic games are used to model dynamic interactions in which the environment changes in response to the behavior of the players. These games are discussed in Solan’s chapter.

Signaling games and Inspection games are also two-player games. Signaling games is the subject of ► [Signaling Games](#) (Joel Sobel). They are games of incomplete in-

formation in which one player is informed and the other is not. Players can use the actions of their opponents to make inferences about hidden information. The earliest work on this subject is Spence’s seminal 1972 work, in which education serves as a signal of ability. Inspection games is covered in ► [Inspection Games](#) (Rudolf Avenhaus and Morton J. Canty). These games deal with the problem faced by an *inspector* who is required to control the compliance of an *inspectee* to some legal or otherwise formal undertaking. They started with the analysis of arms control and disarmament problems in the early 1960s and have been applied to auditing, environmental control, material accountancy, etc.

Inspections cause conflict in many real world situations. In economics, there are services of many kinds the fulfillment or payment of which has to be verified. One example is the problem of principal-agent relationships discussed in detail in ► [Principal-Agent Models](#) (David Perez-Castrillo and Inez Macho-Stadler). The Principal-agent models provide the theory of contracts under asymmetric information, concerning relationships between owner and manager, insurer and insured, etc. The principal, e.g., an employer, delegates work or responsibility to the agent, the employee, and chooses a payment schedule that best exploits the agent’s self-interests. The agent, of course, behaves so as to maximize her own utility given the fee schedule proposed by the principal. The problem faced by the principal is to devise incentives to motivate the agent to act in the principal’s interest. This generates some type of transaction cost for the principal, which includes the task of investigating and selecting appropriate agents, gaining information to set performances standards, monitoring agents, bonding payments by the agents, and residual losses.

The Chap. ► [Differential Games](#) (Marc Quincampoix) is devoted to Differential games with focus on two-player zero sum and antagonist differential games. These are games in which the state of the players depends on time in a continuous way. The positions of the players are solutions to differential equations. Motivated by military applications in the “Cold War” these games have a wide range of applications from Economics to engineer sciences, and recently to biology and behavioral ecology.

Mechanism designed is the subject of ► [Mechanism Design](#) (Ron Lavi). It studies the construction of mechanisms that aim to reach a socially desirable outcome in the presence of rational but selfish players, who care only about their own private utility. More specifically, the question is how to design a mechanism such that the equilibrium behavior of the players in the game induced by the mechanism leads to the socially desired goal.

The theory of mechanism design has been contributed to the development of other research areas as for example auction theory, contract theory and two-sided matching theory. "For having laid the foundations of mechanism design theory" the 2007 Nobel prize in Economics was awarded to Leonid Hurwicz, Eric Maskin and Roger Myerson.

A related theory is the theory of Implementation, the subject of ► [Implementation Theory](#) (Luis Corchon). It reverses the usual procedure, namely, fix a mechanism and see what the outcomes are. More precisely, it investigates the correspondence between normative goals and mechanisms designed to achieve those goals.

A class of "mixed" games is that of two-sided matching games, that has been analyzed since Gale and Shapley, 1962, under both cooperative and non-cooperative game theoretic approaches. The two-sided matching theory is surveyed in ► [Two-Sided Matching Models](#) (Marilda Sotomayor and Ömer Özak) by focusing on the differences and similarities between some matching models. In their paper, Gale and Shapley formulated and solved the Stable Matching problem for the Marriage and the College Admissions markets. The solution of the College Admissions problem was given by a simple deferred-acceptance algorithm which has been adapted and applied in the reorganization of admission processes of many two-sided matching markets.

Another class of problems that have been discussed from the perspective of cooperative and non-cooperative game theory is the Cost sharing problems, treated in ► [Cost Sharing](#) (Maurice Koster). Applications are numerous ranging from environmental issues like pollution, fishing grounds, to sharing multipurpose reservoirs, road systems, communication networks, and the internet. The worth of a "coalition" of such activities is defined as the hypothetical cost of carrying out the activities in that coalition only.

Market games and clubs are treated in ► [Market Games and Clubs](#) (Myrna Wooders) with focus on the equivalence between markets – defined as private goods economies where all participants in the economy have utility functions that are linear in the variable money - and games in characteristic function form.

Learning in Games is surveyed in ► [Learning in Games](#) (John Nachbar). It covers models in which players are "rational" but not necessarily in equilibrium: players forecast, possibly inaccurately, the future behavior of their opponents and optimize or  $\varepsilon$  optimize with respect to their forecasts.

Fair Division is reviewed in ► [Fair Division](#) (Steven Brams). It provides a rigorous analysis of procedures for

allocating goods, or deciding who wins on what issues, in a dispute.

The following two chapters deal with applications to political sciences. The first one, the Chap. ► [Voting](#) (Alvaro Sandroni, Antonio Penta, Jonathan Pogach, Deniz Selman and Michela Tincani), presents a game theoretic analysis of voting systems, which are procedures to choose a winner among a set of candidates from the individual preferences of the voters or more ambitiously, allowing to rank all the candidates or a part of them. Such a situation occurs in the field of social choice and welfare, in the field of elections, and in many other fields as games, sports, artificial intelligence, spam detection, web search engines and more generally Internet applications, statistics, and soon. From a practical point of view, it is crucial to be able to announce who is the winner in a "reasonable" time. This raises the question of the complexity of the voting procedures. The second Chap. ► [Voting Procedures, Complexity of](#) (Olivier Hudry) details the complexity results about several voting procedures.

The Chap. ► [Evolutionary Game Theory](#) (William Sandholm) deals with applications to biology. This field, known as Evolutionary game theory, started in 1972 with the publication of a series of papers by the mathematical biologist John Maynard Smith. Maynard Smith adapted the methods of traditional game theory, which were created to model the behavior of rational economic agents, to the context of biological natural selection.

Network models have a long history in sociology, natural sciences and engineering. However, only recently economists have begun to think of political and economic interactions as network phenomena and to model everything as games of network formation. The Chap. ► [Networks and Stability](#) (Frank Page and Myrna Wooders) is devoted to stable Networks and the game theoretic underpinnings of stable networks.

The Chap. ► [Game Theory and Strategic Complexity](#) (Kalyan Chatterjee and Hamid Sabourian) deals with some aspect of bounded rationality that has generated important work, namely the presence of constraints on the capacities of players. Various constraints could be considered, for example, limits on the ability to plan ahead in intertemporal decision making or on the ability to compute best responses.

This chapter discusses cognitive costs to players of using strategies that depend on long histories of past play. This is done mainly in the context of bargaining and markets. It is shown that such complexity considerations often enable us to make sharp predictions. It is also considered the issue in the context of repeated games.

## Game Theory and Strategic Complexity

KALYAN CHATTERJEE<sup>1</sup>, HAMID SABOURIAN<sup>2</sup>

<sup>1</sup> Department of Economics, The Pennsylvania State University, University Park, USA

<sup>2</sup> Faculty of Economics, University of Cambridge, Cambridge, UK

### Article Outline

Glossary

Definition of the Subject

Introduction

Games, Automata and Equilibrium Concepts

Complexity Considerations in Repeated Games

Complexity and Bargaining

Complexity, Market Games and the Competitive

Equilibrium

Discussion and Future Directions

Acknowledgments

Bibliography

### Glossary

**Game theory** A formal model of interaction, usually in human behavior.

**Repeated games** A series of identical interactions of this kind.

**Strategy** A complete specification of how a player will play the game.

**Strategic complexity** A measure of how complex a strategy is to implement.

**Equilibrium** A solution concept for games in which each player optimizes given his correct prediction of others' behavior.

**Equilibrium path** The outcome in terms of the play of the game if every player uses his equilibrium strategy.

**Continuation game** A description of how the play will proceed in a dynamic game once some part of the game has already occurred.

**Automata** A formal definition of a strategy that captures its complexity.

### Definition of the Subject

The subject of this chapter is at the intersection of economics and computer science and deals with the use of measures of complexity obtained from the study of finite automata to help select among multiple equilibria and other outcomes appearing in game-theoretic models of bargaining, markets and repeated interactions. The impor-

tance of the topic lies in the ability of concepts that employ bounds on available resources to generate more refined predictions of individual behavior in markets.

### Introduction

This chapter is concerned with the concept of strategic complexity and its use in game theory. There are many different meanings associated with the word "complexity", as the variety of topics discussed in this volume makes clear. In this paper, we shall adopt a somewhat narrow view, confining ourselves to notions that measure, in some way, constraints on the ability of economic agents to behave with full rationality in their interactions with other agents in dynamic environments. This will be made more precise a little later. (A more general discussion is available in Rubinstein [45].)

Why is it important to study the effect of such constraints on economic decision-making? The first reason could be to increase the realism of the assumptions of economic models; it is evident from introspection and from observing others that we do not have infinite memory and cannot condition our future actions on the entire corpus of what we once knew; or, for that matter, unlimited computational power. However, only considering the assumptions of a model would not be considered enough if the increased realism were not to expand our ability to explain or to predict. The second reason therefore is that studying the effects of complexity on human decision-making might help us either to make our predictions more precise (by selecting among equilibria) or to generate explanations for behavior that is frequently observed, but incompatible with equilibrium in models that have stronger assumptions about the abilities of agents.

A strategy in a game is an entire plan of how to play the game at every possible history/contingency/eventuality at which the player has to make a move. The particular aspect of complexity that we shall focus on is on the complexity of strategy as a function of the history. One representation of the players' strategies in games is often in terms of (finite) automata. The finiteness need not always be assumed; it can be derived. The ideas of complexity, though often most conveniently represented this way, can also be discussed without referring to finite automata at all but purely in how a strategy depends on the past history of the game.

The number of states in the automaton can be used as a measure of complexity. This may be a natural measure of complexity in a stationary repetitive environment such as repeated games. We shall discuss this measure of complexity as well as other aspects of the complexity of

a strategy that are particularly relevant in non-stationary frameworks.

Note the players are not themselves considered automata in this paper and in the literature it surveys. Also, we do not place restrictions on the ability of players to compute strategies (see [39]), only on the strategies that they can implement. The paper is also not intended as a comprehensive survey of the literature on complexity of implementation in games. The main focus of the paper is inevitably on the works that we have been personally associated with.

The remaining part of this paper is organized as follows: In the next section, we discuss strategies in a game, their representation as finite automata and the basic equilibrium concepts to be used in the paper. Section “[Complexity Considerations in Repeated Games](#)” will consider the use of complexity notions in repeated games. Section “[Complexity and Bargaining](#)” will focus on extensive form bargaining and the effect of complexity considerations in selecting equilibria. Section “[Complexity, Market Games and the Competitive Equilibrium](#)” will extend the analysis of bargaining to markets in which several agents bargain and considers the recent literature that justifies competitive outcomes in market environments by appealing to the aversion of agents to complexity. Section “[Discussion and Future Directions](#)” concludes with some thoughts on future research. This paper draws on an earlier survey paper [14] for some of the material in Sects. “[Games, Automata and Equilibrium Concepts](#)”, “[Complexity Considerations in Repeated Games](#)” and “[Complexity and Bargaining](#)”.

### [Games, Automata and Equilibrium Concepts](#)

As mentioned in the introduction, this paper will be concerned with *dynamic games*. Though the theory of games has diffused from economics and mathematics to several other fields in the last few decades, we include an introduction to the basic concepts to keep this paper as self-contained as possible. A game is a formal model of interaction between individual agents. The basic components of a game are: (i) Players or agents, whose choices will, in general, have consequences for each other. We assume a finite set of players, denoted by  $N$ . We shall also use  $N$  sometimes to represent the cardinality of this set. (ii) A specification of the “rules of the game” or the structure of interaction, described by the sequence of possible events in the game, the order in which the players move, what they can choose at each move and what they know about previous moves. This is usually modeled as a tree and is called the “extensive form” of the game (and will not be formalized

here, though the formalization is standard and found in all the texts on the subject). (iii) Payoffs for each player associated with every path through the tree from the root. It is easier to describe this as a finite tree and ascribe payoffs to the end nodes  $z$ . Let  $u_i(z)$  be the real-valued payoff to Player  $i$  associated with end node  $z$ . The payoffs are usually assumed to satisfy conditions that are sufficient to guarantee that the utility of a probability distribution on a subset of the set of end nodes is the expectation of the utility of the individual end nodes. However, different strands of work on bounded rationality dispense with this assumption. The description above presupposes a tree of finite depth, whilst many of the applications deal with infinite horizon games. However, the definitions are easily modified by associating payoffs with a *play* of the game and defining a node as a set of plays. We shall not pursue this further here.

In the standard model of a game, players are assumed to have all orders of knowledge about the preceding description. Work on bounded rationality also has considered relaxing this assumption.

A strategy is a complete plan of action for playing a game, describing the course of action to be adopted in every possible contingency (or every information set of the player concerned). The plan has to be detailed enough so that it can be played by an agent, even if the principal is not himself or herself in town, and the agent could well be a computer, which is programmed to follow the strategy. Without any loss of generality, a strategy can be represented by an automaton (see below for illustration and Osborne and Rubinstein [38] for a formal treatment in the context of repeated games). Often such a machine description is more convenient in terms of accounting for a complexity of a machine. For example the works that are based on the use of finite automata or Turing machines to represent *strategies* for playing a game impose a natural bound on the set of allowable strategies.

For the types of problem that we shall consider here, it is best to think of a *multistage game with observable actions*, to use the terminology of Fudenberg and Tirole [20]. The game has some temporal structure; let us call each unit of time a period or a stage. In each period, the players choose actions simultaneously and independently. (The actions could include the dummy action.) All the actions taken in a stage are observed and the players then choose actions again. An example is a repeated normal form game, such as the famous Prisoners’ Dilemma being repeated infinitely or finitely often. In each stage, players choose whether to *cooperate* or *defect*. The choices are revealed, payoffs received and the choices repeated again and so on. (The reader will recall that in the Prisoners’ Dilemma played once, *Defect* is better than *Cooperate* for

each player, no matter what the other player does, but both players choosing *Defect* is strictly worse for each than both choosing *Cooperate*.) What a strategy for a given player would do would be to specify the choice in a given stage as a function of the history of the game up to that stage (for every stage). A finite automaton represents a particular strategy in the following way: It partitions all possible histories in the game (at which the player concerned has to move) using a finite number of elements. Each of these elements is a *state* of the machine. Given a state, the automaton prescribes an action (for example, *Cooperate* after all histories in which the other party has cooperated). It also specifies how the state of the machine will change as a result of the action taken by the other player. The state-to-action mapping is called the output mapping and the rule that prescribes the state in the next period as a function of today's state and the action of one's opponent in this period is called the transition mapping. The automaton also needs to prescribe what to do in the first stage, when there is no history of past actions to rely on. Thus for example, the famous 'tit-for-tat' strategy in the repeated Prisoners' Dilemma can be represented by the following automaton.

1. Play *Cooperate* in the first stage. The initial state is denoted as  $q^1$  and in this state the action prescribed is *cooperate*.
2. As long as the other player cooperates, stay in state  $q^1$ .
3. If the other player defects in a state, go to state  $q^2$ . The action specified in  $q^2$  is *Defect*.
4. Stay in  $q^2$  as long as the other player defects. If the other player cooperates in a stage, go to  $q^1$ .

Denoting the output mapping by  $\lambda(\cdot)$ , we get  $\lambda(q^1) = C$  and  $\lambda(q^2) = D$ . The transition mapping,  $\mu(\cdot, \cdot)$  is as follows:  $\mu(q^1, C) = q^1$ ,  $\mu(q^1, D) = q^2$ ,  $\mu(q^2, C) = q^1$ ,  $\mu(q^2, D) = q^2$ . Here, of course, *C* and *D* denote *cooperate* and *defect* respectively. The machine described above has two states and is an instance of a Moore machine in computer science terminology.

The use of a Moore machine to represent a strategy rules out strategies in which histories are arbitrarily finitely partitioned or arbitrarily complex. In fact, the number of states in the machine is a popular measure of the complexity of the machine and the strategy it represents.

Another kind of finite automaton used in the literature is a Mealy machine. The main difference between this and the Moore machine is that now the output is a function both of the state and of an input, unlike the Moore machine where it is only a function of the state. One can always transform a Mealy machine to a Moore machine by making transitions depend on the input and having state transitions after every input. The Mealy machine repre-

sentation is more convenient for the extensive form game we shall consider in Section "Complexity and Bargaining". We shall briefly address why in that section.

The aim of using the machine framework to describe strategies is to take into account explicitly the cost of complexity of strategies. There is the belief for instance that short-term memory (see [33]) is capable of keeping seven things in mind at any given time and if five of them are occupied by how to play the Prisoners' Dilemma there might be less left over for other important activities.

The standard equilibrium concept in game theory is the concept of Nash equilibrium. This requires each player to choose a best strategy (in terms of payoff) given his or her conjectures about other players' strategies and, of course, in equilibrium the conjectures must be correct. Thus, a Nash equilibrium is a profile of strategies, one for each player, such that every player is choosing a best response strategy given the Nash equilibrium strategies of the other players. In dynamic games Nash equilibrium strategies may not be credible (sequentially rational). In multi-stage games, to ensure credibility, the concept of Nash equilibrium is refined by requiring the strategy of each player to be a best response to the strategies of the others at every well-defined history (subgame) within the game. This notion of equilibrium was introduced by Seltten [48] and is called subgame perfect equilibrium. The difference between this concept and that of Nash, which it refines, is that players must specify strategies that are best responses to each other even at nodes in the game tree that would never be reached if the prescribed equilibrium were being played. The Nash concept does not require this. The notion of histories *off the equilibrium path* therefore refers to those that do not occur if every player follows his or her equilibrium strategy. Another useful concept to mention here is that of *payoff in the continuation game*. This refers to the expected payoff from the prescribed strategies in the part of the game remaining to be played after some moves have already taken place. The restriction of the prescribed strategies to the continuation game are referred to here as *continuation strategies*.

Rubinstein [44], Abreu and Rubinstein [1] and others have modified the standard equilibrium concepts to account for complexity costs. This approach is somewhat different from that adopted, for example, by Neyman [35], who restricted strategies to those of bounded complexity. We shall next present the Abreu–Rubinstein definition of *Nash equilibrium with complexity* (often referred to as NEC in the rest of the paper).

The basic idea is a very simple extension of Nash equilibrium. Complexity enters the utility function lexicographically. A player first calculates his or her best re-

sponse to the conjectured strategies of the other players. If there are alternative best responses, the player chooses the less complex one. Thus a Nash equilibrium with complexity has two aspects. First, the strategies chosen by any player must be a best response given his or her conjectures about other players' strategies and, of course, in equilibrium the conjectures must be correct. Second, there must not exist an alternative strategy for a player such that his or her payoff is the same as in the candidate equilibrium strategy, given what other players do, but the alternative strategy is less complex.

In Abreu and Rubinstein [1], the measure of complexity is the number of states in the Moore machine that represents the strategy. The second part of their equilibrium definition restricts the extent to which punishments can be used off the equilibrium path. For example, there is a famous strategy that, if used by all players, gives cooperation in the infinitely repeated Prisoners' Dilemma (for sufficiently high discount factors), namely the "grim" strategy. This strategy can be described by the following machine: Start with *Cooperate*. Play *Cooperate* as long as the other players all cooperate. If in the last period any player has used *Defect*, then switch to playing *Defect* for ever. (That is, never play *Cooperate* again, no matter what the other players do in succeeding periods.) This strategy profile (each player uses the grim strategy) gives an outcome path consisting solely of players cooperating. No one defects because from then until the end of time all the players will be punishing one another.

However, this strategy profile is not a Nash equilibrium with complexity; the grim strategy is a two-state machine in which one state (the one in which a player chooses *Defect*) is never used given that everyone else cooperates on the equilibrium path. Some player can do better, even if lexicographically, by switching to a one-state machine in which he or she cooperates no matter what. Thus even the weak lexicographic requirement has some bite.

Note that the complexity restriction we are considering is on the complexity of *implementation*, not the complexity of *computation*. We know that even a Turing machine, which has potentially infinite memory, might be unable to calculate best responses to all possible strategy profiles of other players in the game (see [2,8]).

To return to the question of defining equilibrium in the machine game, the Abreu–Rubinstein approach is described by them as "buying" states in the machine at the beginning of the game. The complexity cost is therefore a fixed cost per state used. Some recent papers have taken the fixed cost approach further by requiring NEC strategies to be credible. The idea is that players pay an initial fixed cost for the complexity (the no-

tion of complexity in some of these papers differ from counting the states approach) of his/her strategy and then the game is played with strategies being optimal at every contingency as in standard game theory. Chatterjee and Sabourian [15,16] model this by considering Nash equilibrium with complexity costs in (bargaining) games in which machines/strategies can make errors/trembles in output/action. The introduction of errors ensures that the equilibrium strategies are optimal after every history. As the error goes to zero, we are left with subgame perfect equilibria of the underlying game. Chatterjee and Sabourian [15], Sabourian [47], Gale and Sabourian [22] and Lee and Sabourian [31] take a more direct method of introducing credibility into the equilibrium concept with complexity costs by restricting NEC strategies to be subgame perfect equilibrium in the underlying game with no complexity costs. We refer to such an equilibria by perfect equilibrium with complexity costs (PEC).

In contrast to the fixed cost interpretation of complexity cost, Rubinstein in his 1986 paper considers a different approach, namely the choice of "renting" states in the machine for every period the game is played. Formally, the Rubinstein notion of *semi-perfect equilibrium* requires the strategy chosen to have the minimal number of states necessary to play the game at every node on the (candidate) equilibrium outcome path. A state could therefore be dropped if it is not going to be used on the candidate equilibrium path after some period. Thus, to be in the equilibrium machine, it is not sufficient that a state be used on the path, it has to be used in every possible future. Rubinstein called this notion of equilibrium *semi-perfect*, because the complexity of a strategy could be changed in one direction (it could be decreased) after every period. If states could be added as well as deleted every period, we would have yet another definition of equilibrium with complexity, *machine subgame perfect equilibrium*. (See [34].) In contrast, both the NEC and PEC concepts we use here entail a *single choice* of automaton or strategy by players at the beginning of the game.

In all these models, complexity analysis has been facilitated by considering the "machine games". Each player chooses among machines and the complexity of a machine is taken to be the number of states of the machine. In fact, the counting-the-number-of-states measure of complexity has an equivalent measure stated in terms of the underlying strategies that the machine could implement. Kalai and Stanford [29] define complexity of a strategy by the number of *continuation strategies* that the strategy induces at different periods/histories of the game, and establishes that such a measure is equal to the number of the states of the smallest machine that implements the strategy. Thus,

one could equivalently describe any result either in terms of underlying strategies and the cardinality of the set of continuation strategies that they induce or in terms of machines and the number of states in them. The same applies to other measures of complexity discussed in this paper; they can be defined either in terms of the machine specification or in terms of the underlying strategy. In the rest of this paper, to simplify the exposition we shall at times go from one exposition to the other without further explanation.

With this preamble on the concepts of equilibrium used in this literature, we turn to a discussion of a specific game in the next section, the infinitely repeated Prisoners' Dilemma. We will discuss mainly the approach of Abreu and Rubinstein in this section but contrast it with the literature following from Neyman. We also note that the suggestion for using finite automata in games of this kind came originally from Aumann [4].

### Complexity Considerations in Repeated Games

#### Endogenous Complexity

In this subsection we shall first concentrate on the Prisoners' Dilemma and discuss the work of Abreu and Rubinstein, which was introduced briefly in the last section. For concreteness, consider the following Prisoners' Dilemma payoffs:

	$C_2$	$D_2$
$C_1$	3, 3	-1, 4
$D_1$	4, -1	0, 0

This is the “stage game”; each of the two players chooses an action in each stage, their actions are revealed at the end of the stage and then the next stage begins. The game is repeated infinitely often and future payoffs are discounted with a common discount factor  $\delta$ .

The solution concept to be used was introduced in the last section; NEC or Nash equilibrium with complexity. Note that here complexity is *endogenous*. A player has a preference for less complex strategies. This preference comes into play lexicographically, that is for any strategies or machines that give the same payoff against the opponent's equilibrium strategy, a player will choose the one with lowest complexity. Thus the cost of complexity is infinitesimal. One could also consider positive but small costs of more complex strategies, but results will then depend on how large the cost of additional complexity is compared to the additional payoff obtained with a more complex strategy.

We saw in the last section that the “grim trigger” strategy, which is a two-state automaton, is not a NEC. The

reason is that if Player 2 uses such a strategy, Player 1 can be better off by deviating to a one-state strategy in which she always cooperates. (This will give the same payoff with a less complex strategy.) One-state strategies where both players cooperate clearly do not constitute NEC (deviating and choosing a one-state machine that always plays  $D$  is strictly better for a player). However, if both players use a one-state machine that always generates an action of  $D$ , this is a NEC.

The question obviously arises if the cooperative outcome in each stage can be sustained as a NEC and the preceding discussion makes clear that the answer is no. Punishments have to be used on the equilibrium path, but we can get arbitrarily close to the cooperative outcome for a high enough discount factor. For example consider the following two-state machine:

$$\begin{aligned} Q &= \{q^1, q^2\}; \lambda(q^1) = D, \lambda(q^2) = C, \\ \mu(q^1, C) &= q^1, \mu(q^1, D) = q^2, \mu(q^2, D) = q^1, \\ &\quad \mu(q^2, C) = q^2. \end{aligned}$$

Here both players play the same strategy, which starts out playing  $D$ . If both players do as they are supposed to, each plays  $C$  in the next period and thereafter, so the sequence of actions is  $(D, D), (C, C), (C, C) \dots$  If either player plays  $C$  in the first period, the other player keeps playing  $D$  in the next period. The transition rule prescribes that if one plays  $C$  and one's opponent plays  $D$ , one goes back to playing  $D$ , so the sequence with the deviation will be  $(D, C), (D, D), (C, C), (C, C) \dots$

Suppose both players use this machine. First, we check it is a Nash equilibrium in payoffs. We only need to check what happens when a player plays  $C$ . If Player 2 deviates and plays  $D$ , she will get an immediate payoff of 4 followed by payoffs of 0, 3, 3, ... if she thereafter sticks to her strategy for a total payoff of  $4 + \delta^2(3/(1-\delta))$  as opposed to  $3/(1-\delta)$  if she had not deviated. The net gain from deviation is  $1 - 3\delta$ , which is negative for  $\delta > \frac{1}{3}$ . One can check that more complicated deviations are also worse. The second part of the definition needs to be checked as well, so we need to ensure that a player cannot do as well in terms of payoff by moving to a less complex strategy, namely a one-state machine. A one-state machine that always plays  $C$  will get the worst possible payoff, since the other machine will keep playing  $D$  against it. A one-state machine that plays  $D$  will get a payoff of 4 in periods 2, 4, ... or a total payoff of  $4\delta/(1-\delta^2)$  as against  $3\delta/(1-\delta)$ . The second is strictly greater for  $\delta > \frac{1}{3}$ .

This machine gives a payoff close to 3 per stage for  $\delta$  close to 1. As  $\delta \rightarrow 1$ , the payoff of each player goes to 3, the cooperative outcome.

The paper by Abreu and Rubinstein obtains a basic result on the characterization of payoffs obtained as NEC in the infinitely repeated Prisoners' Dilemma. We recall that the "Folk Theorem" for repeated games tells us that all outcome paths that give a payoff per stage strictly greater for each player than the minmax payoff for that player in the stage game can be sustained by Nash equilibrium strategies. Using endogenous complexity, one can obtain a refinement; now only payoffs on a so-called "cross" are sustainable as NEC. This result is obtained from two observations. First, in any NEC of a two-player game, the number of states in the players' machines must be equal. This follows from the following intuitive reasoning (we refer readers to the original paper for the proofs). Suppose we fix the machine used by one of the players (say Player 1), so that to the other player it becomes part of the "environment". For Player 2 to calculate a best response or an optimal strategy to Player 1's given machine, it is clearly not necessary to partition past histories more finely than the other player has done in obtaining her strategy; therefore the number of states in Player 2's machine need not (and therefore will not, if there are complexity costs) exceed the number in Player 1's machine in equilibrium. The same holds true in the other direction, so the number of states must be equal. (This does not hold for more than two players.) Another way of interpreting this result is that it restates the result from Markov decision processes on the existence of an optimal "stationary" policy (that is depending only on the states of the environment, which are here the same as the states of the other player's machine). See also Piccione [40].

Thus there is a one-to-one correspondence between the states of the two machines. (Since the number of states is finite and the game is infinitely repeated, the machine must visit at least one of the states infinitely often for each player.) One can strengthen this further to establish a one-to-one correspondence between *actions*. Suppose Player 1's machine has  $a_t^1 = a_s^1$ , where these denote the actions taken at two distinct periods and states by Player 1, with  $a_t^2 \neq a_s^2$  for Player 2. Since the states in  $t$  and  $s$  are distinct for Player 1 and the actions taken are the same, the transitions must be different following the two distinct states. But then Player 1 does not need two distinct states, he can drop one and condition the transition after, say,  $s$  on the different action used by Player 2. (Recall the transition is a function of the state and the opponent's action.) But then Player 1 would be able to obtain the same payoff with a less complex machine; so the original one could not have been a NEC machine.

Therefore the actions played must be some combination of  $(C, C)$  and  $(D, D)$  (the correspondence is between

the two  $C$ s and the two  $D$ s) or some combination of  $(C, D)$  and  $(D, C)$ . (By combination, we mean combination over time. For example,  $\{C, C\}$  is played, say, 10 times for every 3 plays of  $(D, D)$ ). In the payoff space, sustainable payoffs are either on the line joining  $(3,3)$  and  $(0,0)$  or on the line joining the payoffs on the other diagonal; hence the evocative name chosen to describe the result—the cross of the two diagonals.

While this is certainly a selection of equilibrium outcomes, it does not go as far as we would wish. We would hope that some equilibrium selection argument might deliver us the co-operative outcome  $(3,3)$  uniquely (even in the limit as  $\delta \rightarrow 1$ ), instead of the actual result obtained. There is work that does this, but it uses evolutionary arguments for equilibrium selection (see Binmore and Samuelson [9]). An alternative learning argument for equilibrium selection is used by Maenner [32]. In his model, a player tries to infer what machine is being used by his opponent and chooses the simplest automaton that is consistent with the observed pattern of play as his model of his opponent. A player then chooses a best response to this inference. It turns out complexity is not sufficient to pin down an inference and one must use optimistic or pessimistic rules to select among the simplest inferences. One of these gives only  $(D, D)$  repeated, whilst the other reproduces the Abreu–Rubinstein NEC results. Piccione and Rubinstein [41] show that the NEC profile of 2-player repeated *extensive form games* is unique if the stage game is one of perfect information. This unique equilibrium involves all players playing their one-shot myopic non-cooperative actions at every stage. This is a strong selection result and involves stage game strategies not being observable (only the path of play is) as well as the result on the equilibrium numbers of states being equal in the two players' machines.

In repeated games with more than two players or with more than two actions at each stage the multiplicity problem may be more acute than just not being able to select uniquely a "cooperative outcome". In some such games complexity by itself may not have any bite and the Folk Theorem may survive even when the players care for the complexity of their strategies. (See Bloise [12] who shows robust examples of two-player repeated games with three actions at each stage such that every individually rational payoff can be sustained as a NEC if players are sufficiently patient.)

### Exogenous Complexity

We now consider the different approach taken by Neyman [35,36], Ben Porath [6,7], Zemel [50] and others.

We shall confine ourselves to the papers by Neyman and Zemel on the Prisoners' Dilemma, without discussing the more general results these authors and others have obtained.

Neyman's approach treats complexity as exogenous complexity. Let Player  $i$  be restricted to use strategies/automata with the number of states not to exceed  $m_i$ . He also considers finitely repeated games, unlike the infinitely repeated games we have discussed up to this point. With the stage game being the Prisoners' Dilemma and the number of repetitions being  $T$  (for convenience, this includes the first time the game is played). We can write the game being considered as  $G^T(m_1, m_2)$ . Note that without the complexity restrictions, the finitely repeated Prisoners' Dilemma has a unique Nash equilibrium outcome path (and a unique subgame perfect equilibrium)– $(D, D)$  in all stages. Thus sustaining cooperation in this setting is obtaining non-equilibrium behavior, though one that is frequently observed in real life. This approach therefore is an example of bounded rationality being used to explain observed behavior that is not predicted in equilibrium.

If the complexity restrictions are severe, it turns out that  $(C, C)$  in each period is an equilibrium. For this, we need  $2 \leq m_1, m_2 \leq T - 1$ . To see this consider the grim trigger strategy mentioned earlier-representable as a two-state automaton- and let  $T = 3$ . Here  $\lambda(q^1) = C$ ;  $\lambda(q^2) = D$ ;  $\mu(q^1, C) = q^1$ ;  $\mu(q^1, D) = q^2$ ;  $\mu(q^2, C \text{ or } D) = q^2$ . If each player uses this strategy,  $(C, C)$  will be observed. Such a pair of strategies is clearly not a Nash equilibrium-given Player 1's strategy, Player 2 can do better by playing  $D$  in stage 3. But if Player 2 defects in the second stage, by choosing a two-state machine where  $\mu(q^1, C) = D$ , he will gain 1 in the second stage and lose 3 in the third stage as compared to the machine listed above, so he is worse off. But defecting in stage 3 requires an automaton with three states-two states in which  $C$  is played and one in which  $D$  is played. The transitions in state  $q^1$  will be similar but, if  $q^2$  is the second cooperative state, the transition from  $q^2$  to the defect state will take place no matter whether the other player plays  $C$  or  $D$ . However, automata with three states violate the constraint that the number of states be no more than 2, so the profitable deviation is out of reach.

Whilst this is easy to see, it is not clear what happens when the complexity is high. Neyman shows the following result: *For any integer  $k$ , there exists a  $T_0$ , such that for  $T \geq T_0$  and  $T^{\frac{1}{k}} \leq m_1, m_2 \leq T^k$ , there is a mixed strategy equilibrium of  $G^T(m_1, m_2)$  in which the expected average payoff to each player is at least  $3 - \frac{1}{k}$ .*

The basic idea is that rather than playing  $(C, C)$  at each stage, players are required to play a complex sequence of  $C$

and  $D$  and keeping track of this sequence uses up a sufficient number of states in the automaton so that profitable deviations again hit the constraint on the number of states. But since  $D$  cannot be avoided on the equilibrium path, only something close to  $(C, C)$  each period can be obtained rather than  $(C, C)$  all the time.

Zemel's paper adds a clever little twist to this argument by introducing communication. In his game, there are two actions each player chooses at each stage, either  $C$  or  $D$  as before and a message to be communicated. The message does not directly affect payoffs as the choice of  $C$  or  $D$  does. The communication requirements are now made sufficiently stringent, and deviation from them is considered a deviation, so that once again the states "left over" to count up to  $N$  are inadequate in number and  $(C, C)$  can once again be played in each stage/period. This is an interesting explanation of the rigid "scripts" that many have observed to be followed, for example, in negotiations.

Neyman [36] surveys his own work and that of Ben Porath [6,7]. He also generalizes his earlier work on the finitely-repeated Prisoners' Dilemma to show how small the complexity bounds would have to be in order to obtain outcomes outside the set of (unconstrained) equilibrium payoffs in the finitely-repeated, normal-form game (just as  $(C, C)$  is not part of an unconstrained equilibrium outcome path in the Prisoners' Dilemma). Essentially, if the complexity permitted grows exponentially or faster with the number of repetitions, the equilibrium payoff sets of the constrained and the unconstrained games will coincide. For sub-exponential growth, a version of the Folk-theorem is proved for two-person games. The first result says:

For every game  $G$  in strategic form and with  $m_i$  being the bound on the complexity of  $i$ 's strategy and  $T$  the number of times the game  $G$  is played, there exists a constant  $c$  such that if  $m_i \geq \exp(cT)$ , then  $E(G^T) = E(G^T(m_1, m_2))$  where  $E(\cdot)$  is the set of equilibrium payoffs in the game concerned.

The second result, which generalizes the Prisoners' Dilemma result already stated, considers a sequence of triples  $(m_1(n), m_2(n), T(n))$  for a two-player strategic form game, with  $m_2 \geq m_1$  and shows that the  $\liminf$  of the set of equilibrium payoffs of the automata game as  $n \rightarrow \infty$  includes essentially the strictly individually rational payoffs of the stage game if  $m_1(n) \rightarrow \infty$  and  $(\log m_1(n))/T(n) \rightarrow 0$  as  $n \rightarrow \infty$ . Thus a version of the Folk theorem holds provided the complexity of the players' machines does not grow too fast with the number of repetitions.

## Complexity and Bargaining

### Complexity and the Unanimity Game

The well-known alternating offers bargaining model of Rubinstein has two players alternating in making proposals and responding to proposals. Each period or unit of time consists of one proposal and one response. If the response is “reject”, the player who rejects makes the next proposal but in the following period. Since there is discounting with discount factor  $\delta$  per period, a rejection has a cost. The unanimity game we consider is a multiperson generalization of this bargaining game, with  $n$  players arranged in a fixed order, say  $1, 2, 3 \dots n$ . Player 1 makes a proposal on how to divide a pie of size unity among the  $n$  people; players  $2, 3, \dots n$  respond sequentially, either accepting or rejecting. If everyone accepts, the game ends. If someone rejects, Player 2 now gets to make a proposal but in the next period. The responses to Player 2’s proposal are made sequentially by Players  $3, 4, 5 \dots n, 1$ . If Player  $i$  gets a share  $x_i$  in an eventual agreement at time  $t$ , his payoff is  $\delta^{t-1}x_i$ .

Avner Shaked had shown in 1986 that the unanimity game had the disturbing feature that all individually rational (that is non-negative payoffs for each player) outcomes could be supported as subgame perfect equilibria. Thus the sharp result of Rubinstein [43], who found a unique subgame perfect equilibrium in the two-play stood in complete contrast with the multiplicity of subgame perfect equilibria in the multiplayer game.

Shaked’s proof had involved complex changes in expectations of the players if a deviation from the candidate equilibrium were to be observed. For example, in the three-player game with common discount factor  $\delta$ , the three extreme points  $(1, 0, 0)$ ,  $(0, 1, 0)$ ,  $(0, 0, 1)$  sustain one another in the following way. Suppose Player 1 is to propose  $(0, 1, 0)$ , which is not a very sensible offer for him or her to propose, since it gives everything to the second player. If Player 1 deviates and proposes, say,  $((1 - \delta)/2, \delta, (1 - \delta)/2)$ , then it might be reasoned that Player 2 would have no incentive to reject because in any case he or she can’t get more than 1 in the following period and Player 3 would surely prefer a positive payoff to 0. However, there is a counter-argument. In the subgame following Player 1’s deviation, Player 3’s expectations have been raised so that he (and everyone else, including Player 1) now expect the outcome to be  $(0, 0, 1)$ , instead of the earlier expected outcome. For sufficiently high discount factor, Player 3 would reject Player 1’s insufficiently generous offer. Thus Player 1 would have no incentive to deviate. Player 1 is thus in a bind; if he offers Player 2 less than  $\delta$  and offers Player 3 more in the devi-

ation, the expectation that the outcome next period will be  $(0, 1, 0)$  remains unchanged, so now Player 2 rejects his offer. So no deviation is profitable, because each deviation generates an expectation of future outcomes, an expectation that is confirmed in equilibrium. (This is what equilibrium means.) Summarizing,  $(0, 1, 0)$  is sustained as follows: Player 1 offers  $(0, 1, 0)$ , Player 2 accepts any offer of at least 1 and Player 3 any offer of at least 0. If one of them rejects Player 1’s offer, the next player in order offers  $(0, 1, 0)$  and the others accept. If any proposer, say Player 1, deviates from the offer  $(0, 1, 0)$  to  $(x_1, x_2, x_3)$  the player with the lower of  $\{x_2, x_3\}$  rejects. Suppose it is Player  $i$  who rejects. In the following period, the offer made gives 1 to Player  $i$  and 0 to the others, and this is accepted.

Various attempts were made to get around the continuum of equilibria problem in bargaining games with more than two players; most of them involved changing the game. (See [15,16] for a discussion of this literature.) An alternative to changing the game might be to introduce a cost for this additional complexity, in the belief that players who value simplicity will end up choosing simple, that is history independent, strategies. This seems to be a promising approach because it is clear from Shaked’s construction that the large number of equilibria results from the players choosing strategies that are history-dependent. In fact, if the strategies are restricted to those that are history-independent (also referred to as *stationary* or *Markov*) then it can be shown (see Herrero [27]) that the subgame perfect equilibrium is unique and induces equal division of the pie as  $\delta \rightarrow 1$ .

The two papers ([15,16]) in fact seek to address the issue of complex strategies with players having a preference for simplicity, just as in Abreu and Rubinstein. However, now we have a game of more than two players, and a single extensive form game rather than a repeated game as in Abreu–Rubinstein. It was natural that the framework had to be broadened somewhat to take this into account.

For each of  $n$  players playing the unanimity game, we define a machine or an implementation of the strategy as follows.

A *stage* of the game is defined to be  $n$  periods, such that if a stage were to be completed, each player would play each role at most once. A *role* could be as proposer or  $n - 1$ th responder or  $n - 2$ th responder ... up to first responder (the last role would occur in the period before the player concerned had to make another proposal). An *outcome* of a stage is defined as a sequence of offers and responses, for example  $e = (x, A, A, R; y, R; z, A, R; b, A, A, A)$  in a four-player game where the  $(x, y, z, b)$  are proposals made in the four periods and  $(A, R)$  refer to accept and reject respectively. From

the point of view of the first player to propose (for convenience, let's call him Player 1), he makes an offer  $x$ , which is accepted by Players 2 and 3 but rejected by Player 4. Now it is Player 2's turn to offer, but this offer,  $y$ , is rejected by the first responder Player 3. Player 1 gets to play as second responder in the next period, where he rejects Player 3's proposal. In the last period of this stage, a proposal  $b$  is made by Player 4 and everyone accepts (including Player 1 as first responder). Any partial history within a stage is denoted by  $s$ . For example, when Player 2 makes an offer, he does so after a partial history  $s = (x, A, A, R)$ . Let the set of possible outcomes of a stage be denoted by  $E$  and the set of possible partial histories by  $S$ . Let  $Q_i$  denote the set of states used in the  $i$ th player's machine  $M_i$ . The output mapping is given by  $\lambda_i: S \times Q_i \rightarrow \Lambda$ , where  $\Lambda$  is the set of possible actions (that is the set of possible proposals, plus accept or reject). The transition between states now takes place at the end of each stage, so the transition mapping is given as  $\mu_i: E \times Q_i \rightarrow Q_i$ . As before, in the Abreu–Rubinstein setup, there is an initial state  $q_{\text{initial},i}$  specified for each player. There is also a termination state  $F$ , which is supposed to indicate agreement. Once in the termination state, players will play the null action and make transitions to this state.

Note that our formulation of a strategy naturally uses a Mealy machine. The output mapping  $\lambda_i(., .)$  has two arguments, the state of the machine and the *input*  $s$ , which lists the outcomes of previous moves within the stage. The transitions take place at the end of the stage. The benefit of using this formulation is that the continuation game is the same at the beginning of each stage. In Chatterjee and Sabourian [16], we investigate the effects of modifying this formulation, including studying the effects of having a sub-machine to play each role. The different formulations can all implement the same strategies, but the complexities in terms of various measures could differ. We refer the reader to that paper for details, but emphasize that in the general unanimity game, the results from other formulations are similar to the one developed here, though they could differ for special cases, like three-player games.

We now consider a machine game, where players first choose machines and then the machines play the unanimity game in analogy with Abreu–Rubinstein. Using the same lexicographic utility, with complexity coming after bargaining payoffs, what do we find for Nash equilibria of the machine game?

As it turns out, the addition of complexity costs in this setting has some bite but not much. In particular, any division of the pie can be sustained in some Nash equilibrium of the machine game. Perpetual disagreement can, in fact, be sustained by a stationary machine, that is one

that makes the same offers and responses each time, irrespective of past history. Nor can we prove, for general  $n$ -player games that the equilibrium machines will be one-state. (A three-player counter-example exists in [16]; it does not appear to be possible to generate in games that lasted less than thirty periods.) For two-player games, the result that machines must be one-state in equilibrium can be shown neatly ([16]); another illustration that in this particular area, there is a substantial increase of analytical difficulty in going from two to three players.

One reason why complexity does not appear important here is that the definition of complexity used is too restrictive. Counting the number of states is fine, so long as we don't consider how complex a response might be for partial histories *within* a stage. The next attempt at a solution is based on this observation.

We devise the following definition of complexity: Given the machine and the states, if a machine made the same response to different partial stage histories in different states and another machine made different responses, then the second one was more complex (given that the machines were identical in all other respects). We refer to this notion as *response complexity*. (In [15] the concept of response complexity is in fact stated in terms of the underlying strategy rather than in terms of machines.) It captures the intuition that counting states is not enough; two machines could have the same number of states, for example because each generated the same number of distinct offers, but the complexity of responses in one machine could be much lower than that in the other. Note that this notion would only arise in extensive-form games. In normal form games, counting states could be an adequate measure of complexity. Nor is this notion of complexity derivable from notions of transition complexity, due to Banks and Sundaram, for example, which also apply in normal-form games.

The main result of Chatterjee and Sabourian [15] is that this new aspect of complexity enables us to limit the amount of delay that can occur in equilibrium and hence to infer that only one-state machines are equilibrium machines.

The formal proofs using two different approaches are available in Chatterjee and Sabourian [15,16]. We mention the basic intuition behind these results. Suppose, in the three player game, there is an agreement in period 4 (this is in the second stage). Why doesn't this agreement take place in period 1 instead? It must be because if the same offer and responses are seen in period 1 some player will reject the offer. But of course, he or she does not have to do so because the required offer never happens. But a strategy that accepts the offer in period 4 and rejects it

*off the equilibrium path in period 1 must be more complex, by our definition,* than one that always accepts it whenever it might happen, on or off the expected path. Repeated application of this argument by backwards induction gives the result. (The details are more complicated but are in the papers cited above.) Note that this uses the definition that two machines might have the same number of states and yet one could be simpler than the other. It is interesting, as mentioned earlier, that for two players one can obtain an analogous result without invoking the response simplicity criterion, but from three players on this criterion is essential.

The above result (equilibrium machines have one state each and there are no delays beyond the first stage) is still not enough to refine the set of equilibria to a single allocation. In order to do this, we consider machines that can make errors/trembles in output. As the error goes to zero, we are left with perfect equilibria of our game. With one-state machines, the only subgame perfect equilibria are the ones that give equal division of the pie as  $\delta \rightarrow 1$ . Thus a combination of two techniques, one essentially recognizing that players can make mistakes and the other that players prefer simpler strategies if the payoffs are the same as those given by a more complex strategy, resolves the problem of multiplicity of equilibria in the multiperson bargaining game.

As we mentioned before, the introduction of errors ensures that the equilibrium strategies are credible at every history. We could also take the more direct (and easier) way of obtaining the uniqueness result with complexity costs by considering NEC strategies that are subgame perfect in the underlying game (PEC) (as done in [15]). Then since an history-independent subgame perfect equilibrium of the game is unique and any NEC automaton profile has one state and hence is history-independent, it follows immediately that any PEC is unique and induces equal division as  $\delta \rightarrow 1$ .

### Complexity and Repeated Negotiations

In addition to standard repeated games or standard bargaining games, multiplicity of equilibria often appear in dynamic repeated interactions, where a repeated game is superimposed on an alternating offers bargaining game. For instance, consider two firms, in an ongoing vertical relationship, negotiating the terms of a merger. Such situations have been analyzed in several “negotiation models” by Busch and Wen [13], Fernandez and Glazer [18] and Haller and Holden [25]. These models can be interpreted as combining the features of both repeated and alternating-offers bargaining games. In each period, one of

the two players first makes an offer on how to divide the total available periodic (flow) surplus; if the offer is accepted, the game ends with the players obtaining the corresponding payoffs in the current and every period thereafter. If the offer is rejected, they play some normal form game to determine their flow payoffs for that period and then the game moves on to the next period in which the same play continues with the players’ bargaining roles reversed. One can think of the normal form game played in the event of a rejection as a “threat game” in which a player takes actions that could punish the other player by reducing his total payoffs.

If the bargaining had not existed, the game would be a standard repeated normal form game. Introducing bargaining and the prospect of permanent exit, the negotiation model still admits a large number of equilibria, like standard repeated games. Some of these equilibria involve delay in agreement (even perpetual disagreement) and inefficiency, while some are efficient.

Lee and Sabourian [31] apply complexity considerations to this model. As in Abreu and Rubinstein [1] and others, the players choose among automata and the equilibrium notion is that of NEC and PEC. One important difference however is that in this paper the authors do not assume the automata to be *finite*. Also, the paper introduces a new machine specification that formally distinguishes between the two *roles* – proposer and responder – played by each player in a given period.

Complexity considerations select only *efficient* equilibria in the negotiation model players are sufficiently patient. First, it is shown that if an agreement occurs in some finite period as a NEC outcome then it must occur within the first two periods of the game. This is because if a NEC induces an agreement beyond the first two periods then one of the players must be able to drop the last period’s state of his machine without affecting the outcome of the game. Second, given sufficiently patient players, every PEC in the negotiation model that induces perpetual disagreement is at least *long-run* almost efficient; that is, the game must reach a finite date at which the continuation game then on is almost efficient.

Thus, these results take the study of complexity in repeated games a step further from the previous literature in which complexity or bargaining alone has produced only limited selection results. While, as we discussed above, many inefficient equilibria survive complexity refinement, Lee and Sabourian [31] demonstrate that complexity and bargaining in tandem ensure efficiency in repeated interactions. Complexity considerations also allow Lee and Sabourian to highlight the role of transaction costs in the negotiation game. Transaction costs take the form of pay-

ing a cost to enter the bargaining stage of the negotiation game. In contrast to the efficiency result in the negotiation game with complexity costs, Lee and Sabourian also show that introducing transaction costs into the negotiation game dramatically alters the selection result from efficiency to inefficiency. In particular, they show that, for any discount factor and any transaction cost, every PEC in the costly negotiation game induces perpetual disagreement if the stage game normal form (after any disagreement) has a unique Nash equilibrium.

### Complexity, Market Games and the Competitive Equilibrium

There has been a long tradition in economics of trying to provide a theory of how a competitive market with many buyers and sellers operates. The concept of competitive (Walrasian) equilibrium (see Debreu [17]) is a simple description of such markets. In such an equilibrium each trader chooses rationally the amount he wants to trade taking the prices as given, and the prices are set (or adjust) to ensure that total demanded is equal to the total supplied. The important feature of the set-up is that agents assume that they cannot influence (set) the prices and this is often justified by appealing to the idea that each individual agent is small relative to the market.

There are conceptual as well as technical problems associated with such a justification. First, if no agent can influence the prices then who sets them? Second, even in a large but finite market a change in the behavior of a single individual agent may affect the decisions of some others, which in turn might influence the behavior of some other agents and so on and so forth; thus the market as a whole may end up being affected by the decision of a single individual.

Game theoretic analysis of markets have tried to address these issues (e.g. see [21,47]). This has turned out to be a difficult task because the strategic analysis of markets, in contrast to the simple and elegant model of competitive equilibrium, tends to be complex and intractable. In particular, dynamic market games have many equilibria, in which a variety of different kinds of behavior are sustained by threats and counter-threats.

More than 60 years ago Hayek [26] noted the competitive markets are simple mechanisms in which economic agents only need to know their own endowments, preferences and technologies and the vector of prices at which trade takes place. In such environments, economic agents maximizing utility subject to constraints make efficient choices in equilibrium. Below we report some recent work, which suggests that the converse might also be true:

If rational agents have, at least at the margin, an aversion to complex behavior, then their maximizing behavior will result in simple behavioral rules and thereby in a perfectly competitive equilibrium

(Gale and Sabourian [22]).

### Homogeneous Markets

In a seminal paper, Rubinstein and Wolinsky [46], henceforth RW, considered a market for a single indivisible good in which a finite number of homogeneous buyers and homogeneous sellers are matched in pairs and bargain over the terms of trade. In their set-up, each seller has one unit of an indivisible good and each buyer wants to buy at most one unit of the good. Each seller's valuation of the good is 0 and each buyer's valuation is 1. Time is divided into discrete periods and at each date, buyers and sellers are matched randomly in pairs and one member of the pair is randomly chosen to be the proposer and the other the responder. In any such match the proposer offers a price  $p \in [0, 1]$  and the responder accepts or rejects the offer. If the offer is accepted the two agents trade at the agreed price  $p$  and the game ends with the seller receiving a payoff  $p$  and the buyer in the trade obtaining a payoff  $1 - p$ . If the offer is rejected the pair return to the market and the process continues. RW further assume that there is no discounting to capture the idea that there is no friction (cost to waiting) in the market.

Assuming that the number of buyers and sellers is not the same, RW showed that this dynamic matching and bargaining game has, in addition to a perfectly competitive outcome, a large set of other subgame perfect equilibrium outcomes, a result reminiscent of the Folk Theorem for repeated games. To see the intuition for this, consider the case in which there is one seller  $s$  and many buyers. Since there are more buyers than sellers the price of 1, at which the seller receives all the surplus, is the unique competitive equilibrium; furthermore, since there are no frictions  $p = 1$  seems to be the most plausible price. RW's precise result, however, establishes that for any price  $p^* \in [0, 1]$  and any buyer  $b^*$  there is a subgame perfect equilibrium that results in  $s$  and  $b^*$  trading at  $p^*$ . The idea behind the result is to construct an equilibrium strategy profile such that buyer  $b^*$  is identified as the intended recipient of the good at a price  $p^*$ . This means that the strategies are such that (i) when  $s$  meets  $b^*$ , whichever is chosen as the proposer offers price  $p^*$  and the responder accepts, (ii) when  $s$  is the proposer in a match with some buyer  $b \neq b^*$ ,  $s$  offers the good at a price of  $p = 1$  and  $b$  rejects and (iii) when a buyer  $b \neq b^*$  is the proposer he offers to buy the good at a price of  $p = 0$  and  $s$  rejects. These strategies pro-

duce the required outcome. Furthermore, the equilibrium strategies make use of the following punishment strategies to deter deviations. If the seller  $s$  deviates by proposing to a buyer  $b$  a price  $p \neq p^*$ ,  $b$  rejects this offer and the play continues with  $b$  becoming the intended recipient of the item at a price of zero. Thus, after rejection by  $b$  strategies are the same as those given earlier with the price zero in place of  $p^*$  and buyer  $b$  in place of buyer  $b^*$ . Similarly, if a buyer  $b$  deviates by offering a price  $p \neq p^*$  then the seller rejects, another buyer  $b' \neq b$  is chosen to be the intended recipient and the price at which the unit is traded changes to 1. Further deviations from these punishment strategies can be treated in an exactly similar way.

The strong impression left by RW is that indeterminacy of equilibrium is a robust feature of dynamic market games and, in particular, there is no reason to expect the outcome to be perfectly competitive. However, the strategies required to support the family of equilibria in RW are quite complex. In particular, when a proposer deviates, the strategies are tailor-made so that the responder is rewarded for rejecting the deviating proposal. This requires coordinating on a large amount of information so that at every information set the players know (and agree) what constitutes a deviation.

In fact, RW show that if the amount of information available to the agents is strictly limited so that the agents do not recall the history of past play then the only equilibrium outcome is the competitive one. This suggests that the competitive outcome may result if agents use simple strategies. Furthermore, the equilibrium strategies used described in RW to support non-competitive outcomes are particularly unattractive because they require all players, including those buyers who do not end up trading, to follow complex non-stationary strategies in order to support a non-competitive outcome. But buyers who do not trade and receive zero payoff on the equilibrium path could always obtain at least zero by following a less complex strategy than the ones specified in RW's construction. Thus, RW's construction of non-competitive equilibria is not robust if players prefer, at least at the margin, a simpler strategy to a more complex one.

Following the above observation, Sabourian [47], henceforth S, addresses the role of complexity (simplicity) in sustaining a multiplicity of non-competitive equilibria in RW's model. The concept of complexity in S is similar to that in Chatterjee and Sabourian [15]. It is defined by a partial ordering on the set of individual strategies (or automata) that very informally satisfies the following: If two strategies are otherwise identical except that in some role the second strategy uses more information than that available in the current period of bargaining and the first uses

only the information available in the current period, then the second strategy is said to be more complex than the first. S also introduces complexity costs lexicographically into the RW game and shows that any PEC is history-independent and induces the competitive outcome in the sense that all trades take place at the unique competitive price of 1.

Informally, S's conclusions *in the case of a single seller* and many buyers follows from the following three steps. First, since trading at the competitive price of 1 is the worst outcome for a buyer and the best outcome for the seller, by appealing to complexity type reasoning it can be shown that in any NEC a trader's response to a price offer of 1 is always history-independent and thus he either always rejects 1 or always accepts 1. For example, if in the case of a buyer this were not the case, then since accepting 1 is a worst possible outcome, he could economize on complexity and obtain at least the same payoff by adopting another strategy that is otherwise the same as the equilibrium strategy except that it always rejects 1.

Second, in any non-competitive NEC in which  $s$  receives a payoff of less than 1, there cannot be an agreement at a price of 1 between  $s$  and a buyer at *any* history. For example, if at some history, a buyer is offered  $p = 1$  and he accepts then by the first step the buyer should accept  $p = 1$  whenever it is offered; but this is a contradiction because it means that the seller can guarantee himself an equilibrium payoff of one by waiting until he has a chance to make a proposal to this buyer.

Third, in any non-competitive PEC the continuation payoffs of all buyers are positive at every history. This follows immediately from the previous step because if there is no trade at  $p = 1$  at any history it follows that each buyer can always obtain a positive payoff by offering the seller more than he can obtain in any subgame.

Finally, because of competition between the buyers (there is one seller and many buyers), in any subgame perfect equilibrium there must be a buyer with a zero continuation payoff after some history. To illustrate the basic intuition for this claim, let  $m$  be the worst continuation payoff for  $s$  at any history and suppose that there exists a subgame at which  $s$  is the proposer in a match with a buyer  $\bar{b}$  and the continuation payoff of  $s$  at this subgame is  $m$ . Then if at this subgame  $s$  proposes  $m + \epsilon$  ( $\epsilon > 0$ ),  $\bar{b}$  must reject (otherwise  $s$  can get more than  $m$ ). Since the total surplus is 1,  $\bar{b}$  must obtain at least  $1 - m - \epsilon$  in the continuation game in order to reject  $s$ 's offer and  $s$  gets at least  $m$ , this implies that the continuation payoff of all  $b \neq \bar{b}$  after  $\bar{b}$ 's rejection is less than  $\epsilon$ . The result follows by making  $\epsilon$  arbitrarily small (and by appealing to the finiteness of  $f$ ).

But the last two claims contradict each other unless the equilibrium is competitive. This establishes the result for the case in which there is one seller and many buyers. The case of a market with more than one seller is established by induction on the number of sellers.

The matching technology in the above model is random. RW also consider another market game with the matching is endogenous: At each date each seller (the short side of the market) chooses his trading partner. Here, they show that non-competitive outcomes and multiplicity of equilibria survive even when the players discount the future. By strengthening the notion of complexity S also shows that in the endogenous matching model of RW the competitive outcome is the only equilibrium if complexity considerations are present.

These results suggest perfectly competitive behavior may result if agents have, at least at the margin, preferences for simple strategies. Unfortunately, both RW and S have too simple a market set-up; for example, it is assumed that the buyers are all identical, similarly for the sellers and each agent trades at most one unit of the good. Do the conclusions extend to richer models of trade?

### Heterogeneous Markets

There are good reasons to think that it may be too difficult (or even impossible) to establish a similar set of conclusions as in S in a richer framework. For example, consider a heterogeneous market for a single indivisible good, where buyers (and sellers) have a range of valuations of the good and each buyer wants at most one unit of the good and each seller has one unit of the good for sale. In this case the analysis of S will not suffice. First, in the homogeneous market of RW, except for the special case where the number of buyers is equal to the number of sellers, the competitive equilibrium price is either 0 or 1 and all of the surplus goes to one side of the market. S's selection result crucially uses this property of the competitive equilibrium. By contrast, in a heterogeneous market, in general there will be agents receiving positive payoffs on both sides of the market in a competitive equilibrium. Therefore, one cannot justify the competitive outcome simply by focusing on extreme outcomes in which there is no surplus for one party from trade. Second, in a homogeneous market individually rational trade is by definition efficient. This may not be the case in a heterogeneous market (an inefficient trade between inframarginal and an extramarginal agent can be individually rational). Third, in a homogeneous market, the set of competitive prices remains constant, independently of the set of agents remaining in the market. In the heterogeneous market, this need not be so and in some

cases, the new competitive interval may not even intersect the old one. The change in the competitive interval of prices as the result of trade exacerbates the problems associated with using an induction hypothesis because here future prices may be conditioned on past trades even if prices are restricted to be competitive ones.

Despite these difficulties associated with a market with a heterogeneous set of buyers and sellers, Gale and Sabourian [22], henceforth GS, show that the conclusions of S can be extended to the case of a heterogeneous market in which each agent trades at most one unit of the good. GS, however, focus on *deterministic* sequential matching models in which one pair of agents are matched at each date and they leave the market if they reach an agreement. In particular, they start by considering exogenous matching processes in which the identities of the proposer and responder at each date are an exogenous and deterministic function of the set of agents remaining in the market and the date. The main result of the paper is that a PEC is always competitive in such a heterogeneous market, thus supporting the view that competitive equilibrium may arise in a finite market where complex behavior is costly.

The notion of complexity in GS is similar to that in S [15]. However, in the GS set-up with heterogeneous buyers and sellers the set of remaining agents changes depending who has traded and left the market and who is remaining, and this affects the market conditions. (In the homogeneous case, only the *number* of remaining agents matters.) Therefore, the definition of complexity in GS is with reference to a given set of remaining agents. GS also discuss an alternative notion of complexity that is independent of the set of remaining agents; such a definition may be too strong and may result in an equilibrium set being empty.

To show their result, GS first establish two very useful restrictions on the strategies that form a NEC (similar to the no delay result in Chatterjee and Sabourian [15]). First, they show that if along the equilibrium path a pair of agents  $k$  and  $\ell$  trade at a price  $p$  with  $k$  as the proposer and  $\ell$  as the responder then  $k$  and  $\ell$  always trade at  $p$ , irrespective of the previous history, whenever the two agents are matched in the same way with the same remaining set of agents. To show this consider first the case of the responder  $\ell$ . Then it must be that at every history with the same remaining set of agents  $\ell$  always accepts  $p$  by  $k$ . Otherwise,  $\ell$  could economize on complexity by choosing another strategy that is otherwise identical to his equilibrium strategy except that it always accepts  $p$  from  $k$  without sacrificing any payoff. Such a change of behavior is clearly more simple than sometimes accepting and sometimes re-

jecting the offer and moreover, it results in either agent  $k$  proposing  $p$  and  $\ell$  accepting, so the payoff to agent  $\ell$  is the same as from the equilibrium strategy, or agent  $k$  not offering  $p$ , in which case the change in the strategy is not observed and the play of the game is unaffected by the deviation. Furthermore, it must also be that at every history with the same remaining set of agents agent  $k$  proposes  $p$  in any match with  $\ell$ . Otherwise,  $k$  could economize on complexity by choosing another strategy that is otherwise identical to his equilibrium strategy except that it always proposes  $p$  to  $\ell$  without sacrificing any payoff on the equilibrium path: Such a change of behavior is clearly more simple and moreover  $k$ 's payoff is not affected because either agent  $k$  and  $\ell$  are matched and  $k$  proposes  $p$  and  $\ell$  by the previous argument accepts, so the payoff to agent  $k$  is the same as from the equilibrium strategy, or agent  $k$  and  $\ell$  are not matched with  $k$  as the proposer, in which case the change in the strategy is not observed and the play of the game is unaffected by the deviation.

GS's shows a second restriction, again with the same remaining set of agents, namely that in any NEC for any pair of agents  $k$  and  $\ell$ , player  $\ell$ 's response to  $k$ 's (on or off-the-equilibrium path) offer is always the same. Otherwise, it follows that  $\ell$  sometimes accepts an offer  $p$  by  $k$  and sometimes rejects (with the same remaining set of agents). Then by the first restriction it must be that if such an offer is made by  $k$  to  $\ell$  on the equilibrium path it is rejected. But then  $\ell$  can could economize on complexity by always rejecting  $p$  by  $k$  without sacrificing any payoff on the equilibrium path: Such a change of behavior is clearly more simple and furthermore  $\ell$ 's payoff is not affected because such a behavior is the same as what the equilibrium strategy prescribes on the equilibrium path.

By appealing to the above two properties of NEC and to the competitive nature of the market GS establish, using a complicated induction argument, that every PEC induces a competitive outcome in which each trade occurs at the same competitive price.

The matching model we have described so far is deterministic and exogenous. The selection result of GS however extends to richer deterministic matching models. In particular, GS also consider a semi-endogenous sequential matching model in which the choice of partners is endogenous but the identity of the proposer at any date is exogenous. Their results extends to this variation, with an endogenous choice of responders. A more radical departure change would be to consider the case where at any date any agent can choose his partner and make a proposal. Such a totally endogenous model of trade generates new conceptual problems. In a recent working paper Gale and Sabourian [24] consider a continuous time version of

such a matching model and show that complexity considerations allows one to select a competitive outcome in the case of totally endogenous matching. Since the selection result holds for all the different matching models we can conclude that complexity considerations inducing a competitive outcome seem to be a robust result in deterministic matching and bargaining market games with heterogeneous agents.

Random matching is commonly used in economic models because of its tractability. The basic framework of GS, however, does not extend to such a framework if either the buyers or the sellers are not identical. This is for two different reasons. First, in general in any random framework there is more than one outcome path that can occur in equilibrium with a positive probability; as a result introducing complexity lexicographically may not be enough to induce agents to behave in a simple way (they will have to be complex enough to play optimally along all paths that occur with a positive probability). Second, in Gale and Sabourian [23] it is shown that subgame perfect equilibria in Markov strategies are not necessarily perfectly competitive for the random matching model with heterogeneous agents. Since the definition of complexity in GS is such that Markov strategies are the least complex ones, it follows that with random matching the complexity definition used in GS is not sufficient to select a competitive outcome.

### Complexity and Off-The-Equilibrium Path Play

The concept of the PEC (or NEC) used in S, GS and elsewhere was defined to be such that for each player the strategy/automaton has minimal complexity amongst all strategies/automata that are best responses to the equilibrium strategies/automata of others. Although, these concepts are very mild in the treatment of complexity, it should be noted that there are other ways of introducing complexity into the equilibrium concept. One extension of the above set-up is to treat complexity as a (small) positive fixed cost of choosing a more complex strategy and define a Nash (subgame perfect) equilibrium with a fixed positive complexity costs accordingly. All the selection results based on lexicographic complexity in the papers we discuss in this survey also hold for positive small complexity costs. This is not surprising because with positive costs complexity has at least as much bite as in the lexicographic case; there is at least as much refinement of the equilibrium concept with the former as with the latter. In particular, in the case of a NEC (or a PEC), *in considering complexity*, players ignore any consideration of payoffs off the equilibrium path and the trade-off is between the equilibrium payoffs of two strategies and the complexity of the two.

As a result these concepts put more weight on complexity costs than on being “prepared” for off-the-equilibrium-path moves. Therefore, although complexity costs are insignificant, they take priority over optimal behavior after deviations. (See [16] for a discussion.)

A different approach would be to assume that complexity is a less significant criterion than the off-the-equilibrium payoffs. In the extreme case, one would require agents to choose minimally complex strategies among the set of strategies that are best responses on and off the equilibrium path (see Kalai and Neme [28]).

An alternative way of illustrating the differences between the different approaches is by introducing two kinds of vanishingly small perturbations into the underlying game. One perturbation is to impose a small but positive cost of choosing a more complex strategy. Another perturbation is to introduce a small but positive probability of making an error (off-the-equilibrium-path move). Since a PEC requires each agents to choose a minimally complex strategy within the set of best responses, it follows that the limit points of Nash equilibria of the above perturbed game correspond to the concept of PEC if we first let the probability of making an off-the-equilibrium-path move go to zero and then let the cost of choosing a more complex strategy go to zero (this is what Chatterjee and Sabourian [15] do). On the other hand, in terms of the above limiting arguments, if we let the cost of choosing a more complex strategy go to zero and then let the probability of making an off-the-equilibrium-path move go to zero then any limit corresponds to the equilibrium definition in Kalai and Neme [28] where agents choose minimally complex strategies among the set of strategies that are best responses on and off the equilibrium path.

Most of the results reported in this paper on refinement and endogenous complexity (for example Abreu-Rubinstein [1]), Chatterjee and Sabourian [15], Gale and Sabourian [22] and Lee and Sabourian [31] hold only for the concept of NEC and its variations and thus depend crucially on assuming that complexity costs are more important than off-the-equilibrium payoffs. This is because these results always appeal to an argument that involves economizing on complexity if the complexity is not used off the equilibrium path. Therefore, they may be a good predictor of what may happen only if complexity costs are more significant than the perturbations that induce off-the-equilibrium-path behavior. The one exception is the selection result in S [47]. Here, although the result we have reported is stated for NEC and its variations, it turns out that the selection of competitive equilibrium does not in fact depend on the relative importance of complexity costs and off-the-equilibrium path payoffs. It remains true even

for the case where the strategies are required to be least complex amongst those that are best responses at every information set. This is because in S’s analysis complexity is only used to show that every agent’s response to the price offer of 1 is always the same irrespective of the past history of play. This conclusion holds irrespective of the relative importance of complexity costs and off-the-equilibrium payoff because trading at the price of 1 is the best outcome that any seller can achieve at any information set (including those off-the-equilibrium) and a worst outcome for any buyer. Therefore, irrespective of the order, the strategy of sometimes accepting a price of 1 and sometimes rejecting cannot be an equilibrium for a buyer (similar arguments applies for a seller) because the buyer can economize on complexity by always rejecting the offer without sacrificing any payoff off or on-the-equilibrium path (accepting  $p = 1$  is a worse possible outcome).

## Discussion and Future Directions

The use of finite automata as a model of players in a game has been criticized as being inadequate, especially because as the number of states becomes lower it becomes more and more difficult for the small automaton to do routine calculations, let alone the best response calculations necessary for game-theoretic equilibria. Some of the papers we have explored address other aspects of complexity that arise from the concrete nature of the games under consideration. Alternative models of complexity are also suggested, such as computational complexity and communication complexity.

While our work and the earlier work on which it builds focuses on equilibrium, an alternative approach might seek to see whether simplicity evolves in some reasonable learning model. Maenner [32] has undertaken such an investigation with the infinitely repeated Prisoners’ Dilemma (studied in the equilibrium context by Abreu and Rubinstein). Maenner provides an argument for “learning to be simple”. On the other hand, there are arguments for increasing complexity in competitive games ([42]). It is an open question, therefore, whether simplicity could arise endogenously through learning, though it seems to be a feature of most human preferences and aesthetics (see [11]).

The broader research program of explicitly considering complexity in economic settings might be a very fruitful one. Auction mechanisms are designed with an eye towards how complex they are – simplicity is a desideratum. The complexity of contracting has given rise to a whole literature on incomplete contracts, where some models postulate a fixed cost per contingency described in the con-

tract. All this is apart from the popular literature on complexity, which seeks to understand complex, adaptive systems from biology. The use of formal complexity measures such as those considered in this survey and the research we describe might throw some light on whether incompleteness of contracts, or simplicity of mechanisms, is an assumption or a result (of explicitly considering choice of level of complexity).

## Acknowledgments

We wish to thank an anonymous referee and Jihong Lee for valuable comments that improved the exposition of this chapter. We would also like to thank St. John's College, Cambridge and the Pennsylvania State University for funding Dr Chatterjee's stay in Cambridge at the time this chapter was written.

## Bibliography

1. Abreu D, Rubinstein A (1988) The structure of Nash equilibria in repeated games with finite automata. *Econometrica* 56:1259–1282
2. Anderlini L (1990) Some notes on Church's thesis and the theory of games. *Theory Decis* 29:19–52
3. Anderlini L, Sabourian H (1995) Cooperation and effective computability. *Econometrica* 63:1337–1369
4. Aumann RJ (1981) Survey of repeated games. In: Essays in game theory and mathematical economics in honor of Oskar Morgenstern. Bibliographisches Institut, Mannheim/Vienna/Zurich, pp 11–42
5. Banks J, Sundaram R (1990) Repeated games, finite automata and complexity. *Games Econ Behav* 2:97–117
6. Ben Porath E (1986) Repeated games with bounded complexity. Mimeo, Stanford University
7. Ben Porath E (1993) Repeated games with finite automata. *J Econ Theory* 59:17–32
8. Binmore KG (1987) Modelling rational players I. *Econ Philos* 3:179–214
9. Binmore KG, Samuelson L (1992) Evolutionary stability in repeated games played by finite automata. *J Econ Theory* 57:278–305
10. Binmore KG, Piccione M, Samuelson L (1998) Evolutionary stability in alternating-offers bargaining games. *J Econ Theory* 80:257–291
11. Birkhoff GD (1933) Aesthetic measure. Harvard University Press, Cambridge
12. Bloise G (1998) Strategic complexity and equilibrium in repeated games. Unpublished doctoral dissertation, University of Cambridge
13. Busch L-A, Wen Q (1995) Perfect equilibria in a negotiation model. *Econometrica* 63:545–565
14. Chatterjee K (2002) Complexity of strategies and multiplicity of Nash equilibria. *Group Decis Negot* 11:223–230
15. Chatterjee K, Sabourian H (2000) Multiperson bargaining and strategic complexity. *Econometrica* 68:1491–1509
16. Chatterjee K, Sabourian H (2000) N-person bargaining and strategic complexity. Mimeo, University of Cambridge and the Pennsylvania State University
17. Debreu G (1959) Theory of value. Yale University Press, New Haven/London
18. Fernandez R, Glazer J (1991) Striking for a bargain between two completely informed agents. *Am Econ Rev* 81:240–252
19. Fudenberg D, Maskin E (1990) Evolution and repeated games. Mimeo, Harvard/Princeton
20. Fudenberg D, Tirole J (1991) Game theory. MIT Press, Cambridge
21. Gale D (2000) Strategic foundations of general equilibrium: Dynamic matching and bargaining games. Cambridge University Press, Cambridge
22. Gale D, Sabourian H (2005) Complexity and competition. *Econometrica* 73:739–770
23. Gale D, Sabourian H (2006) Markov equilibria in dynamic matching and bargaining games. *Games Econ Behav* 54:336–352
24. Gale D, Sabourian H (2008) Complexity and competition II; endogenous matching. Mimeo, New York University/University of Cambridge
25. Haller H, Holden S (1990) A letter to the editor on wage bargaining. *J Econ Theory* 52:232–236
26. Hayek F (1945) The use of knowledge in society. *Am Econ Rev* 35:519–530
27. Herrero M (1985) A Strategic theory of market institutions. Unpublished doctoral dissertation, London School of Economics
28. Kalai E, Neme A (1992) The strength of a little perfection. *Int J Game Theory* 20:335–355
29. Kalai E, Stanford W (1988) Finite rationality and interpersonal complexity in repeated games. *Econometrica* 56:397–410
30. Klemperer P (ed) (2000) The economic theory of auctions. Elgar, Northampton
31. Lee J, Sabourian H (2007) Coase theorem, complexity and transaction costs. *J Econ Theory* 135:214–235
32. Maenner E (2008) Adaptation and complexity in repeated games. *Games Econ Behav* 63:166–187
33. Miller GA (1956) The magical number seven plus or minus two: Some limits on our capacity to process information. *Psychol Rev* 63:81–97
34. Neme A, Quintas L (1995) Subgame perfect equilibrium of repeated games with implementation cost. *J Econ Theory* 66:599–608
35. Neyman A (1985) Bounded complexity justifies cooperation in the finitely-repeated Prisoners' Dilemma. *Econ Lett* 19:227–229
36. Neyman A (1997) Cooperation, repetition and automata in co-operation: Game-theoretic approaches. In: Hart S, Mas-Colell A (eds) NATO ASI Series F, vol 155. Springer, Berlin, pp 233–255
37. Osborne M, Rubinstein A (1990) Bargaining and markets. Academic, New York
38. Osborne M, Rubinstein A (1994) A course in game theory. MIT Press, Cambridge
39. Papadimitriou CH (1992) On games with a bounded number of states. *Games Econ Behav* 4:122–131
40. Piccione M (1992) Finite automata equilibria with discounting. *J Econ Theory* 56:180–193
41. Piccione M, Rubinstein A (1993) Finite automata play a repeated extensive game. *J Econ Theory* 61:160–168
42. Robson A (2003) The evolution of rationality and the Red Queen. *J Econ Theory* 111:1–22

43. Rubinstein A (1982) Perfect equilibrium in a bargaining model. *Econometrica* 50:97–109
44. Rubinstein A (1986) Finite automata play the repeated Prisoners’ Dilemma. *J Econ Theory* 39:83–96
45. Rubinstein A (1998) Modeling bounded rationality. MIT Press, Cambridge
46. Rubinstein A, Wolinsky A (1990) Decentralized trading, strategic behaviour and the Walrasian outcome. *Rev Econ Stud* 57:63–78
47. Sabourin H (2003) Bargaining and markets: Complexity and the competitive outcome. *J Econ Theory* 116:189–228
48. Selten R (1965) Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Z gesamte Staatswiss* 12:201–324
49. Shaked A (1986) A three-person unanimity game. In: The Los Angeles national meetings of the Institute of Management Sciences and the Operations Research Society of America, Mimeo, University of Bonn
50. Zemel E (1989) Small talk and cooperation: A note on bounded rationality. *J Econ Theory* 49:1–9

## GARCH Modeling

CHRISTIAN M. HAFNER  
Université catholique de Louvain,  
Louvain-la-Neuve, Belgium

### Article Outline

- [Glossary](#)
- [Definition of the Subject](#)
- [Introduction](#)
- [Properties of the GARCH\(1,1\) Model](#)
- [Estimation and Inference](#)
- [Testing for ARCH](#)
- [Asymmetry, Long Memory, GARCH-in-Mean](#)
- [Non- and Semi-parametric Models](#)
- [Multivariate GARCH Models](#)
- [Stochastic Volatility](#)
- [Aggregation](#)
- [Future Directions](#)
- [Bibliography](#)

### Glossary

- ACF** Autocorrelation Function
- ARMA** Autoregressive Moving Average
- BEKK** A multivariate GARCH model named after an early unpublished paper by Baba, Engle, Kraft and Kroner.
- CCC** Constant Conditional Correlation
- DCC** Dynamic Conditional Correlation
- CAPM** Capital Asset Pricing Model

**GARCH** Generalized Autoregressive Conditional Heteroskedasticity

**Heteroskedasticity** A non-constant variance that depends on the observation or on time.

**i.i.d.** independent, identically distributed

**Kurtosis** A standardized fourth moment of a random variable that tells something about the shape of the distribution. A Gaussian distribution has a kurtosis of three. If the kurtosis is larger than three, then typically the distribution will have tails that are thicker than those of the Gaussian distribution.

**Lag** An operation that shifts the time index of a time series. For example, the first lag of  $y_t$  is  $y_{t-1}$ .

**Long memory** Property of covariance stationary processes without absolutely summable ACF, meaning that the ACF decays slowly.

**Realized volatility** Sum of intra-day squared returns as a measure for daily volatility.

**Skewness** A standardized third moment of a random variable that tells something about the asymmetry of the distribution. Symmetric distributions have skewness equal to zero.

**Volatility** Degree of fluctuation of a time series around its mean.

### Definition of the Subject

GARCH (Generalized Autoregressive Conditional Heteroskedasticity) is a time series model developed by [44] and [21] to describe the way volatility changes over time. In a GARCH model, the volatility at a given time  $t$ ,  $\sigma_t^2$  say, is a function of lagged values of the observed time series  $y_t$ . The GARCH model can be written as  $y_t = \sigma_t \xi_t$ , with  $\xi_t$  being an independent, identically distributed (i.i.d.) error term with mean zero and variance one, and where

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (1)$$

with constant parameters  $\omega$ ,  $\alpha_1, \dots, \alpha_q$  and  $\beta_1, \dots, \beta_p$ . Model (1) is also called GARCH( $p, q$ ), analogous to ARMA( $p, q$ ), as it includes  $p$  lagged volatilities and  $q$  lagged squared values of  $y_t$ . In this model,  $\sigma_t^2$  is the variance of  $y_t$  conditional on the observations until time  $t-1$ . It is specified as a linear function of lagged squared  $y_t$  and lagged conditional variances. Many extensions or modifications of the basic model in (1) have been proposed, the most prominent being the exponential GARCH model of [99] and the threshold GARCH models of [123] and [59]. [71] and [42] provided classes of models that contain a large number of suggested models of the GARCH type.

## Introduction

In the late seventies of the last century it became obvious that volatilities of financial assets are indeed not constant, nor deterministic or seasonal, but rather stochastic in nature. There is an unsystematic change between periods of high volatility and periods of low volatility. This ‘volatility clustering’ had already been remarked in the early works of [93] and [54]. It was one of several stylized facts of financial asset returns, another of which was the observation that the distribution of returns is not Gaussian. Of course, these features were not necessarily treated in an independent way, and in fact it was soon discovered that very likely one of the effects was causing another, such as volatility clustering causing leptokurtosis, or fat tailed distributions. For example, consider the simple model for asset returns  $y_t$ ,

$$y_t = \sigma_t \xi_t$$

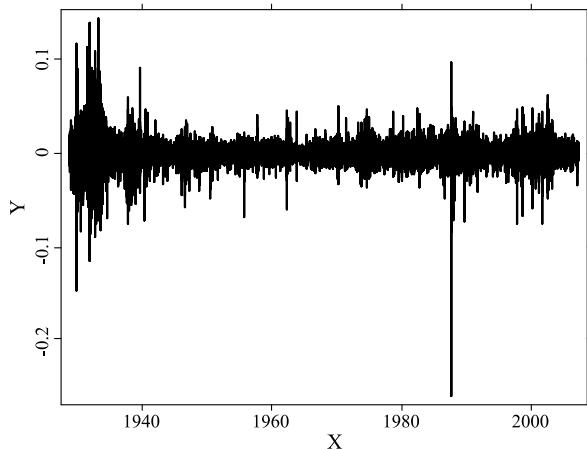
where  $\xi_t \sim N(0, 1)$  and  $\sigma_t$  is stochastic with  $E[\sigma_t^2] = \sigma^2$ , say, and independent of present and future  $\xi_t$ . Then it is straightforward to show that the kurtosis of  $y_t$  is given by

$$\kappa = \frac{E[y_t^4]}{E[y_t^2]^2} = 3 + 3 \frac{\text{Var}(\sigma_t^2)}{\sigma^4}. \quad (2)$$

Thus, returns in this model are Gaussian distributed if and only if  $\text{Var}(\sigma_t^2) = 0$ , i.e., volatility is non-stochastic. Moreover, as the second term on the right hand side of (2) is always positive, the kurtosis will be larger than three under stochastic volatility, which often means that its tails are fatter than those of the Gaussian distribution. In other words, extreme events are more likely under stochastic volatility compared with constant volatility.

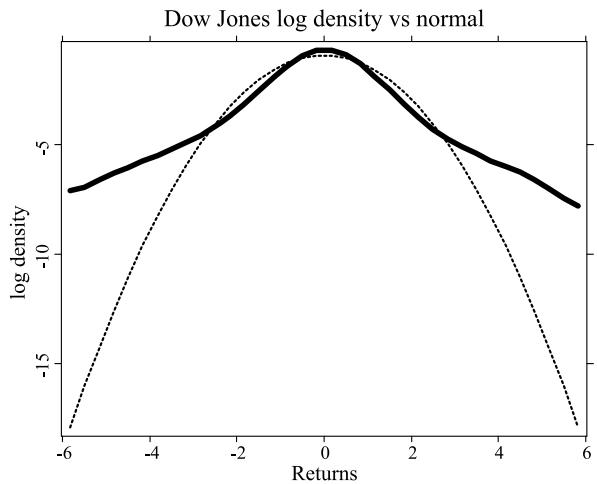
To illustrate the effects of volatility clustering and fat tails, consider the daily returns on the Dow Jones Industrial Index over the period October 1928 to April 2007. A graph of the (log) index  $X_t$  and returns, defined as  $y_t = X_t - X_{t-1}$ , is given in Fig. 1. Clearly visible is the volatility clustering in the beginning of the sample period and around the year 2000, while the years at the end of the sample showed less volatility. Also visible is the crash of October 1987 where the index dropped by 22 percent.

Figure 2 shows a nonparametric estimator of the logarithmic density of returns, compared with the analogue of a Gaussian distribution. Clearly, the Dow Jones returns distributions has fat tails, i.e., there are more extreme events than one would expect under normality. There are also more returns close to zero than under normality. Concerning volatility clustering, Fig. 3 shows the autocorrelation function of returns and squared returns. While there is very little structure in the ACF of returns, the ACF of



**GARCH Modeling, Figure 1**

Daily returns of the Dow Jones Index, 1928 to 2007, defined as first difference of the log index

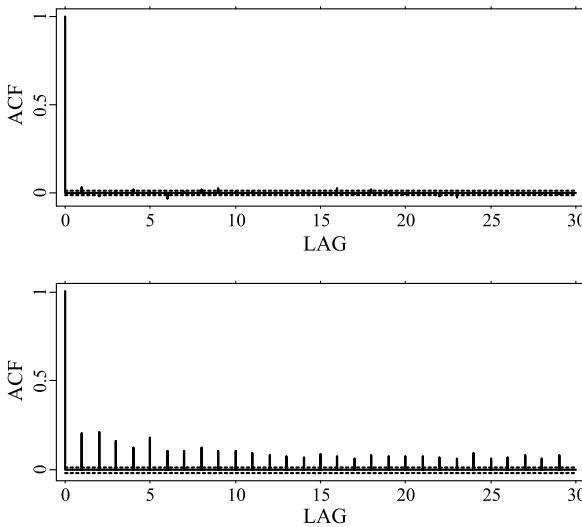


**GARCH Modeling, Figure 2**

Dow Jones log density versus the Gaussian log density

squared returns are all positive and highly significant. This positive autocorrelation is explained by the fact that large returns tend to be followed by large returns and small returns tend to be followed by small returns.

Yet, realizing that volatility is stochastic does not tell us which model we should use for it. In practice, people are sometimes debating whether they should take historical volatilities over 20 or 100 days, say. They notice that calculating the standard deviation over a shorter period is more accurate when recent upturns or downturns want to be captured, while it is far less efficient than a longer time window when volatility has not changed much. Thus, there is some kind of bias-variance trade-off. The problem is that the optimal window length typically changes



**GARCH Modeling, Figure 3**  
Dow Jones autocorrelation function of returns (upper panel) and squared returns (lower panel)

over time, and it is virtually impossible to adjust historical volatility windows automatically to market developments. A related problem is that historical volatilities imply a weighting scheme that is highly questionable: Why should  $k$  days be incorporated in the calculation with equal weight, but no weights are put to days up to  $k + 1$  days ago? A smoother weighting scheme seems more natural, and in particular, an exponential scheme seems attractive. Thus, for example, we may specify for  $\sigma_t^2$

$$\sigma_t^2 = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^i y_{t-1-i}^2 \quad (3)$$

with parameter  $\lambda \in (0, 1)$ . Equation (3) can be rewritten as

$$\sigma_t^2 = (1 - \lambda) y_{t-1}^2 + \lambda \sigma_{t-1}^2, \quad (4)$$

which looks more familiar. It is actually the model used by RiskMetrics of JP Morgan, when the smoothing parameter is fixed to 0.94. RiskMetrics is often used in practice as a means to calculate the Value-at-Risk (VaR) of a portfolio and to assess the market risk of a bank, required by the Basel Committee for Banking Supervision, see e.g., [78] and [95]. The VaR is essentially an extreme quantile of the distribution of portfolio returns. Under Gaussianity, for example, the VaR is a direct function of volatility. The RiskMetrics model is a special case of the integrated GARCH model of [47].

The generalized autoregressive conditional heteroskedasticity – GARCH – model is based on the seminal work of [44] and [21]. The idea is to do exponential

smoothing in a more flexible way than RiskMetrics but keeping the model parsimonious. The particular specification reveals many similarities to autoregressive moving average (ARMA) time series models. In its most often used form, the standard GARCH model of order (1,1) reads

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (5)$$

where  $\omega$ ,  $\alpha$  and  $\beta$  are parameters to be estimated from the data. Thus, the conditional variance is a linear function of lagged squared observations  $y_t$  and lagged conditional variances. Comparing (5) with the RiskMetrics model (4), it becomes clear that in the GARCH(1,1) model a constant is added, the parameter  $\alpha$  takes the role of  $1 - \lambda$  and  $\beta$  that of  $\lambda$ . But since  $\alpha$  and  $\beta$  can be chosen independently, the GARCH model is more flexible than RiskMetrics. In (5), substituting successively for  $\sigma_{t-i}^2$ , one obtains the analogue representation of (3),

$$\sigma_t^2 = \frac{\omega}{1 - \beta} + \alpha \sum_{i=0}^{\infty} \beta^i y_{t-1-i}^2, \quad (6)$$

which clearly shows the exponential smoothing feature of the GARCH(1,1) model. The basic model can now be extended to allow for more lags. The GARCH( $p, q$ ) model is given by

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i y_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (7)$$

extending the number of parameters to  $p + q + 1$ . GARCH models of order higher than (1,1) allow for more complex autocorrelation structures of the squared process. However, in most empirical studies coefficients corresponding to higher lags turned out to be insignificant and thus simple GARCH(1,1) have clearly dominated models of higher order.

Although extremely successful due to its simplicity and yet accurate description of volatility changes, a thorough understanding of its stochastic properties such as stationarity or positivity constraints took many years. For example, [98] shows that, paradoxically at first sight, conditions for strict stationarity are less rigid than those for covariance stationarity if error terms are Gaussian, the reason being that covariance stationarity requires finite variances whereas strict stationarity does not. Moreover, the often given parameter restrictions  $\omega > 0$ ,  $\alpha_i, \beta_j \geq 0$  are only sufficient but not necessary for  $\sigma_t > 0$  almost surely as demonstrated by [100]. These are just two examples for the subtleties of the theory of univariate GARCH processes.

Nevertheless, the immense success of the simple GARCH(1,1) model to explain many sorts of financial and

macroeconomic time series was irreversible, partly also because it became available in standard statistical programming packages. The theory of estimation and inference developed rapidly, although perhaps still being underway, and estimation time of a GARCH(1,1) model for a thousand or so observations decreased from minutes in the eighties over seconds in the nineties to just fractions of a second nowadays. With these developments it became available to a broad public, and more and more practitioners started using the model, be it for option pricing, portfolio optimization, risk management, or other purposes. Monographs and reviews appeared such as [14,20,60] and [13]. Anniversary issues of renowned journals such as Journal of Applied Econometrics, 2002, were dedicated entirely to new ideas in GARCH modeling. The Nobel price for economics in 2003 was awarded to two time series econometricians, Clive Granger and Robert Engle. The latter has mainly driven the development of a new financial econometrics discipline, based on volatility modeling but spreading also to other areas such as modeling of extreme events and risk management.

The pricing of options and other derivatives is perhaps the most typical example for where models for the volatility of an asset matter. For example, the celebrated option pricing formula of [18] does not depend on the drift of the underlying stock but well on its volatility. In fact, among the ingredients of the Black and Scholes formula, volatility is the most crucial one, the other ones being either fixed such as time of maturity or strike price, or relatively easy to determine such as a riskfree interest rate. Volatility, however, has always been subject to debates about how exactly to find accurate measures for it. The Black and Scholes assumption of constant volatility is actually less crucial to their formula than one often thinks. Actually, if volatility is time-varying but in a deterministic way, then the Black and Scholes formula remains valid. One just has to replace the volatility parameter by the mean of the volatility function from today until the time of maturity of the option contract, see e.g., [90]. If, however, volatility is stochastic, i. e., it has an additional source of randomness, then markets are no longer complete and the Black and Scholes formula breaks down. In that case, assumptions about the volatility risk premium have to be made. In continuous time stochastic volatility models a classical paper is [74], while in a discrete time GARCH framework, [41] derives results for option pricing.

### Properties of the GARCH(1,1) Model

For the sake of simplicity let us consider the univariate GARCH(1,1) model given in (5), where we additionally

assume that the conditional distribution of  $y_t$  is Gaussian. The model can be written as

$$y_t = \sigma_t \xi_t, \quad \xi_t \sim \text{i.i.d. } N(0, 1) \quad (8)$$

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2. \quad (9)$$

In the following we discuss a few properties of model (8). First, the GARCH model specifies the *conditional* variance, where the condition is the information set generated by the process  $y_t$ . Formally, it is given by the sigma-algebra  $\mathcal{F}_t = \sigma(y_t, y_{t-1}, \dots)$ . With this notation we can write  $\sigma_t^2 = \text{Var}(y_t | \mathcal{F}_{t-1})$ , since  $\sigma_t^2$  is  $\mathcal{F}_{t-1}$ -measurable. As the information set changes over time, the conditional variance also changes. On the other hand, this does not imply that the *unconditional* variance is also time-varying. In fact, for model (8) it is quite straightforward to show that the unconditional variance, if it exists, is constant and given by

$$\text{Var}(y_t) = \frac{\omega}{1 - \alpha - \beta}.$$

A necessary and sufficient condition for the existence of the unconditional variance is

$$\alpha + \beta < 1, \quad (10)$$

see [21]. He also shows that condition (10) is necessary and sufficient for the process  $\{y_t\}$  to be covariance stationary. In that case, the autocorrelation function of  $\{y_t\}$  is given by  $\rho_\tau(y_t) = 0, \forall \tau \geq 1$ . Moreover, both the conditional and unconditional mean of  $y_t$  are zero, so that the process  $\{y_t\}$  has the properties of a white noise without being an i.i.d. process. The dependence occurs in higher moments of the process. For example, the autocorrelation function of the squared process, provided that fourth moments exist, is given by

$$\rho_1(y_t^2) = \alpha \frac{1 - \alpha\beta - \beta^2}{1 - 2\alpha\beta - \beta^2} \quad (11)$$

$$\rho_\tau(y_t^2) = (\alpha + \beta)\rho_{\tau-1}(y_t^2), \quad \tau \geq 2. \quad (12)$$

From (11) and (12) it is obvious that in the GARCH(1,1) model all autocorrelations of squared returns are positive with an exponential decay. This decay is slow if  $\alpha + \beta$  is close to one, as often found for empirical data. One also characterizes this coefficient as the “persistence” parameter of the GARCH(1,1) model. The closer the persistence parameter is to one, the longer will be the periods of volatility clustering. On the other hand, the larger  $\alpha$  relative to  $\beta$ , the higher will be the immediate impact of lagged squared returns on volatility.

The necessary and sufficient condition for finite fourth moments is given by  $\beta^2 + 2\alpha\beta + 3\alpha^2 < 1$ , see [21]. In that case, the kurtosis of  $y_t$  is given by

$$\kappa = 3 + \frac{6\alpha^2}{1 - \beta^2 - 2\alpha\beta - 3\alpha^2},$$

which is larger than three since the second term on the right hand side is positive. Hence, the GARCH(1,1) exhibits fat tails compared with a normal distribution.

A link between the GARCH model and an ARMA model is given by considering the squared process,  $\{y_t^2\}$ . By simply adding  $y_t^2$  and subtracting  $\sigma_t^2$  on both sides of Eq. (9), one obtains

$$y_t^2 = \omega + (\alpha + \beta)y_{t-1}^2 - \beta u_{t-1} + u_t \quad (13)$$

where  $u_t = y_t^2 - \sigma_t^2$ . Equation (13) is an ARMA(1,1) in  $y_t^2$ , since  $u_t$  is a white noise error term: We have  $E[u_t | \mathcal{F}_{t-1}] = 0$ , which implies that all autocorrelations of  $u_t$  are zero.

It is possible that the process  $\{y_t\}$  is strictly stationary without being covariance stationary, simply because a finite variance is not necessary for strict stationarity. If the process starts in the infinite past, a necessary and sufficient condition for strict stationarity of the GARCH(1,1) process as shown by [98] is given by

$$E[\log(\alpha\xi_t^2 + \beta)] < 0, \quad (14)$$

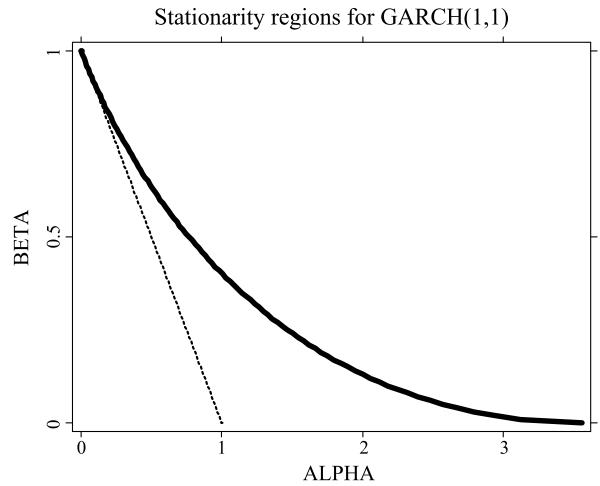
which is indeed weaker than condition (10). This follows directly by noting that (10) is equivalent to  $\log(\alpha + \beta) = \log(E[\alpha\xi_t^2 + \beta]) < 0$ . Thus, by Jensen's inequality,  $E[\log(\alpha\xi_t^2 + \beta)] < \log(E[\alpha\xi_t^2 + \beta]) < 0$ . For example, for an ARCH(1) model (i.e., a GARCH(1,1) model with  $\beta = 0$ ),  $\alpha$  can be as large as 3.56 and still the process is strictly stationary. Figure 4 shows the different stationarity regions as a function of the two parameters. [25] generalized condition (14) to the GARCH( $p, q$ ) case.

Under the sufficient condition (10), the GARCH(1,1) process with Gaussian innovations is also geometrically ergodic and  $\beta$ -mixing with exponential decay as shown by [28].

If condition (14) holds, then the process  $\{y_t\}$  has a stationary distribution whose tails are of the Pareto type. That is, for large  $x$  and some  $a, k > 0$ ,

$$p(x) = Pr(y_t > x) = kx^{-a}. \quad (15)$$

The coefficient  $a$  is known as the tail index. The smaller  $a$ , the fatter the tail of the distribution. For all  $c, 0 \leq c < \alpha$ ,  $E[|y_t|^c] < \infty$ . [43] showed that a stationary ARCH model



GARCH Modeling, Figure 4

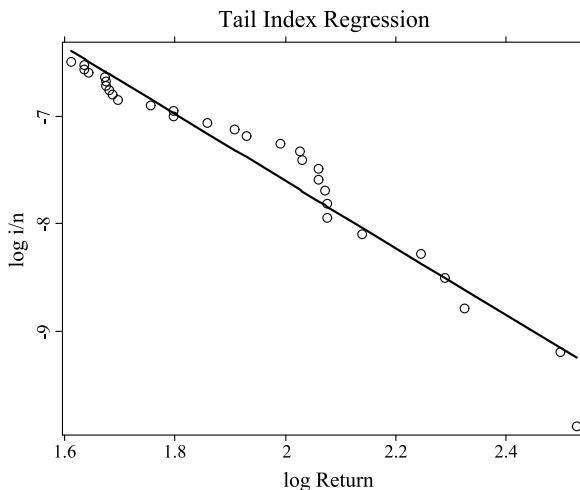
Stationarity regions for a GARCH(1,1) process with Gaussian innovations. To the left of the dashed line is the region of covariance stationarity, to the left of the thick line is the region of strict stationarity, and to the right of the thick line is the region of non-stationarity

has Pareto-like tails. Knowledge of the tail index is important for risk management in order to assess the risk of extreme events. The theoretical tail index of a fitted ARCH or GARCH model can be compared with an estimate of the empirical tail index in order to diagnose the goodness-of-fit with respect to the tails. For example, taking logarithms of (15), one obtains  $\log p(x) = \log(k) - a \log(x)$  for large  $x$ . Replacing  $x$  by the largest  $m$  order statistics of  $y_t$ , and introducing an error term, one obtains the regression

$$\log \frac{i}{n} = \log k - a \log X_{(i)} + \varepsilon_i, \quad i = 1, \dots, m \quad (16)$$

where  $X_{(i)}$  are the largest  $m$  order statistics of  $y_t$  and  $\varepsilon_i$  is an error term. One can estimate the coefficients of the linear regression (16) simply by ordinary least squares. More problematic is the choice of  $m$ , which involves a bias-variance trade-off. For the Dow Jones returns, Fig. 5 shows the tail index regression using  $m = 30$ . The OLS estimator of  $a$  is 3.12, indicating that fourth moments of returns may not exist. Another simple estimator is the Hill estimator proposed by [72], which is based on a likelihood principle. For the Dow Jones returns, the Hill estimator of  $a$  using  $m = 30$  is 2.978, which is close to the OLS estimator, suggesting that even third moments may not exist. More elaborate estimators have been proposed and we refer to the detailed discussion in [43].

The presence of autoregressive conditional heteroskedasticity has an effect on the forecast intervals for



**GARCH Modeling, Figure 5**  
Tail index regression for the Dow Jones returns

predicted  $y_{t+k}$  given information at time  $t$ . If volatility at time  $t$  as measured by the GARCH model is high (low), these will be larger (smaller) than if GARCH effects are ignored. Furthermore, forecasting the volatility itself is easily possible with the standard GARCH model, since analytical expressions can be found for the conditional mean of future volatility as a function of today's information. The conditional mean is the optimal predictor in a mean square prediction error sense. For example, to forecast  $\sigma_{t+k}^2$ , one derives for a forecast horizon of  $k \geq 2$ ,

$$\begin{aligned} E[\sigma_{t+k}^2 | \mathcal{F}_t] &= \omega(1 + (\alpha + \beta) + \dots + (\alpha + \beta)^{k-2}) \\ &\quad + (\alpha + \beta)^{k-1}\sigma_{t+1}^2. \end{aligned}$$

If the process is covariance stationary, i. e.,  $\alpha + \beta < 1$ , then volatility forecasts converge to the unconditional variance:

$$\lim_{k \rightarrow \infty} E[\sigma_{t+k}^2 | \mathcal{F}_t] = \frac{\omega}{1 - \alpha - \beta} = \text{Var}(y_t).$$

In the early literature on GARCH models, these were criticized for not providing good forecasts in terms of conventional forecast criteria. For example, when regressing the ex post squared daily returns on the forecasted conditional variance, the obtained  $R^2$  is typically small, of the order of about ten percent. [4] found that the daily squared return is not really the targeted value, but that daily volatility should rather be measured by the sum of *intra-day* squared returns, e. g., on intervals of five minute returns, which they called *realized volatility*. In terms of realized volatility, the forecasting performance of GARCH models improved substantially to levels of about fifty percent

$R^2$ . Later, a new branch of volatility modeling opened by noticing that if intra-day data are available, then it is indeed more efficient to measure daily volatility directly by realized volatility and then do forecasting of daily volatility using models fitted to realized volatility, see e. g., [5].

### Estimation and Inference

The principal estimation method for GARCH models is maximum likelihood (ML). In most cases one assumes a conditional Gaussian distribution. If the true distribution is Gaussian, then ML estimators are consistent and efficient under quite general conditions. On the other hand, if the true distribution is not Gaussian, then one loses efficiency but again under quite general conditions, consistency is retained if at least the first two conditional moments are correctly specified, see [23]. In the case of misspecification of the conditional distribution one also speaks of *quasi* maximum likelihood (QML), distinguishing it from ML where the true distribution is used, which however in general is unknown.

The log likelihood function, up to an additive constant and conditional on some starting value for the volatility process, reads  $L(\theta) = \sum_{t=1}^n l_t(\theta)$ , where

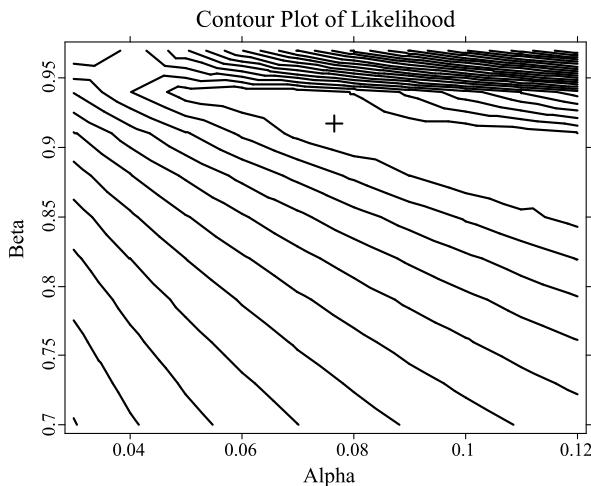
$$l_t(\theta) = -\frac{1}{2} \log \sigma_t^2(\theta) - \frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{\sigma_t^2(\theta)}$$

and where  $\theta = (\omega, \alpha, \beta)'$  is the parameter vector. The maximum likelihood estimator is then defined as the maximizer of  $L(\theta)$  over some compact set  $\Theta$ ,

$$\hat{\theta} = \arg \max_{\theta \in \Theta} L(\theta). \quad (17)$$

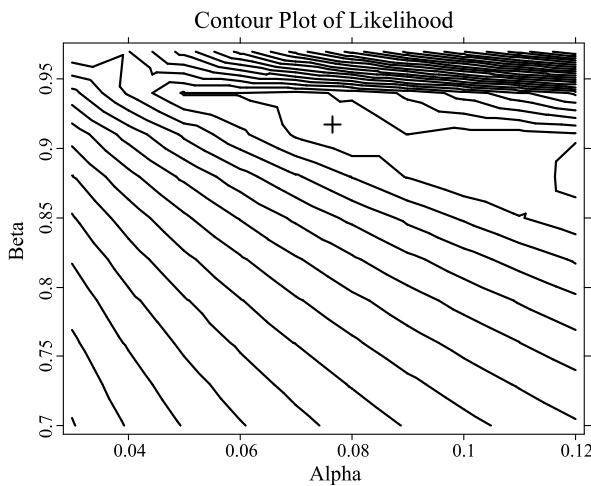
Unfortunately, there is no closed form solution to (17) but many numerical optimization procedures exist. For example, a popular algorithm is that of [15].

Figure 6 shows the likelihood function of a GARCH(1,1) process generated using the parameter estimates of the Dow Jones index returns (see Sect. “Asymmetry, Long Memory, GARCH-in-Mean”), Gaussian innovations, and the same sample size of  $n = 19727$ . The parameter  $\omega$  has been determined by the variance targeting technique of [50], i. e.,  $\omega = \sigma^2(1 - \alpha - \beta)$ , where  $\sigma^2$  is the sample variance of observed returns. In regions where  $\alpha + \beta \geq 1$ ,  $\omega$  is set to zero. Note the steep decline of the likelihood for values of  $\alpha$  and  $\beta$  that lie beyond the covariance stationarity region ( $\alpha + \beta \geq 1$ ). Figure 7 shows the same function for the observed Dow Jones index returns. No major difference can be detected between both graphs, indicating an appropriate specification of the Gaussian likelihood function.



**GARCH Modeling, Figure 6**

Contour plot of the likelihood function of a generated GARCH(1,1) process using Gaussian innovations and a sample size of  $n = 19727$ . The abscissa is the parameter  $\alpha$ , the ordinate is  $\beta$ . True values, marked by a cross in the figure, are  $\alpha = 0.0766$  and  $\beta = 0.9173$



**GARCH Modeling, Figure 7**

Contour plot of the likelihood function of the GARCH(1,1) model fitted to observed Dow Jones index returns, 1928 to 2007, with sample size  $n = 19727$ . The abscissa is the parameter  $\alpha$ , the ordinate is  $\beta$ . The maximum, marked by a cross in the figure, is obtained for  $\alpha = 0.0766$  and  $\beta = 0.9173$

If the first two moments of  $y_t$  are correctly specified and under further regularity conditions given by [118] and [23], the QML estimator is consistent with asymptotic distribution given by

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{L} N(0, \mathcal{J}^{-1} \mathcal{I} \mathcal{J}^{-1}) \quad (18)$$

where

$$\mathcal{I} = E \left[ \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'} \right], \quad \mathcal{J} = -E \left[ \frac{\partial^2 l_t}{\partial \theta \partial \theta'} \right],$$

and where the derivatives are evaluated at the true parameter values. In case the conditional distribution is indeed Gaussian, one has the identity  $\mathcal{I} = \mathcal{J}$  and the asymptotic covariance matrix reduces to the inverse of the information matrix,  $\mathcal{I}^{-1}$ . Note that consistency is retained if the conditional distribution is not Gaussian, but efficiency is lost in that case.

It is straightforward to obtain analytical formula for the score vector, the outer product of the score and the Hessian matrix, with which inference on parameter estimates can be done using the result in (18). More primitive conditions than those of [23] have been derived, e.g., by [57, 85, 91] and [66].

Maximum likelihood estimation using other than Gaussian distributions has been considered, e.g., by [101]. He shows that if the distribution is misspecified, then consistency is no longer guaranteed. In particular, if a symmetric distribution is assumed but the true distribution is asymmetric, then maximum likelihood estimators are inconsistent. In practice, a common distribution used for maximum likelihood estimation is the Student  $t$  distribution. Given the results of [101], one should be careful in interpreting parameter estimates if there is evidence for skewness in standardized residuals.

Another estimation strategy based on maximum likelihood is a nonparametric estimation of the error density, which has been advocated by [48]. They suggest to use a first stage estimator of the model parameters, which is consistent but not efficient such as the Gaussian MLE, to construct residuals and then to use nonparametric methods to estimate the error density. Given the estimated error density, one can maximize the likelihood corresponding to this nonparametric density function. These estimators will under regulatory conditions be consistent and more efficient than the Gaussian ML estimator, provided that the true density is different from Gaussian.

A potential practical problem of maximum likelihood estimators is its dependence on numerical optimization routines. Recently, a closed form estimator based on the autocorrelation structure of squared returns has been suggested by [82]. Their estimator is inefficient compared to ML but has the advantage of being uniquely determined by the data. Further Monte Carlo evidence is necessary to see whether it is a serious practical competitor for ML-type estimators. Least squares type estimators of ARCH( $q$ ) have been considered by [118] and [103]. Again, these are inefficient compared with maximum likelihood estimators

but simpler to compute. [104] suggest a least absolute deviation estimator for GARCH models that is robust with respect to outliers but does not allow for a closed form. Finally, Bayesian estimation of GARCH-type models has been investigated, e.g., by [12,115] and [32].

## Testing for ARCH

In a regression such as

$$y_t = \mu_t + \varepsilon_t \quad (19)$$

where  $\mu_t$  is measurable w.r.t.  $\mathcal{F}_{t-1} = \sigma(y_{t-1}, y_{t-2}, \dots)$  and  $\varepsilon_t$  is a white noise sequence, inference on  $\mu_t$  typically depends on the properties of the error term  $\varepsilon_t$ . For example, if  $\varepsilon_t$  is i.i.d. Gaussian and  $\mu_t$  is linear such as an AR( $p$ ) model, then estimation by least squares of the autoregressive coefficients in  $\mu_t$  is efficient. If, however,  $\varepsilon_t$  is not i.i.d. and for example conditionally heteroskedastic, then estimation by ordinary least square (OLS) is no longer efficient and some kind of generalized least squares may be employed. Moreover, inference on the parameters in  $\mu_t$  will be erroneous if homoskedasticity of  $\varepsilon_t$  is assumed but, in reality,  $\varepsilon_t$  is conditionally heteroskedastic. In particular, standard errors in that case are typically underestimated. To avoid this, it is essential to test for ARCH type effects in  $\varepsilon_t$ . The following testing procedure, based on the Lagrange multiplier principle, has been proposed in the original ARCH paper by [44]. The null hypothesis is that  $\varepsilon_t$  is i.i.d. white noise, the alternative is the presence of ARCH. One first estimated the model (19) by least squares, obtains residuals  $\hat{\varepsilon}_t$ , and then runs the regression

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2 + \eta_t \quad (20)$$

where  $\eta_t$  is an error term. Under the null hypothesis  $H_0$ ,  $\alpha_1 = \dots = \alpha_q = 0$ . The test statistic is  $\lambda = nR^2$ , where  $n$  is the sample size and  $R^2$  the coefficient of determination of the regression (20). Under  $H_0$ , the test statistic follows asymptotically a  $\chi^2$  distribution with  $q$  degrees of freedom. Hence, it is an elementary exercise to test for ARCH effects in the error term of regression models. Historically, it is remarkable that prior to the introduction of the ARCH model, the above LM test was used by Prof. Clive Granger as an LM test for a bilinear error term, for which it has some power. Then, Prof. Robert Engle discovered that it has more power for another model, which he then introduced as the ARCH model.

An alternative to the LM test of [44] would be a Wald-type test of the hypothesis  $H_0: \alpha = 0$  in the GARCH(1,1) model (5) using, e.g., the t-ratio as test statistic. However, this test is non-standard since under the null hypothesis

the parameter  $\alpha$  is on the boundary of the parameter space and the parameter  $\beta$  is not identified. [6] treats this test in a general framework.

## Asymmetry, Long Memory, GARCH-in-Mean

In the standard GARCH model in (7), positive and negative values of lagged returns  $y_{t-i}$  have the same impact on volatility, since they appear in squares in the equation for  $\sigma_t^2$ . Empirically, it has been frequently noted since [17] that for stock markets, negative returns increase volatility more than positive returns do. Essentially, this so-called leverage effect means that negative news have a stronger impact on volatility than positive ones. To account for this empirical observation, several extensions of the standard GARCH model have been proposed in the literature. The most commonly used are the exponential GARCH model of [99] and the threshold GARCH model of [59] and [123]. The threshold model in its first order variant is given by the process

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \alpha^* y_{t-1}^2 I(y_{t-1} < 0) + \beta \sigma_{t-1}^2$$

where  $\alpha^*$  is an additional parameter and  $I(\cdot)$  is the indicator function. If  $\alpha^* = 0$ , then the threshold model reduces to the standard GARCH model. If  $\alpha^* > 0$ , then negative returns have a stronger impact on volatility than positive ones, which corresponds to the empirical observation for stock markets.

Secondly, the exponential GARCH (EGARCH) model of [99] specifies log-volatility as

$$\log \sigma_t^2 = \omega + \theta \xi_{t-1} + \alpha(|\xi_{t-1}| - E|\xi_{t-1}|) + \beta \log \sigma_{t-1}^2$$

where  $\xi_t = y_t/\sigma_t$  is i.i.d. with a generalized error distribution (GED) which nests the Gaussian and allows for slightly fatter tails. Due to the specification of log-volatility, no parameter restrictions are necessary to keep volatility positive. Moreover, the conditions for weak and strong stationarity coincide. Note that if  $\theta \neq 0$ , then  $\text{Cov}(y_t^2, y_{t-j}) \neq 0$  such that a leverage effect can be captured. A drawback of the EGARCH model is that asymptotic theory for maximum likelihood estimation and inference under primitive conditions are not available yet, but [110] are making much progress in this respect.

Another model allowing for asymmetry is the asymmetric power ARCH (APARCH) model [36]. In its (1,1) order form it specifies volatility as

$$\sigma_t^\delta = \omega + \alpha(|y_{t-1}| - \phi y_{t-1})^\delta + \beta \sigma_{t-1}^\delta$$

where  $\delta$  is a positive parameter. If  $\delta = 2$  and  $\phi = 0$ , the standard GARCH model is retained. For  $\phi \neq 0$ , there is

an asymmetric impact of positive and negative lagged returns on volatility. The additional flexibility due to the parameter  $\delta$  allows to better reproduce the so-called ‘Taylor property,’ originally noted by [111], which says that the autocorrelations of  $|y_t|^d$  are positive even at long lags, and when viewed as a function of  $d$  take a maximum for  $d \approx 1$  for many financial returns  $y_t$ . [70] provide a formal discussion of this issue.

The standard GARCH( $p, q$ ) model (7) implies that the decay of the autocorrelation function (ACF) of squared returns is geometrically fast. However, one often finds evidence for a slow hyperbolical decay in financial time series, see for example Fig. 3. The decay pattern of the ACF is related to the structure of coefficients  $c_j$  in the ARCH( $\infty$ ) representation of GARCH models,

$$\sigma_t^2 = c_0 + \sum_{j=1}^{\infty} c_j y_{t-j}^2. \quad (21)$$

For example, in the GARCH(1,1) model, these are given by  $c_j = \alpha\beta^{j-1}$ . Covariance stationary GARCH models have the property that the autocovariance function of squared returns,  $\gamma(\tau) = \text{Cov}(y_t^2, y_{t-\tau}^2)$ , is absolutely summable, i.e.,  $\sum_{\tau} |\gamma(\tau)| < \infty$ . Such a property is commonly called *short memory* as opposed to *long memory* processes for which the ACF is not absolutely summable. Long memory GARCH models have been proposed by [8] and [19], see also the review of long memory processes in econometrics by [7]. An example of a long memory GARCH process would be given by (21) with  $c_j = C j^{-\theta}$  for some constant  $C$  and parameter  $\theta > 0$ . A particular example for such a process is the fractionally integrated GARCH (FIGARCH) model of [8], which can be written as

$$(1 - L)^d \sigma_t^2 = \omega + \alpha y_{t-1}^2$$

where  $L$  is the lag operator and  $d$  a positive parameter. When  $d = 1$  one obtains the integrated GARCH (IGARCH) model of [47]. For  $d \neq 1$  one can use a binomial extension to obtain after inverting

$$(1 - L)^{-d} = \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} L^j = \sum_{j=0}^{\infty} c_j L^j \quad (22)$$

where  $\Gamma(\cdot)$  is the Gamma function. The coefficient  $c_j$  in (22) can be shown to be of the long memory type. A similar long memory EGARCH model has been introduced by [19]. The drawback of these particular specifications is that they share the property with the IGARCH model to have infinite variance. [105] has proposed a long memory GARCH type model that allows for finite variance.

Finally, in the finance literature a link is often made between the expected return and the risk of an asset, since investors are willing to hold risky assets only if their expected returns compensate for the risk. A model that incorporates this link is the GARCH-in-mean or GARCH-M model of [52], given by

$$y_t = \delta g(\sigma_t^2) + \varepsilon_t$$

where  $\varepsilon_t$  is an ARCH or GARCH error process,  $\delta$  a parameter, and  $g$  a known function such as square root or logarithm. If  $\delta > 0$  and  $g$  is monotone increasing, then the term  $\delta g(\sigma_t^2)$  can be interpreted as a risk premium that increases expected returns  $E[y_t]$  if volatility  $\sigma_t^2$  is high. It can be shown that such a model, when applied to the market index, is consistent with the capital asset pricing model (CAPM) of [108] and [87], see [24].

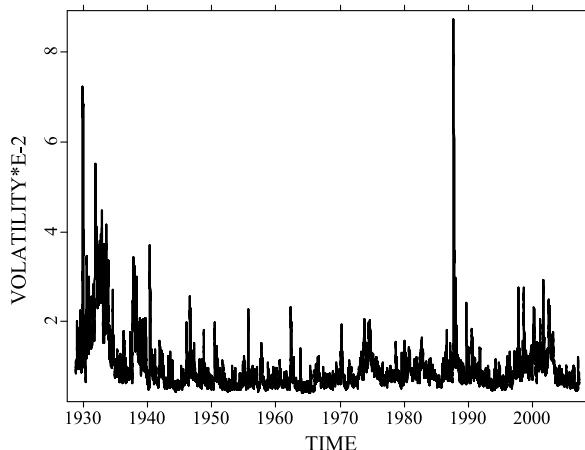
As an empirical illustration we estimate alternative models for the Dow Jones index discussed in the introduction. To recall, we have daily returns from October 1928 to April 2007. First order autocorrelation of returns is 0.03, which due to the large number of observations is significant at the level 1%. However, we refrain here from fitting an autoregressive or moving average model to the returns as the results concerning volatility estimation do not change substantially. We only consider a constant conditional mean in the model  $y_t = \mu + \varepsilon_t$ , where  $\varepsilon_t$  is one of the discussed GARCH-type models and  $\mu$  takes into account a non-zero trend in returns. Six alternative GARCH models are considered, all of them being of order (1,1): standard GARCH, TGARCH, EGARCH, GARCH-M, TGARCH-M and EGARCH-M. For the “in-mean” versions, we have chosen the square root specification for the function  $g(\cdot)$ , which seems to work better than the logarithm or the identity function. Moreover, for all “in-mean” models the constant  $\mu$  turned out to be insignificant and hence was suppressed from the model. Table 1 summarizes the estimation results.

Note first that all estimated risk premia are positive  $\mu > 0$  and  $\delta > 0$ , as theory would predict. Second, for all models allowing for asymmetry, the leverage effect of negative returns is confirmed, i.e.,  $\alpha^* > 0$  for the TGARCH models and  $\theta < 0$  for the EGARCH models. Third, in all cases persistence of shocks to volatility is very high, measured by  $\alpha + \beta$  in the GARCH model,  $\alpha + \alpha^*/2 + \beta$  in the TGARCH model (assuming a symmetric innovation distribution), and by  $\beta$  in the EGARCH model. Thus, all models are short memory with exponential decay of the ACF of squared returns, but the models try to adapt to the empirically observed slow decay of the ACF by pushing the persistence parameter close to one. This near-IGARCH

**GARCH Modeling, Table 1**

Estimation results for the following models: GARCH, EGARCH, TGARCH, GARCH-in-mean, TGARCH-in-mean and EGARCH-in-mean, applied to daily Dow Jones returns from 1928 to 2007. All parameters are significant at the one percent level

	G	TG	EG	GM	TGM	EGM
$\mu$	4.25E-04	2.64E-04	2.36E-04			
$\delta$				0.0591	0.0354	0.0286
$\omega$	8.74E-07	1.03E-06	-0.2184	8.75E-07	1.06E-06	-0.2227
$\alpha$	0.0766	0.0308	0.1391	0.0766	0.0306	0.1378
$\alpha^*$		0.0769			0.0761	
$\theta$			-0.0599			-0.0595
$\beta$	0.9173	0.9208	0.9879	0.9172	0.9205	0.9874
$L$	65456.47	65600.53	65589.32	65462.45	65603.79	65590.69

**GARCH Modeling, Figure 8**

Estimated conditional standard deviation of daily Dow Jones index returns, 1928 to 2007, using the TGARCH-in-mean model

behavior is typical for daily returns. Finally, the goodness-of-fit seems to be best for the TGARCH-in-mean model, taking the log-likelihood as criterion. The estimation results strongly confirm the presence of the leverage effect, high persistence, and positive risk premium in the data. Figure 8 shows the estimated conditional standard deviation of the TGARCH-M model. For the other models, the graph would look quite similar and is therefore not shown here. Notice the very volatile periods at the beginning of the sample in the 1930s, around the year 2000 corresponding to the “new economy” boom and following crash, as well as the spike in 1987 due to the crash of October 17, 1987.

**Non- and Semi-parametric Models**

Nonparametric methods refrain from associating particular parametric forms to functions or distributions. Instead,

only the class of functions is determined, for example the class of squared integrable functions or the degree of smoothness. The price for the flexibility is typically slower convergence rates than parametric models. A combination of the two approaches is often called semiparametric. One such approach has already been mentioned in Sect. “[Estimation and Inference](#)” in the context of estimation by maximum likelihood using nonparametric estimates of the error density, as proposed by [48] for GARCH models. [88] shows that this procedure leads to adaptive estimation of the identifiable parameters ( $\alpha$  and  $\beta$  in a GARCH(1,1) model) in the sense of [16]. That is, it is possible to achieve the Cramer–Rao lower bound and do as good as if one knew the true error distribution. The scale parameter  $\omega$ , however, is not adaptively estimable. See also [40] and [37] for related results for univariate GARCH models, and [65] for an extension to semiparametric estimation of multivariate GARCH models.

A different approach is to directly model the volatility process in a nonparametric way. Early models were proposed by [61] and [51]. The qualitative threshold ARCH model of [61] specifies models of the type

$$y_t = \sum_{j=1}^J \sigma_j I(y_{t-1} \in A_j) \xi_t, \quad (23)$$

where  $(A_j)$  is a partition of the real line and  $\sigma_j, j = 1, \dots, J$ , are positive parameters. Thus, volatility is modeled as a piecewise constant function of lagged returns. Note that the threshold ARCH model of [59] and [123] is not a special case of (23) as there the volatility function is piecewise quadratic in lagged returns. Extensions to the ARCH( $q$ ) and GARCH( $p, q$ ) are straightforward. [51] replaced the piecewise constant functions by piecewise linear functions. In both cases, one may consider their models as nonparametric if the partition becomes finer as the sample size increases.

Consider the model

$$y_t = \sigma(y_{t-1})\xi_t$$

where  $\sigma(\cdot)$  is an unknown smooth function, and  $\xi_t \sim \text{i.i.d. } N(0, 1)$ . For  $\sigma^2(x) = \alpha x^2$  one obtains the parametric ARCH(1) model of [44]. An example of a nonparametric estimator of  $\sigma(\cdot)$  is the Nadaraya–Watson estimator given by

$$\hat{\sigma}^2(x) = \frac{\sum_{t=2}^n K\{(y_{t-1} - x)/h\}y_t^2}{\sum_{t=2}^n K\{(y_{t-1} - x)/h\}}$$

where  $K$  is a kernel function satisfying  $\int K(x)dx = 1$  and  $\int xK(x)dx = 0$ , and where  $h > 0$  is a bandwidth that determines the degree of smoothing. As the Nadaraya–Watson estimator can be interpreted as fitting a constant locally, a generalization consists of fitting a local polynomial instead. This has been derived by [68] for the volatility case.

A general problem of nonparametric methods is the so-called *curse of dimensionality* when smoothing has to operate in high dimensions. Considering a nonparametric ARCH( $q$ ) model,

$$y_t = \sigma(y_{t-1}, \dots, y_{t-q})\xi_t$$

this problem is apparent and in practice very large data sets are required to estimate the function  $g$  with appropriate precision. One may be inclined to impose more structure on the  $g$  function such as additive or multiplicative separability. Nonparametric multiplicative ARCH models have been proposed by [63] and [120]. Semi-parametric additive ARCH models of the type  $\sigma^2(y_{t-1}, \dots, y_{t-q}) = \sum_{j=1}^p \beta^{j-1} g(y_{t-j})$  with some unknown function  $g$  and parameter  $\beta \in (0, 1)$  have been considered by [29].

Extension of nonparametric ARCH( $q$ ) models to nonparametric GARCH(1,1) models have also been proposed. However, in its general form  $y_t = \sigma_t \xi_t$  with  $\sigma_t = g(y_{t-1}, \sigma_{t-1})$ , the model is difficult to estimate due to lack of structure. One might consider iterative estimation algorithms, based on some initial estimate of volatility as in [26].

Imposing a similar semi-parametric structure as for the semi-parametric ARCH( $q$ ) model of [29], one can write  $\sigma_t^2 = g(y_{t-1}) + \beta \sigma_{t-1}^2$ , where again  $g(\cdot)$  is an unknown smooth function. Note that this model nests many of the proposed parametric models. It has been considered by [119] and [89].

In practice, nonparametric methods may be used whenever it is not a priori clear what functional form fits best the data, either by using them directly, or as a tool to specify a parametric model in a second stage.

## Multivariate GARCH Models

In economics and finance, one typically deals with multiple time series that are fluctuating in a non-systematic manner and are considered as realizations of stochastic processes. The interest for applied econometricians is therefore to model their risk, that is, their volatility, but also their inter-dependencies. For example, if one has reasons to assume that the underlying stochastic processes are Gaussian, then the inter-dependencies may be completely described by the correlation structure. In fact, when we say ‘multivariate volatility models’ we usually mean the modeling of volatilities but also that of correlations. This is also the reason why the extension of univariate volatility to multivariate volatility models is much more complex than that of univariate models for the conditional mean, such as ARMA models, to the multivariate case.

It will be immediately clear that the multivariate case is the one that is by far more relevant in practice when financial markets are under study. The reason is, first, the large number of different assets, or even different types of contracts, assets, exchange rates, interest rates, options, futures, etc. Second, there is usually a strong link between these variables, at least within one group. For example, asset returns in one stock market tend to be quite strongly correlated. One would make big approximation errors when treating the variables as independent by, e.g., using univariate volatility models for the conditional variances and set conditional covariances to zero. Note that, setting conditional covariance to zero is much stronger an assumption than setting the unconditional covariance to zero. Ways must be found to treat the dependence of the series in a flexible yet parsimonious way.

A first step would again be to do exponential smoothing à la RiskMetrics, which can be used not only to obtain the individual variances according to (4), but also to obtain the correlations. To see this, we define just as in (4) an exponential smoother for the covariances as

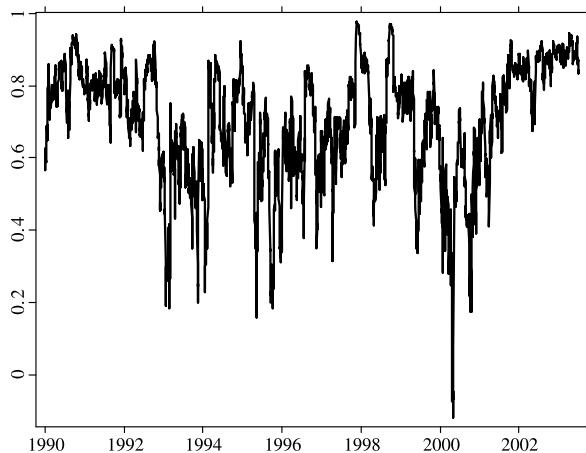
$$\sigma_{12,t} = (1 - \lambda)\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \lambda\sigma_{12,t-1}$$

and then obtain as usual the conditional correlation as

$$\rho_t = \frac{\sigma_{12,t}}{\sigma_{1,t}\sigma_{2,t}},$$

which is guaranteed to be between minus one and one if the same parameter  $\lambda$  is used, typically  $\lambda = 0.94$ . Figure 9 depicts the RiskMetrics conditional correlation series for the DOW and NASDAQ return series.

Obviously, conditional correlations are not constant, although it is difficult from the graph to verify such a statement statistically. However, one thing to observe is that



**GARCH Modeling, Figure 9**

Conditional correlations of the Dow Jones IA and NASDAQ index returns, daily, using the RiskMetrics model

during the New Economy boom in 1999 and 2000, estimated correlations have been substantially lower, sometimes even negative, than at other times. The reason is probably a decoupling due to the higher vulnerability of the NASDAQ with respect to the bubble in high tech and internet stocks. A more thorough analysis of this data set which also compares this model with other, more flexible models is provided by [45]. We see that the RiskMetrics tool, even though very simple, can give some guidelines. One of the objectives of the econometrician is to enhance the model in terms of flexibility (e.g., why should  $\lambda$  be fixed to 0.94?), and to establish a statistical framework in which hypotheses such as constant conditional correlations can be tested.

From an econometrical viewpoint, modeling the volatility of multiple time series is, for several reasons, challenging both theoretically and practically. For the sake of illustration, consider a bivariate GARCH model of the Vec type that was introduced by [24]. Denote by  $H_t$  the conditional variance matrix of the asset return vector  $y_t$ . Then a bivariate ARCH(1) model reads

$$H_t = \begin{pmatrix} h_{1,t} & h_{12,t} \\ h_{12,t} & h_{2,t} \end{pmatrix}$$

and where

$$\begin{aligned} h_{1t} &= \omega_1 + \alpha_{11}\varepsilon_{1,t-1}^2 + \alpha_{12}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \alpha_{13}\varepsilon_{2,t-1}^2 \\ h_{12,t} &= \omega_2 + \alpha_{21}\varepsilon_{1,t-1}^2 + \alpha_{22}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \alpha_{23}\varepsilon_{2,t-1}^2 \\ h_{2t} &= \omega_3 + \alpha_{31}\varepsilon_{1,t-1}^2 + \alpha_{32}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + \alpha_{33}\varepsilon_{2,t-1}^2. \end{aligned}$$

Each conditional variance,  $h_{1t}$  and  $h_{2t}$ , and conditional covariance,  $h_{12,t}$ , depends on all lagged squared returns

(two in the bivariate case) and all lagged cross-products (one in the bivariate case). The main reason for the rapidly increasing complexity of the model when the dimension is increased lies in the fact that not only all conditional variances with their cross-dependencies have to be modeled, but also all conditional correlations. It is in fact the latter that poses the main problem, as there are a total of  $N(N - 1)/2$  such correlations when the dimension is  $N$ , but only  $N$  variances. Thus, modeling variances and correlations simultaneously, a total of  $N(N + 1)/2$  entries of the conditional covariance matrix need to be modeled. For example, if  $N = 10$  (a moderate dimension for many economic or financial problems) this number is 55, if  $N = 100$  (such as modeling all stocks of a common stock index), then 5050 series, conditional variances and covariances, are at stake.

It is clear that this is too much to allow for a flexible cross-dependence of the individual series. Without imposing any structure except linearity, the multivariate generalization of the standard GARCH model, the so-called Vec model introduced by [24], is feasible only for low dimensions, two or three say, as otherwise the number of parameters becomes too high relative to the number of observations typically encountered in economic practice. Another problem is that the Vec model does not guarantee a positive definite covariance matrix. Necessary conditions for the latter desirable property are as yet unknown in the general Vec specification.

These are some reasons to look for other models, and in fact, over recent years a broad variety of different approaches to the problem have been suggested in the literature. Roughly speaking, one can divide them into two groups. The first one tries to simplify the problem by imposing more structure on the Vec model. Examples are the BEKK model of [49] and the factor GARCH model by [53]. More recently, the second group tries to separate the problem of modeling the conditional variances and conditional correlations. An early and simple version of this group is to say that conditional variances are just univariate GARCH and conditional correlations are constant over time, as suggested by [22]. In its simplicity, this constant conditional correlation (CCC) model basically does not add any complexity beyond univariate GARCH to the multivariate estimation problem, which renders the model extremely useful in empirical practice. It also introduced the idea of two-step estimation, where in the first step conditional variances are modeled, and in the second step the conditional correlations using the standardized residuals of the first step. However, starting with [45] there have been plenty of arguments in favor of time-varying conditional correlations in financial markets. In par-

ticular, a common finding is that correlations are higher when the market moves up than when it moves down. A test for this correlation asymmetry has been suggested by [73]. Using a dynamic conditional correlation model (DCC), [45] shows that time varying correlations are not uncommon even in normal market situations. In the following we sketch these two branches of the multivariate GARCH literature. It should however be mentioned that there are models that do not fall into these two categories such as a multivariate version of the exponential GARCH model proposed by [79].

### Factor GARCH Models

In the following factor GARCH models are discussed as an example of multivariate GARCH models. The main idea of factor models is to reduce the dimension of the system to a tractable two or three factors, which can then be modeled in a standard way. It should be noted that also ‘full-factor’ models with number of factors equal to the number of variables have been proposed in the literature. For example, [116] propose the model

$$y_t = Wf_t$$

where  $W$  is a  $N \times N$  parameter matrix and  $f_t$  is a  $N$ -vector with conditional mean zero and diagonal conditional variance matrix,  $\Sigma_t$  say. The individual conditional variances of  $f_t$  can be modeled by univariate GARCH(1,1), for example. One can restrict  $W$  to be lower triangular, as it is well known that the Choleski decomposition of a positive definite matrix always exists and is unique. Thus, the conditional variance matrix of  $y_t$  is given by  $H_t = W\Sigma_t W' = L_t L_t'$ , where  $L_t = W\Sigma_t^{1/2}$  is lower triangular. In this model, the parameters in  $W$  and those in  $\Sigma_t$  need to be estimated jointly, which may be cumbersome in high dimensions. The empirical performance of such full factor models still remains to be investigated.

It is more common to specify only a few factors and allow for idiosyncratic noise. We will look at such models in the following. Suppose that there are  $K$  (observed or unobserved) factors, collected in a  $K$ -vector  $f_t$ , with  $K < N$ . Then a simple factor model can be written as

$$y_t = Wf_t + \nu_t \quad (24)$$

where  $\nu_t$  is a white noise vector with  $\text{Var}(\nu_t) = \Omega$  that represents the idiosyncratic noise. Typically, one assumes that  $\Omega$  is diagonal so that components of the idiosyncratic noise are uncorrelated. In that case, correlation between components of  $y_t$  is induced only through the common factors  $f_t$ . If  $y_t$  represents the error of a time series system,

one may constrain  $f_t$  to have conditional mean zero. The matrix  $W$  is of dimension  $N \times K$ , of full column rank, and contains the so-called *factor loadings*, the weights of a factor associated with the individual components of  $y_t$ .

In finance, model (24) is well known from the arbitrage pricing theory (APT) of [106], where  $y_t$  are excess returns of financial assets,  $f_t$  are systematic risk factors and  $\nu_t$  is unsystematic risk. It can also be viewed as a generalization of the capital asset pricing model (CAPM) developed by [108] and [87]. For simplicity we assume here that factors are observed. If they are unobserved, identification issues arise that are discussed, e. g., by [107].

For the factors, a low-dimensional GARCH model can be assumed:  $\text{Var}(f_t | \mathcal{F}_{t-1}) = \Sigma_t$ , where  $\Sigma_t$  is a  $(K \times K)$  covariance matrix. The conditional covariance matrix of  $y_t$  is given by

$$H_t = \text{Var}(y_t | \mathcal{F}_{t-1}) = W\Sigma_t W' + \Omega. \quad (25)$$

In the case of just one factor, the matrix  $W$  reduces to a vector  $w$  and the factor volatility,  $\sigma_t^2$  say, can be modeled by univariate GARCH and the conditional variance of  $y_t$  simplifies to

$$H_t = ww' \sigma_t^2 + \Omega.$$

If the factors are conditionally uncorrelated, i. e.,  $\Sigma_t$  is diagonal with  $\Sigma_t = \text{diag}(\sigma_{1t}^2, \dots, \sigma_{Kt}^2)$ , then one can write

$$H_t = \sum_{k=1}^K w_k w_k' \sigma_{kt}^2 + \Omega$$

where  $w_k$  is the  $k$ th column of  $W$ . [83] propose methods to test for the number of factors  $K$  and derive results for maximum likelihood estimation. For the more general BEKK model class, [31] derived asymptotic theory but assuming moments of order eight of the process, which may exclude many of the typically fat-tailed financial time series.

A popular factor GARCH model is the orthogonal GARCH (OGARCH) model of [3]. In the OGARCH model, factors  $f_t$  are the  $K$  largest principal components obtained from the (unconditional) sample covariance matrix, and the loading matrix  $W$  is the matrix of associated eigenvectors. The loadings represent the sensitivity of an individual series on a specific factor. By construction, the unconditional correlation between the factors is zero, due to the orthogonality of the principal components. However, the *conditional* correlation may be different from zero. Denote the (empirical) covariance matrix of  $y_t$  by  $\Sigma$ . The decomposition  $\Sigma = \Gamma\Lambda\Gamma'$  gives  $\Gamma = (\gamma_1, \dots, \gamma_N)$  with the eigenvectors  $\gamma_i$  and  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N)$  with corresponding eigenvalues  $\lambda_i$ . We order the columns of  $\Gamma$

according to the magnitude of the corresponding eigenvalues such that  $\lambda_1 > \lambda_2 > \dots > \lambda_N$ . Let us assume here that all eigenvalues are distinct, otherwise  $\Gamma$  may not be identified. For the case of non-distinct eigenvalues, one may use the more general singular value decomposition and go for the GO-GARCH (generalized orthogonal GARCH) model of [114].

The vector of principal components, given by

$$\begin{pmatrix} f_t \\ \varepsilon_t \end{pmatrix} = \Gamma' y_t$$

is partitioned into the first  $K$  components  $f_t$ , whose volatility will assumed to be stochastic, and the last  $N - K$  components  $\varepsilon_t$ , whose volatility will assumed to be constant. One could speak of  $K$  *dynamic* and  $N - K$  *static* factors.

Now decompose the matrices as follows:

$$\Gamma_{(N \times N)} = (\Gamma_{1(N \times K)}, \Gamma_{2(N \times (N-K))}),$$

and  $\Lambda_1 = \text{diag}(\lambda_1, \dots, \lambda_K)$ ,  $\Lambda_2 = \text{diag}(\lambda_{K+1}, \dots, \lambda_N)$ .

The model can then be written as

$$y_t = \Gamma_1 f_t + \Gamma_2 \varepsilon_t \quad (26)$$

where  $\text{Var}(f_t | \mathcal{F}_{t-1}) = \Sigma_t$  and  $\text{Var}(\varepsilon_t | \mathcal{F}_{t-1}) = \Lambda_2$ . For example,  $\Sigma_t$  may be diagonal or some  $K$ -variate GARCH model. Note that this representation is equivalent to that of (24) with  $W = \Gamma_1$  and  $v_t = \Gamma_2 \varepsilon_t$ , except that  $\Omega = \Gamma_2 \Lambda_2 \Gamma_2'$  will not be diagonal in general. The conditional variance of  $y_t$  is given by

$$\begin{aligned} H_t &= \text{Var}(y_t | \mathcal{F}_{t-1}) = \Gamma_1 \Sigma_t \Gamma_1' + \Gamma_2 \Lambda_2 \Gamma_2' \\ &= \Gamma \begin{pmatrix} \Sigma_t & 0 \\ 0 & \Lambda_2 \end{pmatrix} \Gamma'. \end{aligned}$$

If  $\Sigma_t$  follows a  $K$ -variate BEKK process, then it can be shown that  $H_t$  will follow an  $N$ -variate BEKK process with restrictions on the parameter matrices. However, the classical OGARCH assumes that factors are conditionally orthogonal, hence  $\Sigma_t$  is diagonal, additional to the fact that they are unconditionally orthogonal by construction. This assumption is crucial and may not always be justified in practice. It should be emphasized that in the OGARCH model, the factor loadings contained in the matrix  $\Gamma$  and the factor variances contained in  $\Lambda$  are considered as fixed for a given sample covariance matrix. This contrasts the general factor model (24) where factor loadings  $W$  are estimated jointly with the parameters describing the factor dynamics.

Instead of using unconditionally orthogonal factors, [55] proposed to use conditionally orthogonal factors

by searching numerically for linear combinations of the data such that the conditional correlation between these combinations is minimized under norm constraints. The existence of such linear combinations is tested using bootstrap methods.

### Constant and Dynamic Conditional Correlation Models

[22] suggests a multivariate GARCH model with constant conditional correlations. Let  $H_t$  be the conditional covariance matrix of a series  $y_t$ , and  $V_t$  be a diagonal matrix with the conditional standard deviations of  $y_t$  on its diagonal. Then the model is simply

$$H_t = V_t R V_t \quad (27)$$

where  $R$  is the constant correlation matrix.  $H_t$  is positive definite as long as the conditional variances are positive and  $R$  is positive definite. For instance, one could specify univariate GARCH models for the individual conditional variances. On the other hand, it is possible to allow for spill-over of volatilities from one series to other series. Note that the CCC model is not nested in the Vec specification. Theory of maximum likelihood estimation for CCC-type models has been established by [77] for consistency and [86] for asymptotic normality.

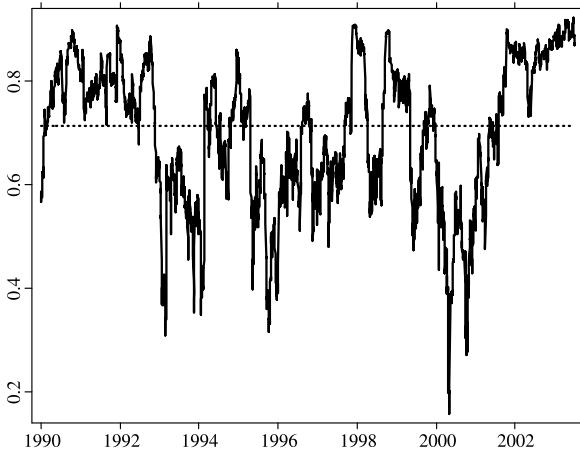
The assumption of constant correlations simplifies strongly the estimation problem. However, it might sometimes be too restrictive. For example, it is often observed that correlations between financial time series increase in turbulent periods, and are very high in crash situations. A Lagrange Multiplier test against the CCC model has been suggested by [112]. An extension of the CCC model to allow for time-varying correlations is the dynamic conditional correlations (DCC) model introduced by [45]. The DCC model renders the conditional correlation matrix  $R$  dependent on time,  $R_t$  say. The conditional correlation between the  $i$ th and  $j$ th component of  $y_t$  is modeled as

$$R_{ij,t} = \frac{Q_{ij,t}}{\sqrt{Q_{ii,t} Q_{jj,t}}}$$

where  $Q_{ij,t}$  is the  $ij$ th element of the matrix  $Q_t$  given by

$$Q_t = S(1 - \alpha - \beta) + \alpha v_{t-1} v_{t-1}' + \beta Q_{t-1} \quad (28)$$

where  $\alpha$  and  $\beta$  are parameters and  $v_t = V_t^{-1} y_t$  are the standardized but correlated residuals. That is, the conditional variances of the components of  $v_t$  are one, but the conditional correlations are given by  $R_t$ . The matrix  $S$  is the sample correlation matrix of  $v_t$ , so a consistent esti-



**GARCH Modeling, Figure 10**

Conditional correlations of the Dow Jones IA and NASDAQ index returns, daily, using the DCC model. Dashed line: constant conditional correlation

mate of the unconditional correlation matrix. If  $\alpha$  and  $\beta$  are zero, we get the above CCC model. If they are different from zero one gets a kind of ARMA structure for all correlations. Note however that all correlations would follow the same kind of dynamics, since the ARMA parameters are the same for all correlations. The specification of the first term of  $Q_t$  ensures that the unconditional mean of  $Q_t$  is equal to the sample covariance matrix of  $v_t$ , similar to the *variance targeting* technique of [50]. Also it facilitates the estimation, since that can be done in two steps: First, the conditional variances in  $V_t$  are estimated using univariate GARCH models, for example, then  $v_t$ , the standardized (but correlated) residuals and their covariance matrix  $S$  are computed, before in the second step only two remaining parameters,  $\alpha$  and  $\beta$ , need to be estimated. A model similar to DCC has been proposed by [113].

Figure 10 depicts the estimated conditional correlations for the DOW Jones and NASDAQ time series, using the DCC and CCC models. Comparing the former with the RiskMetrics estimates of Fig. 9, no substantial difference can be detected visually. However, the parameter estimates of  $\alpha$  and  $\beta$  are 0.0322 and 0.9541 with standard errors 0.0064 and 0.0109, respectively, so that the null hypothesis  $H_0: \alpha = 0.06$  is clearly rejected. Whether or not the difference in estimated conditional correlations matters in empirical applications has been addressed, e.g., by [30], who consider the problem of portfolio selection. [96] compare the performance of CCC, DCC, OGARCH and a model of [84] in forecasting and portfolio selection in high dimensions. They find that the difference is not substantial, but that the CCC model is too restrictive.

To summarize, the whole challenge of multivariate volatility modeling is to balance model complexity and simplicity in such a way that the model is flexible enough to capture all stylized facts in the second moments (and perhaps beyond that) of the series while keeping it simple for estimation and inference.

In the following we sketch some applications of multivariate GARCH models in finance. As an early example, [24] estimate a capital asset pricing model (CAPM) with time-varying betas. The beta is defined as the ratio of the asset return's covariance with the market return, divided by the variance of the market return. Denote by  $r_{it}$  the excess return of asset  $i$ , and by  $r_{mt}$  the excess return of the market. Then the beta-form of the CAPM can be written as

$$r_{it} = \beta_{it} r_{mt} + \varepsilon_i = \frac{\text{Cov}(r_{it}, r_{mt})}{\text{Var}(r_{mt})} + \varepsilon_i$$

where  $\varepsilon_i$  is idiosyncratic noise whose risk cannot be diversified away and is therefore called unsystematic risk. As we observe time varying second moments, it is clear that betas will also be time varying, not only due to the variance of the market but also due to the covariances of the assets with the market. However, if both returns are covariance stationary, then by definition the unconditional second moments will be constant, and only after conditioning on suitable information sets such as historical returns will second moments become time varying.

Secondly, correlations between exchange rates have been substantially time-varying, as for example in Europe the European exchange rate mechanism enforced increasing correlations. The correlation of the DEM/USD and FRF/USD rates, for instance, increased steadily in the late 1990s until it was virtually one just before the launch of the Euro. See, e.g., [45], who models these data, among others, with alternative correlation models. Thirdly, portfolio selection is another type of application. If, for example, one is interested in the minimum variance portfolio of  $n$  assets with covariance matrix  $\Sigma$ , then the well known formula for the optimal weight vector  $\alpha$  is given by

$$\alpha = \frac{\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota}$$

where  $\iota$  is an  $n$ -vector of ones, see, e.g., [30]. Obviously, if  $\Sigma$  is allowed to be time-varying, then the optimal portfolio weights will in general also depend on time. This has many important practical implications, e.g., for portfolio managers. One of the problems is to determine an optimal reallocation frequency. If variances and covariances change daily and the objective is to minimize the portfo-

lio variance over the next ten days, then one could follow at least two strategies: either calculate the optimal portfolio weights daily and reallocate accordingly. Or, calculate the return distribution over ten days, obtain thus a covariance matrix for ten-day returns, find the optimal weights using this covariance matrix and leave the corresponding portfolio unchanged throughout the ten days. If the objective is to minimize the variance over the ten days, then the first method will usually outperform the second. The intuitive reason is that the second method aggregates data, thus losing valuable information. However, in practice one may still prefer the second method for various reasons, one of which could be the higher transaction costs of the first method.

### Stochastic Volatility

GARCH models discussed so far explain the conditional variance at time  $t$  as a function of the information set at time  $t - 1$ . In other words, it is measurable with respect to this information set. This is not the case for models of the stochastic volatility (SV) type, which introduce an extra error term in the volatility equation. For example, in the univariate case such a model could take the form

$$\begin{aligned} y_t &= \sigma_t \xi_t \\ \log \sigma_t^2 &= \omega + \beta \log \sigma_{t-1}^2 + \eta_t \end{aligned} \quad (29)$$

where  $\xi_t$  and  $\eta_t$  are i.i.d. mean zero random variables with variance equal to one and  $\sigma_\eta^2$ , respectively. Here, log volatility follows an AR(1) process. Since volatility is unobserved, model (29) is a particular case of a latent variable model. Note that, if the information set at time  $t - 1$  consists of all lagged values of  $y_t$  up to  $y_{t-1}$ , then volatility at time  $t$  is not measurable with respect to this information set. [27] compare moment properties such as kurtosis and persistence of SV and GARCH models. [58] propose a model that encompasses both GARCH and stochastic volatility and thus allows for testing against each of them.

Estimation is more complicated than for GARCH models because the likelihood is an integral of dimension equal to the sample size, given by

$$L(Y; \theta) = \int p(Y | H, \theta) p(H | \theta) dH \quad (30)$$

where  $Y = (y_1, \dots, y_n)$ ,  $H = (\sigma_1^2, \dots, \sigma_n^2)$ , and  $\theta = (\omega, \beta, \sigma_\eta^2)$ . Maximization of (30) has no closed form and numerical optimization is difficult due to the high dimensional integral. Therefore, other estimation methods have been considered in the literature, for example generalized method of moments (GMM), simulated maximum

likelihood with Markov Chain Monte Carlo (MCMC) or Bayesian methods, see e.g., [75,76] and [80]. An application to currency options by [92] compares three alternative estimation algorithms and finds that the estimation error of the volatility series is large for all methods.

In the multivariate case, without imposing structure, estimating a highly dimensional stochastic volatility model seems difficult. One way of imposing structure in multivariate SV models is to assume a factor model as, e.g., in [34,69,81] and [56], or constant correlations. To consider a bivariate extension of stochastic volatility models, one suggestion of [69] is to say that the stochastic variances  $\sigma_{1,t}$  and  $\sigma_{2,t}$  of the two assets follow univariate stochastic variance processes as in (29), and the stochastic covariance is given by

$$\sigma_{12,t} = \rho \sigma_{1,t} \sigma_{2,t},$$

where  $\rho$  is a constant parameter between  $-1$  and  $1$ . This model, very much in the spirit of the constant conditional correlation GARCH model of [22], is quite parsimonious and can be efficiently estimated using simulated maximum likelihood as demonstrated in [33]. It is straightforward to generalize this specification to higher dimensions. However, estimation may then become trickier. Also the restriction of constant correlation parameters may not be innocuous. More empirical tests are required about goodness of fits, comparing the non-nested GARCH and SV type models of about the same model complexity.

SV models lend themselves naturally to continuous time stochastic volatility models and realized volatility. Indeed, as shown by [9], realized volatility can be used to estimate the volatility of SV models. The monograph of [109] collects influential papers of the stochastic volatility literature.

### Aggregation

The frequency at which financial time series are sampled is often not unique. For example, one researcher may be interested in the behavior of returns to the Dow Jones index at a daily frequency, but another one at a weekly or monthly frequency. Considering log-returns, weekly returns can be directly obtained from daily returns by simply summing up intra-week returns. If a model is fitted to daily returns, an important question is what this implies for the weekly returns. In particular, one may ask if the model remains in the same class, which would then be called closed under temporal aggregation. For the univariate GARCH model, [38] have shown that only a weak version of it is closed under temporal aggregation. Instead

of modeling the conditional variance, weak GARCH models the best linear predictor of squared returns in terms of a constant, lagged returns and lagged squared returns. In the weak GARCH(1,1) case, they show how to obtain the parameters of the aggregated process as a function of the parameters of the high frequency process. In particular, denoting the parameters of the aggregated process by  $\alpha^{(m)}$  and  $\beta^{(m)}$ , where  $m$  is the aggregation level, then the persistence parameter of the aggregated level is given by  $\alpha^{(m)} + \beta^{(m)} = (\alpha + \beta)^m$ . Thus, the persistence of the aggregated process declines geometrically fast with the aggregation level. Asymptotically, the process will reduce to white noise. One would therefore expect to see much less conditional heteroskedasticity in monthly returns than in weekly or daily returns. The link between parameters at different frequencies also provides a means for model diagnostics. The results of [38] have been extended to the multivariate case by [64].

Instead of aggregating, one could go the other way and look at “disaggregating” the process temporally, i.e., sampling the underlying process at finer intervals. [97] showed that GARCH models can be viewed as approximations of continuous time stochastic volatility models, see also [39]. However, [117] has shown that the GARCH model and its diffusion limit are not equivalent in a statistical experiment sense.

Rather than aggregating temporally, one may alternatively be interested in aggregating contemporaneously in a multivariate context. For example, stock indices are constructed as linear combinations of individual stocks. [102] show that again the aggregated process is only weak GARCH. Rather than aggregating multivariate GARCH models, one can alternatively consider aggregation of univariate heterogenous GARCH processes with random coefficients. In linear ARMA models, this aggregation scheme is known to produce long memory type behavior of the aggregate, see [62]. [35] conjectured that this holds in a similar way for GARCH models. However, [122] shows that although the ACF of the squared aggregate decays hyperbolically, it may be absolutely summable and hence there is no long memory. For the general model class of [94] which includes GARCH, weak GARCH and stochastic volatility as special cases, [121] shows that contemporaneous aggregation leads to long memory properties of the aggregate.

## Future Directions

The theory of univariate GARCH models is now well developed and understood. For example, theory of maximum likelihood estimation is available under weak con-

ditions that allow for integrated and even mildly explosive processes. However, theory of multivariate GARCH is still in its infancy and far from closed, due to arising technical difficulties. For general specification such as the BEKK model, no results on asymptotic normality of estimates are available yet that would allow for integrated processes. Most available results on general specifications are high level and only for some special cases, primitive conditions are established. This is certainly one of the main directions for future research.

On the modeling side, there is no clear general answer how to deal with the problem of high dimensions, and in particular how to balance model flexibility with econometric feasibility. More practical experience is necessary to see what type of model performs best for what kind of data. On the application side, a still open issue is how to evaluate the volatility risk for option pricing, and how to efficiently use multivariate GARCH models in portfolio selection or risk management. Other frontiers for GARCH models are discussed by [46].

An interesting new field is the combination of GARCH models with nonparametric distributions to obtain more accurate estimates of the Value-at-Risk, mentioned in Sect. “Introduction”. In the univariate case this is quite obvious, but in the multivariate case one has to deal with the “curse of dimensionality”, common in the nonparametrics literature. Furthermore, issues such as tail dependence need to be modeled accurately in that case. A joint framework that captures volatilities, correlations, other distributional shape features and tail dependence would be an interesting target for applied research.

Finally, *realized volatilities* (RV) have been mentioned at the end of Sect. “Properties of the GARCH(1,1) Model” as a means to use intra-day data to generate accurate ex post measures of daily volatilities. Using these RV measures, one can build time series models that predict daily volatilities one or more steps ahead, see e.g., [5] for a detailed analysis. It seems that RV provides better forecasts than GARCH, which is not surprising as it uses more information, namely the intra-day returns. The RV literature has evolved as an important second branch of volatility modeling next to the discrete time GARCH or SV models. One direction of research is the treatment of microstructure noise, present in most high frequency data, as e.g., in [1,2] and [67]. Another one is the modeling of jumps using the so-called bi-power variation and the generalization to the multivariate case using realized covariances and bipower co-variation, see e.g., [10] and [11]. Other directions are possible and it seems likely that RV will become the dominant econometric tool to model volatilities provided that high frequency data are available.

## Bibliography

### Primary Literature

1. Ait-Sahalia Y, Mykland P, Zhang L (2005) A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J Am Stat Assoc* 100:1394–1411
2. Ait-Sahalia Y, Mykland P, Zhang L (2005) How often to sample a continuous time process in the presence of market microstructure noise. *Rev Financial Stud* 18:351–416
3. Alexander C (2001) Orthogonal GARCH. In: *Mastering Risk, Financial Times*, vol 2. Prentice Hall, London, pp 21–38
4. Andersen TG, Bollerslev T (1998) Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *Int Econ Rev* 39:885–905
5. Andersen TG, Bollerslev T, Diebold FX, Labys P (2003) Modeling and forecasting realized volatility. *Econometrica* 71:579–625
6. Andrews DWK (1999) Estimation when a parameter is on a boundary. *Econometrica* 67:1341–1383
7. Baillie RT (1996) Long memory processes and fractional integration in econometrics. *J Econom* 73:5–59
8. Baillie RT, Bollerslev T, Mikkelsen HO (1996) Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J Econom* 74:3–30
9. Barndorff-Nielsen O, Shephard N (2002) Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J Roy Stat Soc* 64:253–280
10. Barndorff-Nielsen O, Shephard N (2004) Econometric analysis of realized covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica* 72:885–925
11. Barndorff-Nielsen O, Shephard N (2004) Power and bipower variation with stochastic volatility and jumps (with discussion). *J Financial Econom* 2:1–48
12. Bauwens L, Lubrano M (1998) Bayesian inference on garch models using the gibbs sampler. *Econom J* 1:C23–C46
13. Bauwens L, Laurent S, Rombouts J (2006) Multivariate garch models: a survey. *J Appl Econom* 21:79–109
14. Bera A, Higgins M (1993) A survey of ARCH models: properties, estimation and testing. *J Econ Surv* 7:305–366
15. Berndt EK, Hall BH, Hall RE, Hausman JA (1974) Estimation and inference in nonlinear structural models. *Ann Econ Soc Meas* 3:653–665
16. Bickel PJ (1982) On adaptive estimation. *Ann Stat* 10:647–671
17. Black F (1976) Studies in stock price volatility changes. In: *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, American Statistical Association*, pp 177–181
18. Black F, Scholes M (1973) The pricing of options and corporate liabilities. *J Political Econ* 81:637–654
19. Bollerslev T, Mikkelsen HO (1996) Modeling and pricing long-memory in stock market volatility. *J Econom* 73:151–184
20. Bollerslev T, Engle R, Nelson D (1994) ARCH models. In: Engle R, McFadden D (eds) *Handbook of Econometrics*. North Holland, Amsterdam, pp 2959–3038
21. Bollerslev TP (1986) Generalized autoregressive conditional heteroscedasticity. *J Econom* 31:307–327
22. Bollerslev TP (1990) Modelling the coherence in short-run nominal exchange rates: A multivariate generalized arch model. *Rev Econ Stat* 72:498–505
23. Bollerslev TP, Wooldridge JM (1992) Quasi maximum likelihood estimation of dynamic models with time-varying covariances. *Econom Rev* 11:143–172
24. Bollerslev TP, Engle RF, Wooldridge JM (1988) A capital asset pricing model with time-varying covariances. *J Political Econ* 96:116–131
25. Bougerol P, Picard N (1992) Stationarity of garch processes and some nonnegative time series. *J Econom* 52:115–127
26. Bühlmann P, McNeil AJ (2002) An algorithm for nonparametric GARCH modelling. *Comput Stat Data Anal* 40:665–683
27. Carriero MA, Pena D, Ruiz E (2004) Persistence and kurtosis in GARCH and stochastic volatility models. *J Financial Econom* 2:319–342
28. Carrasco M, Chen X (2002) Mixing and moment properties of various garch and stochastic volatility models. *Econom Theory* 18:17–39
29. Carroll R, Härdle W, Mammen E (2002) Estimation in an additive model when the components are linked parametrically. *Econom Theory* 18:886–912
30. Chan LKC, Karceski J, Lakonishok J (1999) On portfolio optimization: Forecasting covariances and choosing the risk model. *Rev Financial Stud* 12:937–674
31. Comte F, Lieberman O (2003) Asymptotic theory for multivariate garch processes. *J Multivar Anal* 84:61–84
32. Concepcion Ausian M, Galeano P (2007) Bayesian estimation of the gaussian mixture GARCH model. *Comput Stat Data Anal* 51:2636–2652
33. Danielsson J (1998) Multivariate stochastic volatility models: Estimation and a comparison with vgarch models. *J Empir Finance* 5:155–174
34. Diebold FX, Nerlove M (1989) The dynamics of exchange rate volatility: A multivariate latent factor arch model. *J Appl Econom* 4:1–21
35. Ding Z, Granger C (1996) Modelling volatility persistence of speculative returns: a new approach. *J Econom* 73:185–215
36. Ding Z, Granger CWJ, Engle RF (1993) A long memory property of stock market returns and a new model. *J Empir Finance* 1:83–106
37. Drost FC, Klaassen CAJ (1997) Efficient estimation in semi-parametric garch models. *J Econom* 81:193–221
38. Drost FC, Nijman T (1993) Temporal aggregation of GARCH processes. *Econometrica* 61:909–927
39. Drost FC, Werker BJM (1996) Closing the garch gap: Continuous garch modeling. *J Econom* 74:31–57
40. Drost FC, Klaassen CAJ, Werker BJM (1997) Adaptive estimation in time series models. *Ann Stat* 25:786–817
41. Duan JC (1995) The GARCH option pricing model. *Math Finance* 5:13–32
42. Duan JC (1997) Augmented garch( $p, q$ ) process and its diffusion limit. *J Econom* 79:97–127
43. Embrechts P, Klüppelberg C, Mikosch T (1997) *Modelling Extremal Events*. Springer, Berlin
44. Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50:987–1008
45. Engle RF (2002) Dynamic conditional correlation – a simple class of multivariate garch models. *J Bus Econ Stat* 20: 339–350
46. Engle RF (2002) New frontiers of ARCH models. *J Appl Econom* 17:425–446

47. Engle RF, Bollerslev TP (1986) Modelling the persistence of conditional variances. *Econom Rev* 5:1–50, 81–87
48. Engle RF, Gonzalez-Rivera G (1991) Semiparametric ARCH models. *J Bus Econ Stat* 9:345–360
49. Engle RF, Kroner KF (1995) Multivariate simultaneous generalized arch. *Econom Theory* 11:122–150
50. Engle RF, Mezrich J (1996) GARCH for groups. *RISK* 9:36–40
51. Engle RF, Ng VK (1993) Measuring and testing the impact of news on volatility. *J Finance* 48:1749–1778
52. Engle RF, Lilien DM, Robins RP (1987) Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica* 55:391–407
53. Engle RF, Ng VK, Rothschild M (1990) Asset pricing with a factor-ARCH covariance structure. *J Econom* 45:213–237
54. Fama EF (1965) The behavior of stock market prices. *J Bus* 38:34–105
55. Fan J, Wang M, Yao Q (2008) Modelling multivariate volatilities via conditionally uncorrelated components. *J Royal Stat Soc Ser B* 70:679–702
56. Fiorentini G, Sentana E, Shephard N (2004) Likelihood-based estimation of latent generalized ARCH structures. *Econometrica* 72:1481–1517
57. Francq C, Zakoian JM (2004) Maximum likelihood estimation of pure garch and arma-garch processes. *Bernoulli* 10: 605–637
58. Fridman M, Harris L (1998) A maximum likelihood approach for non-gaussian stochastic volatility models. *J Bus Econ Stat* 16:284–291
59. Glosten LR, Jagannathan R, Runkle DE (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. *J Finance* 48:1779–1801
60. Gouriéroux C (1992) Modèles ARCH et Applications Financières. *Economica*
61. Gouriéroux C, Monfort A (1992) Qualitative threshold ARCH models. *J Econom* 52:159–199
62. Granger CWJ (1980) Long memory relationships and the aggregation of dynamic models. *J Econom* 14:227–238
63. Hafner CM (1998) Estimating high frequency foreign exchange rate volatility with nonparametric ARCH models. *J Stat Plan Inference* 68:247–269
64. Hafner CM (2008) Temporal aggregation of multivariate GARCH processes. *J Econom* 142:467–483
65. Hafner CM, Rombouts JV (2007) Semiparametric multivariate volatility models. *Econom Theory* 23:251–280
66. Hall P, Yao P (2003) Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* 71:285–317
67. Hansen PR, Lunde A (2006) Realized variance and market microstructure noise, with comments and rejoinder. *J Bus Econ Stat* 24:127–218
68. Härdle W, Tsybakov A (1997) Local polynomial estimation of the volatility function. *J Econom* 81:223–242
69. Harvey AC, Ruiz E, Sheppard N (1994) Multivariate stochastic variance models. *Rev Econ Stud* 61:247–264
70. He C, Teräsvirta T (1999) Statistical Properties of the Asymmetric Power ARCH Process. In: Engle RF, White H (eds) *Cointegration, Causality, and Forecasting. Festschrift in honour of Clive W.J. Granger*, Oxford University Press, pp 462–474
71. Hentschel L (1995) All in the family: Nesting symmetric and asymmetric garch models. *J Financial Econ* 39:71104
72. Hill B (1975) A simple general approach to inference about the tail of a distribution. *Ann Stat* 3:1163–1174
73. Hong Y, Tu J, Zhou G (2007) Asymmetries in stock returns: Statistical tests and economic evaluation. *Rev Financial Stud* 20:1547–1581
74. Hull J, White A (1987) The pricing of options on assets with stochastic volatilities. *J Finance* 42:281–300
75. Jacquier E, Polson NG, Rossi PE (1994) Bayesian analysis of stochastic volatility models (with discussion). *J Bus Econ Stat* 12:371–417
76. Jacquier E, Polson NG, Rossi PE (2004) Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *J Econom* 122:185–212
77. Jeantheau T (1998) Strong consistency of estimators for multivariate arch models. *Econom Theory* 14:70–86
78. Jorion P (2000) Value-at-Risk: The New Benchmark for Managing Financial Risk. McGraw-Hill, New York
79. Kawakatsu H (2006) Matrix exponential GARCH. *J Econom* 134:95–128
80. Kim S, Shephard N, Chib S (1998) Stochastic volatility: likelihood inference and comparison with ARCH models. *Rev Econ Stud* 65:361–393
81. King M, Sentana E, Wadhwani S (1994) Volatility and links between national stock markets. *Econometrica* 62:901–933
82. Kristensen D, Linton O (2006) A closed-form estimator for the garch(1,1) model. *Econom Theory* 323–337
83. Lanne M, Saikkonen P (2007) A multivariate generalized orthogonal factor GARCH model. *J Bus Econ Stat* 25:61–75
84. Ledoit O, Santa-Clara P, Wolf M (2003) Flexible multivariate GARCH modeling with an application to international stock markets. *Rev Econ Stat* 85:735–747
85. Lee SW, Hansen BE (1994) Asymptotic properties of the maximum likelihood estimator and test of the stability of parameters of the GARCH and IGARCH models. *Econom Theory* 10:29–52
86. Ling S, McAleer M (2003) Asymptotic theory for a vector ARMA-GARCH model. *Econom Theory* 19:280–310
87. Lintner J (1965) Security prices, risk and maximal gains from diversification. *J Finance* 20:587–615
88. Linton O (1993) Adaptive estimation in ARCH models. *Econom Theory* 9:539–569
89. Linton O, Mammen E (2005) Estimating semiparametric ARCH models by kernel smoothing methods. *Econometrica* 73:771–836
90. Lo A, Wang J (1995) Implementing option pricing models when asset returns are predictable. *J Finance* 50:87–129
91. Lumsdaine RL (1996) Asymptotic properties of the quasi maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models. *Econometrica* 64:575–596
92. Mahieu R, Schotman P (1998) An empirical application of stochastic volatility models. *J Appl Econom* 13:333–360
93. Mandelbrot B (1963) The variation of certain speculative prices. *J Bus* 36:394–419
94. Meddahi N, Renault E (2004) Temporal aggregation of volatility models. *J of Econom* 119:355–379
95. Morgan JP (1996) Riskmetrics Technical Document, 4th edn. J.P. Morgan, New York
96. Morillo D, Pohlman L (2002) Large scale multivariate GARCH risk modelling for long-horizon international equity portfolios. Proceedings of the 2002 Forecasting Financial Markets conference, London
97. Nelson DB (1990) ARCH models as diffusion approximations. *J Econom* 45:7–38

98. Nelson DB (1990) Stationarity and persistence in the GARCH(1,1) model. *Econom Theory* 6:318–334
99. Nelson DB (1991) Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* 59:347–370
100. Nelson DB, Cao CQ (1992) Inequality constraints in the univariate garch model. *J Bus Econ Stat* 10:229–235
101. Newey WK, Steigerwald DS (1997) Asymptotic bias for quasi maximum likelihood estimators in conditional heteroskedasticity models. *Econometrica* 3:587–599
102. Nijman T, Sentana E (1996) Marginalization and contemporaneous aggregation in multivariate GARCH processes. *J Econom* 71:71–87
103. Pantula SG (1988) Estimation of autoregressive models with ARCH errors. *Sankhya Indian J Stat B* 50:119–138
104. Peng L, Yao Q (2003) Least absolute deviations estimation for ARCH and GARCH models. *Biometrika* 90:967–975
105. Robinson PM (1991) Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *J Econom* 47:67–84
106. Ross SA (1976) The arbitrage theory of capital asset pricing. *J Econ Theory* 13:341–360
107. Sentana E, Fiorentini G (2001) Identification, estimation and testing of conditionally heteroskedastic factor models. *J Econom* 102:143–164
108. Sharpe WF (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *J Finance* 19:425–442
109. Shephard N (2005) Stochastic Volatility: Selected Readings. Oxford University Press, Oxford
110. Straumann D, Mikosch T (2006) Quasi-mle in heteroscedastic times series: a stochastic recurrence equations approach. *Ann Stat* 34:2449–2495
111. Taylor SJ (1986) Modelling Financial Time Series. Wiley, New York
112. Tse YK (2000) A test for constant correlations in a multivariate GARCH model. *J Econom* 98:107–127
113. Tse YK, Tsui AKC (2002) A multivariate GARCH model with time-varying correlations. *J Bus Econ Stat* 20:351–362
114. van der Weide R (2002) Go-garch: A multivariate generalized orthogonal GARCH model. *J Appl Econom* 17:549–564
115. Vrontos ID, Dellaportas P, Politis DN (2000) Full bayesian inference for GARCH and EGARCH models. *J Bus Econ Stat* 18:187198
116. Vrontos ID, Dellaportas P, Politis D (2003) A full-factor multivariate garch model. *Econom J* 6:311–333
117. Wang Y (2002) Asymptotic nonequivalence of GARCH models and diffusions. *Ann Stat* 30:754–783
118. Weiss AA (1986) Asymptotic theory for ARCH models: Estimation and testing. *Econom Theory* 2:107–131
119. Yang L (2006) A semiparametric GARCH model for foreign exchange volatility. *J Econom* 130:365–384
120. Yang L, Härdle W, Nielsen P (1999) Nonparametric autoregression with multiplicative volatility and additive mean. *J Time Ser Anal* 20:579–604
121. Zaffaroni P (2007) Aggregation and memory of models of changing volatility. *J Econom* 136:237–249
122. Zaffaroni P (2007) Contemporaneous aggregation of GARCH processes. *J Time Series Anal* 28:521–544
123. Zakoian JM (1994) Threshold heteroskedastic functions. *J Econ Dyn Control* 18:931–955

## Books and Reviews

- Andersen T, Bollerslev T, Diebold F (2004) Parametric and nonparametric measurement of volatility. In: Ait-Sahalia L, Hansen LP (eds) *Handbook of Financial Econometrics*. Amsterdam (forthcoming)
- Bauwens L, Laurent S, Rombouts J (2006) Multivariate GARCH models: A survey. *J Appl Econom* 21:79–109
- Bera A, Higgins M (1993) A survey of ARCH models: properties, estimation and testing. *J Econ Surv* 7:305–366
- Bollerslev T, Chou R, Kroner K (1992) ARCH modelling in finance: a review of the theory and empirical evidence. *J Econom* 52:5–59
- Bollerslev T, Engle R, Nelson D (1994) ARCH models. In: Engle R, McFadden D (eds) *Handbook of Econometrics*. North Holland Press, Amsterdam, pp 2959–3038
- Engle R (1995) ARCH: Selected Readings. Oxford University Press, Oxford
- Gouriéroux C (1997) ARCH Models and Financial Applications. Springer, New York
- Shephard N (1996) Statistical aspects of ARCH and stochastic volatility. In: Cox DR, Hinkley DV, Barndorff-Nielsen OE (eds) *Time Series Models in Econometrics, Finance and Other Fields*. Chapman & Hall, London, pp 1–67
- Shephard N (2005) Stochastic Volatility: Selected Readings. Oxford University Press, Oxford
- Taylor S (1986) Modelling Financial Time Series. Wiley, Chichester

---

## Genetic and Evolutionary Algorithms and Programming: General Introduction and Application to Game Playing

MICHAEL ORLOV, MOSHE SIPPER, AMI HAUPTMAN  
Department of Computer Science,  
Ben-Gurion University, Beer-Sheva, Israel

### Article Outline

- [Glossary](#)
- [Definition of the Subject](#)
- [Introduction](#)
- [Evolutionary Algorithms](#)
- [A Touch of Theory](#)
- [Extensions of the Basic Methodology](#)
- [Lethal Applications](#)
- [Evolutionary Games](#)
- [Future Directions](#)
- [Bibliography](#)

### Glossary

#### Evolutionary algorithms/evolutionary computation

A family of algorithms inspired by the workings of

evolution by natural selection whose basic structure is to:

1. produce an initial **population** of individuals, these latter being candidate solutions to the problem at hand
2. evaluate the **fitness** of each individual in accordance with the problem whose solution is sought
3. *while* termination condition not met *do*
  - (a) **select** fitter individuals for reproduction
  - (b) **recombine (crossover)** individuals
  - (c) **mutate** individuals
  - (d) **evaluate** fitness of modified individuals*end while*

**Genome/chromosome** An individual's makeup in the population of an evolutionary algorithm is known as a genome, or chromosome. It can take on many forms, including bit strings, real-valued vectors, character-based encodings, and computer programs. The representation issue – namely, defining an individual's genome (well) – is critical to the success of an evolutionary algorithm.

**Fitness** A measure of the quality of a candidate solution in the population. Also known as *fitness function*. Defining this function well is critical to the success of an evolutionary algorithm.

**Selection** The operator by which an evolutionary algorithm selects (usually probabilistically) higher-fitness individuals to contribute *genetic* material to the next generation.

**Crossover** One of the two main genetic operators applied by an evolutionary algorithm, wherein two (or more) candidate solutions (*parents*) are combined in some pre-defined manner to form *offspring*.

**Mutation** One of the two main genetic operators applied by an evolutionary algorithm, wherein one candidate solution is randomly altered.

### Definition of the Subject

Evolutionary algorithms are a family of search algorithms inspired by the process of (Darwinian) evolution in nature. Common to all the different family members is the notion of solving problems by evolving an initially random population of candidate solutions, through the application of operators inspired by natural genetics and natural selection, such that in time *fitter* (i. e., better) solutions emerge. The field, whose origins can be traced back to the 1950s and 1960s, has come into its own over the past two decades, proving successful in solving multitudinous

problems from highly diverse domains including (to mention but a few): optimization, automatic programming, electronic-circuit design, telecommunications, networks, finance, economics, image analysis, signal processing, music, and art.

### Introduction

The first approach to artificial intelligence, the field which encompasses evolutionary computation, is arguably due to Turing [31]. Turing asked the famous question: "Can machines think?" Evolutionary computation, as a subfield of AI, may be the most straightforward answer to such a question. In principle, it might be possible to evolve an algorithm possessing the functionality of the human brain (this has already happened at least once: in nature).

In a sense, nature is greatly inventive. One often wonders how so many magnificent solutions to the problem of existence came to be. From the intricate mechanisms of cellular biology, to the sandy camouflage of flatfish; from the social behavior of ants to the diving speed of the peregrine falcon – nature created versatile solutions, at varying levels, to the problem of survival. Many ingenious solutions were *invented* (and still are), without any obvious intelligence directly creating them. This is perhaps the main motivation behind evolutionary algorithms: creating the settings for a dynamic environment, in which solutions can be created and improved in the course of time, advancing in new directions, with minimal direct intervention. The gain to problem solving is obvious.

### Evolutionary Algorithms

In the 1950s and the 1960s several researchers independently studied evolutionary systems with the idea that evolution could be used as an optimization tool for engineering problems. Central to all the different methodologies is the notion of solving problems by evolving an initially random population of candidate solutions, through the application of operators inspired by natural genetics and natural selection, such that in time *fitter* (i. e., better) solutions emerge [9,16,19,28]. This thriving field goes by the name of *evolutionary algorithms* or *evolutionary computation*, and today it encompasses two main branches – genetic algorithms [9] and genetic programming [19] – in addition to less prominent (though important) offshoots, such as evolutionary programming [10] and evolution strategies [26].

A *genetic algorithm* (GA) is an iterative procedure that consists of a *population* of individuals, each one represented by a finite string of symbols, known as the *genome*, encoding a possible solution in a given problem space.

This space, referred to as the *search space*, comprises all possible solutions to the problem at hand. Generally speaking, the genetic algorithm is applied to spaces which are too large to be exhaustively searched. The symbol alphabet used is often binary, but may also be character-based, real-valued, or any other representation most suitable to the problem at hand.

The standard genetic algorithm proceeds as follows: an initial population of individuals is generated at random or heuristically. Every evolutionary step, known as a *generation*, the individuals in the current population are *decoded* and *evaluated* according to some predefined quality criterion, referred to as the *fitness*, or *fitness function*. To form a new population (the next generation), individuals are *selected* according to their fitness. Many selection procedures are available, one of the simplest being *fitness-proportionate selection*, where individuals are selected with a probability proportional to their relative fitness. This ensures that the expected number of times an individual is chosen is approximately proportional to its relative performance in the population. Thus, high-fitness (*good*) individuals stand a better chance of *reproducing*, while low-fitness ones are more likely to disappear.

Selection alone cannot introduce any new individuals into the population, i. e., it cannot find new points in the search space; these are generated by genetically-inspired operators, of which the most well known are *crossover* and *mutation*. Crossover is performed with probability  $p_{\text{cross}}$  (the *crossover probability* or *crossover rate*) between two selected individuals, called *parents*, by exchanging parts of their genomes (i. e., encodings) to form one or two new individuals, called *offspring*. In its simplest form, substrings are exchanged after a randomly-selected crossover point. This operator tends to enable the evolutionary process to move toward *promising* regions of the search space. The mutation operator is introduced to prevent premature convergence to local optima by randomly sampling new points in the search space. It is carried out by flipping bits at random, with some (small) probability  $p_{\text{mut}}$ . Genetic algorithms are stochastic iterative processes that are not guaranteed to converge. The termination condition may be specified as some fixed, maximal number of generations or as the attainment of an acceptable fitness level. Figure 1 presents the standard genetic algorithm in pseudo-code format.

Let us consider the following simple example, demonstrating the GA's workings. The population consists of four individuals, which are binary-encoded strings (genomes) of length 10. The fitness value equals the number of ones in the bit string, with  $p_{\text{cross}} = 0.7$  and  $p_{\text{mut}} = 0.05$ . More typical values of the population size and the

```
begin GA
    g:=0 { generation counter }
    Initialize population P(g)
    Evaluate population P(g) { i.e., compute fitness values }
    while not done do
        g:=g+1
        Select P(g) from P(g - 1)
        Crossover P(g)
        Mutate P(g)
        Evaluate P(g)
    end while
end GA
```

**Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Figure 1**  
**Pseudo-code of the standard genetic algorithm**

**Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Table 1**  
**The initial population**

Label	Genome	Fitness
$p_1$	0000011011	4
$p_2$	1110111101	8
$p_3$	0010000010	2
$p_4$	0011010000	3

genome length are in the range 50–1000. Note that fitness computation in this case is extremely simple, since no complex decoding or evaluation is necessary. The initial (randomly generated) population might look as shown in Table 1.

Using fitness-proportionate selection we must choose four individuals (two sets of parents), with probabilities proportional to their relative fitness values. In our example, suppose that the two parent pairs are  $\{p_2, p_4\}$  and  $\{p_1, p_2\}$  (note that individual  $p_3$  did not get selected as our procedure is probabilistic). Once a pair of parents is selected, crossover is effected between them with probability  $p_{\text{cross}}$ , resulting in two offspring. If no crossover is effected (with probability  $1 - p_{\text{cross}}$ ), then the offspring are exact copies of each parent. Suppose, in our example, that crossover takes place between parents  $p_2$  and  $p_4$  at the (randomly chosen) third bit position:

111|0111101

001|1010000

This results in offspring  $p'_1 = 1111010000$  and  $p'_2 = 0010111101$ . Suppose no crossover is performed between parents  $p_1$  and  $p_2$ , forming offspring that are exact copies of  $p_1$  and  $p_2$ . Our interim population (after crossover) is thus as depicted in Table 2:

### Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Table 2

The interim population

Label	Genome	Fitness
$p'_1$	1111010000	5
$p'_2$	0010111101	6
$p'_3$	0000011011	4
$p'_4$	1110111101	8

### Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Table 3

The resulting population

Label	Genome	Fitness
$p''_1$	1111010000	5
$p''_2$	0010101101	5
$p''_3$	0000011011	4
$p''_4$	1110111111	9

Next, each of these four individuals is subject to mutation with probability  $p_{\text{mut}}$  per bit. For example, suppose offspring  $p'_2$  is mutated at the sixth position and offspring  $p'_4$  is mutated at the ninth bit position. Table 3 describes the resulting population.

The resulting population is that of the next generation (i. e.,  $p''_i$  equals  $p_i$  of the next generation). As can be seen, the transition from one generation to the next is through application of selection, crossover, and mutation. Moreover, note that the best individual's fitness has gone up from eight to nine, and that the average fitness (computed over all individuals in the population) has gone up from 4.25 to 5.75. Iterating this procedure, the GA will eventually find a perfect string, i. e., with maximal fitness value of ten.

Another prominent branch of the evolutionary computation tree is that of *genetic programming*, introduced by Cramer [7], and transformed into a field in its own right in large part due to the efforts of Koza [19]. Basically, genetic programming (GP) is a GA (genetic algorithm) with individuals in the population being programs instead of bit strings.

In GP we evolve a population of individual LISP expressions<sup>1</sup>, each comprising *functions* and *terminals*. The functions are usually arithmetic and logic operators that receive a number of arguments as input and compute a result as output; the terminals are zero-argument functions that serve both as constants and as sensors, the latter be-

ing a special type of function that queries the domain environment.

The main mechanism behind GP is precisely that of a GA, namely, the repeated cycling through four operations applied to the entire population: evaluate-select-crossover-mutate. However, the evaluation of a single individual in GP is usually more complex than with a GA since it involves running a program. Moreover, crossover and mutation need to be made to work on trees (rather than simple bit strings), as shown in Fig. 2.

### A Touch of Theory

Evolutionary computation is mostly an experimental field. However, over the years there have been some notable theoretical treatments of the field, gaining valuable insights into the properties of evolving populations.

Holland [17] introduced the notion of *schemata*, which are abstract properties of binary-encoded individuals, and analyzed the growth of different schemas when fitness-proportionate selection, point mutation and one-point crossover are employed. Holland's approach has since been enhanced and more rigorous analysis performed; however, there were not many practical consequences on the existing evolutionary techniques, since most of the successful methods are usually much more complex in many aspects. Moreover, the schematic analysis suffers from an important approximation of infinite population size, while in reality schemata can vanish.

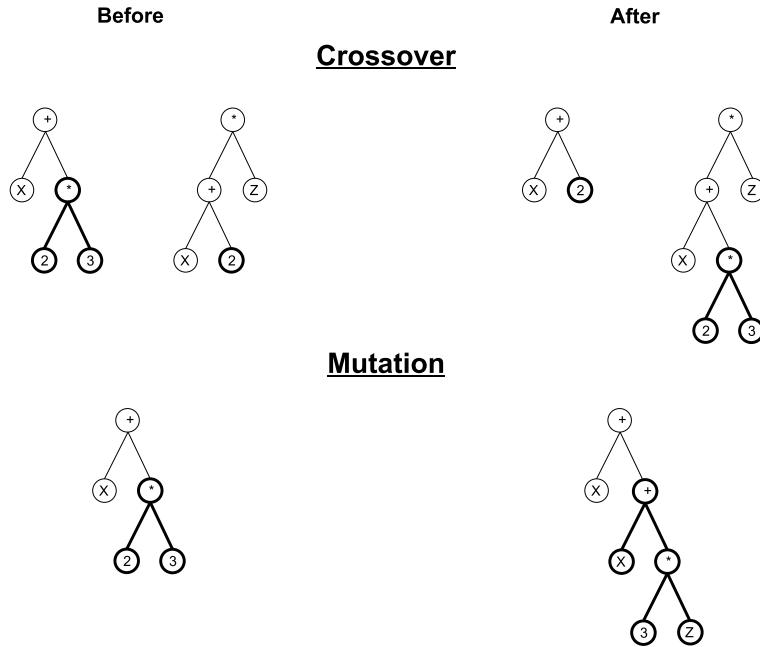
Note that the *No Free Lunch* theorem states that "...for any [optimization] algorithm, any elevated performance over one class of problems is exactly paid for in performance over another class" [32].

### Extensions of the Basic Methodology

We have reviewed the basic evolutionary computation methods. More advanced techniques are used to tackle complex problems, where an approach of a single population with homogeneous individuals does not suffice. One such advanced approach is coevolution [24].

Coevolution refers to the simultaneous evolution of two or more species with coupled fitness. Such coupled evolution favors the discovery of complex solutions whenever complex solutions are required. Simplistically speaking, one can say that coevolving species can either compete (e. g., to obtain exclusivity on a limited resource) or cooperate (e. g., to gain access to some hard-to-attain resource). In a competitive coevolutionary algorithm the fitness of an individual is based on direct competition with individuals of other species, which in turn evolve separately in their own populations. Increased fitness of one of the species

<sup>1</sup>Languages other than LISP have been used, although LISP is still by far the most popular within the genetic programming domain.



**Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Figure 2**

Genetic operators in genetic programming. LISP programs are depicted as trees. Crossover (top): Two sub-trees (marked in **bold**) are selected from the parents and swapped. Mutation (bottom): A sub-tree (marked in **bold**) is selected from the parent individual and removed. A new sub-tree is grown instead

implies a diminution in the fitness of the other species. This evolutionary pressure tends to produce new strategies in the populations involved so as to maintain their chances of survival. This *arms race* ideally increases the capabilities of each species until they reach an optimum.

Cooperative (also called symbiotic) coevolutionary algorithms involve a number of independently evolving species which together form complex structures, well suited to solve a problem. The fitness of an individual depends on its ability to collaborate with individuals from other species. In this way, the evolutionary pressure stemming from the difficulty of the problem favors the development of cooperative strategies and individuals.

Single-population evolutionary algorithms often perform poorly – manifesting stagnation, convergence to local optima, and computational costliness – when confronted with problems presenting one or more of the following features: 1) the sought-after solution is complex, 2) the problem or its solution is clearly decomposable, 3) the genome encodes different types of values, 4) strong interdependencies among the components of the solution, and 5) components-ordering drastically affects fitness [24]. Cooperative coevolution addresses effectively these issues, consequently widening the range of applications of evolutionary computation.

Consider, for instance, the evolution of neural networks [33]. A neural network consists of simple units called neurons, each having several inputs and a single output. The inputs are assigned weights, and a weighted sum of the inputs exceeding a certain threshold causes the neuron to fire an output signal. Neurons are usually connected using a layered topology.

When we approach the task of evolving a neural network possessing some desired property naively, we will probably think of some linearized representation of a neural network, encoding both the neuron locations in the network, and their weights. However, evolving such a network with a simple evolutionary algorithm might prove quite a frustrating task, since much information is encoded in each individual, and it is not homogeneous, which presents us with the difficult target of evolving the individuals as single entities.

On the other hand, this task can be dealt with more sagely via evolving two independently encoded populations of neurons and network topologies. Stanley and Miikkulainen [30] evaluate the fitness of an individual in one of the populations using the individuals of the other. In addition to the simplification of individuals in each population, the fitness is now dynamic, and an improvement in the evolution of topologies triggers a correspond-

ing improvement in the population of neurons, and vice versa.

### Lethal Applications

In this section we review a number of applications that – though possibly not killer (death being in the eye of the beholder...) – are most certainly lethal. These come from a sub-domain of evolutionary algorithms, which has been gaining momentum over the past few years: human-competitive machine intelligence. Koza et al. [20] recently affirmed that the field of evolutionary algorithms “now routinely delivers high-return human-competitive machine intelligence”, meaning, according to [20]:

- Human-competitive: Getting machines to produce human-like results, e.g., a patentable invention, a result publishable in the scientific literature, or a game strategy that can hold its own against humans.
- High-return: Defined by Koza et al. as a high *artificial-to-intelligence ratio* (A/I), namely, the ratio of that which is delivered by the automated operation of the artificial method to the amount of intelligence that is supplied by the human applying the method to a particular system.
- Routine: The successful handling of new problems once the method has been *jump-started*.
- Machine intelligence: To quote Arthur Samuel, getting “machines to exhibit behavior, which if done by humans, would be assumed to involve the use of intelligence.”

Indeed, as of 2004 the major annual event in the field of evolutionary algorithms – GECCO (Genetic and Evolutionary Computation Conference; see [www.sigevo.org](http://www.sigevo.org)) – boasts a prestigious competition that awards prizes to human-competitive results. As noted at [www.human-competitive.org](http://www.human-competitive.org): “Techniques of genetic and evolutionary computation are being increasingly applied to difficult real-world problems – often yielding results that are not merely interesting, but competitive with the work of creative and inventive humans.”

We now describe some winners from the HUMIES competition at [www.human-competitive.org](http://www.human-competitive.org). Lohn et al. [22] won a Gold Medal in the 2004 competition for an evolved X-band antenna design and flight prototype to be deployed on NASA’s Space Technology 5 (ST5) spacecraft:

The ST5 antenna was evolved to meet a challenging set of mission requirements, most notably the combination of wide beamwidth for a circularly-polarized wave and wide bandwidth. Two evolutionary

algorithms were used: one used a genetic algorithm style representation that did not allow branching in the antenna arms; the second used a genetic programming style tree-structured representation that allowed branching in the antenna arms. The highest performance antennas from both algorithms were fabricated and tested, and both yielded very similar performance. Both antennas were comparable in performance to a hand-designed antenna produced by the antenna contractor for the mission, and so we consider them examples of human-competitive performance by evolutionary algorithms [22].

Preble et al. [25] won a gold medal in the 2005 competition for designing photonic crystal structures with large band gaps. Their result is “an improvement of 12.5% over the best human design using the same index contrast platform.”

Recently, Kilinç et al. [18] was awarded the Gold Medal in the 2006 competition for designing oscillators using evolutionary algorithms, where the oscillators possess characteristics surpassing the existing human-designed analogs.

### Evolutionary Games

Evolutionary games is the application of evolutionary algorithms to the evolution of game-playing strategies for various games, including chess, backgammon, and Robocode.

### Motivation and Background

Ever since the dawn of artificial intelligence in the 1950s, games have been part and parcel of this lively field. In 1957, a year after the Dartmouth Conference that marked the official birth of AI, Alex Bernstein designed a program for the IBM 704 that played two amateur games of chess. In 1958, Allen Newell, J. C. Shaw, and Herbert Simon introduced a more sophisticated chess program (beaten in thirty-five moves by a ten-year-old beginner in its last official game played in 1960). Arthur L. Samuel of IBM spent much of the fifties working on game-playing AI programs, and by 1961 he had a checkers program that could play rather decently. In 1961 and 1963 Donald Michie described a simple trial-and-error learning system for learning how to play Tic-Tac-Toe (or Noughts and Crosses) called MENACE (for Matchbox Educable Noughts and Crosses Engine). These are but examples of highly popular games that have been treated by AI researchers since the field’s inception.

Why study games? This question was answered by Susan L. Epstein, who wrote:

There are two principal reasons to continue to do research on games.... First, human fascination with game playing is long-standing and pervasive. Anthropologists have catalogued popular games in almost every culture.... Games intrigue us because they address important cognitive functions.... The second reason to continue game-playing research is that some difficult games remain to be won, games that people play very well but computers do not. These games clarify what our current approach lacks. They set challenges for us to meet, and they promise ample rewards [8].

Studying games may thus advance our knowledge in both cognition and artificial intelligence, and, last but not least, games possess a competitive angle which coincides with our human nature, thus motivating both researcher and student alike.

Even more strongly, Laird and van Lent [21] proclaimed that,

...interactive computer games are the killer application for human-level AI. They are the application that will soon need human-level AI, and they can provide the environments for research on the right kinds of problems that lead to the type of the incremental and integrative research needed to achieve human-level AI [21].

### Evolving Game-Playing Strategies

Recently, evolutionary algorithms have proven a powerful tool that can automatically *design* successful game-playing strategies for complex games [2,3,13,14,15,27,29].

1. **Chess** (endgames) Evolve a player able to play endgames [13,14,15,29]. While endgames typically contain but a few pieces, the problem of evaluation is still hard, as the pieces are usually free to move all over the board, resulting in complex game trees – both deep and with high branching factors. Indeed, in the chess lore much has been said and written about endgames.
2. **Backgammon** Evolve a full-fledged player for the non-doubling-cube version of the game [2,3,29].
3. **Robocode** A simulation-based game in which robotic tanks fight to destruction in a closed arena ([robocode.alphaworks.ibm.com](#)). The programmers implement their robots in the Java programming language, and can test their creations either by using a graphical environment in which battles are held, or by submitting

them to a central web site where online tournaments regularly take place. Our goal here has been to evolve Robocode players able to rank high in the international league [27,29].

A strategy for a given player in a game is a way of specifying which choice the player is to make at every point in the game from the set of allowable choices at that point, given all the information that is available to the player at that point [19]. The problem of discovering a strategy for playing a game can be viewed as one of seeking a computer program. Depending on the game, the program might take as input the entire history of past moves or just the current state of the game. The desired program then produces the next move as output. For some games one might evolve a complete strategy that addresses every situation tackled. This proved to work well with Robocode, which is a dynamic game, with relatively few parameters, and little need for past history.

Another approach is to couple a current-state evaluator (e.g., board evaluator) with a next-move generator. One can go on to create a minimax tree, which consists of all possible moves, counter moves, counter counter-moves, and so on; for real-life games, such a tree's size quickly becomes prohibitive. The approach we used with backgammon and chess is to derive a very shallow, single-level tree, and evolve *smart* evaluation functions. Our artificial player is thus had by combining an evolved board evaluator with a simple program that generates all next-move boards (such programs can easily be written for backgammon and chess).

In what follows we describe the definition of six items necessary in order to employ genetic programming: program architecture, set of terminals, set of functions, fitness measure, control parameters, and manner of designating result and terminating run.

### Example: Chess

As our purpose is to create a schema-based program that analyzes single nodes thoroughly, in a way reminiscent of human thinking, we did not perform deep lookahead.

We evolved individuals represented as LISP programs. Each such program receives a chess endgame position as input, and, according to its sensors (terminals) and functions, returns an evaluation of the board, in the form of a real number.

Our chess endgame players consist of an evolved LISP program, together with a piece of software that generates all possible (legal) next-moves and feeds them to the program. The next-move with the highest score is selected

(ties are broken stochastically). The player also identifies when the game is over (either by a draw or a win).

**Program Architecture** As most chess players would agree, playing a winning position (e.g., with material advantage) is very different than playing a losing position, or an even one. For this reason, each individual contains not one but three separate trees: an advantage tree, an even tree, and a disadvantage tree. These trees are used according to the current status of the board. The disadvantage tree is smaller, since achieving a stalemate and avoiding exchanges requires less complicated reasoning. Most terminals and functions were used for all trees.

The structure of three trees per individual was preserved mainly for simplicity reasons. It is actually possible to coevolve three separate populations of trees, without binding them to form a single individual before the end of the experiment. This would require a different experimental setting, and is one of our future-work ideas.

**Terminals and Functions** While evaluating a position, an expert chess player considers various aspects of the board. Some are simple, while others require a deep understanding of the game. Chase and Simon found that experts recalled meaningful chess formations better than novices [6]. This led them to hypothesize that chess skill depends on a large knowledge base, indexed through thousands of familiar chess patterns.

We assumed that complex aspects of the game board are comprised of simpler units, which require less game knowledge, and are to be combined in some way. Our chess programs use terminals, which represent those relatively simple aspects, and functions, which incorporate no game knowledge, but supply methods of combining those aspects. As we used strongly typed GP [23], all functions and terminals were assigned one or more of two data types: *Float* and *Boolean*. We also included a third data type, named *Query*, which could be used as any of the former two. We also used ephemeral random constants (ERCs).

**The Terminal Set** We developed most of our terminals by consulting several high-ranking chess players.<sup>2</sup> The terminal set examined various aspects of the chessboard, and may be divided into three groups:

*Float values*, created using the ERC mechanism. ERCs were chosen at random to be one of the following six values:  $\pm 1 \cdot \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}\} \cdot MAX$  ( $MAX$  was empirically set to 1000), and the inverses of these numbers. This guaranteed

that when a value was returned after some group of features has been identified, it was distinct enough to engender the outcome.

*Simple terminals*, which analyzed relatively simple aspects of the board, such as the number of possible moves for each king, and the number of attacked pieces for each player. These terminals were derived by breaking relatively complex aspects of the board into simpler notions. More complex terminals belonged to the next group (see below). For example, a player should capture his opponent's piece if it is not sufficiently protected, meaning that the number of attacking pieces the player controls is greater than the number of pieces protecting the opponent's piece, and the material value of the defending pieces is equal to or greater than the player's. Adjudicating these considerations is not simple, and therefore a terminal that performs this entire computational feat by itself belongs to the next group of complex terminals.

The simple terminals comprising this second group were derived by refining the logical resolution of the previous paragraphs' reasoning: Is an opponent's piece attacked? How many of the player's pieces are attacking that piece? How many pieces are protecting a given opponent's piece? What is the material value of pieces attacking and defending a given opponent's piece? All these questions were embodied as terminals within the second group. The ability to easily embody such reasoning within the GP setup, as functions and terminals, is a major asset of GP.

Other terminals were also derived in a similar manner. See Table 4 for a complete list of simple terminals. Note that some of the terminals are inverted – we would like terminals to always return positive (or true) values, since these values represent a favorable position. This is why we used, for example, a terminal evaluating the player's king's *distance* from the edges of the board (generally a favorable feature for endgames), while using a terminal evaluating the *proximity* of the opponent's king to the edges (again, a positive feature).

*Complex terminals*: these are terminals that check the same aspects of the board a human player would. Some prominent examples include: the terminal *OppPieceCanBeCaptured* considering the capture of a piece; checking if the current position is a draw, a mate, or a stalemate (especially important for non-even boards); checking if there is a mate in one or two moves (this is the most complex terminal); the material value of the position; comparing the material value of the position to the original board – this is important since it is easier to consider change than to evaluate the board in an absolute manner. See Table 5 for a full list of complex terminals.

<sup>2</sup>The highest-ranking player we consulted was Boris Gutkin, ELO 2400, International Master, and fully qualified chess teacher.

**Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Table 4**  
 Simple terminals for evolving chess endgame players. *Opp*: opponent, *My*: player

Terminal	Description
B=NotMyKingInCheck()	Is the player's king not being checked?
B=IsOppKingInCheck()	Is the opponent's king being checked?
F=MyKingDistEdges()	The player's king's distance from the edges of the board
F=OppKingProximityToEdges()	The opponent's king's proximity to the edges of the board
F=NumMyPiecesNotAttacked()	The number of the player's pieces that are not attacked
F=NumOppPiecesAttacked()	The number of the opponent's attacked pieces
F=ValueMyPiecesAttacking()	The material value of the player's pieces which are attacking
F=ValueOppPiecesAttacking()	The material value of the opponent's pieces which are attacking
B=IsMyQueenNotAttacked()	Is the player's queen not attacked?
B=IsOppQueenAttacked()	Is the opponent's queen attacked?
B=IsMyFork()	Is the player creating a fork?
B=IsOppNotFork()	Is the opponent not creating a fork?
F=NumMovesMyKing()	The number of legal moves for the player's king
F=NumNotMovesOppKing()	The number of illegal moves for the opponent's king
F=MyKingProxRook()	Proximity of my king and rook(s)
F=OppKingDistRook()	Distance between opponent's king and rook(s)
B=MyPiecesSameLine()	Are two or more of the player's pieces protecting each other?
B=OppPiecesNotSameLine()	Are two or more of the opponent's pieces protecting each other?
B=IsOppKingProtectingPiece()	Is the opponent's king protecting one of his pieces?
B=IsMyKingProtectingPiece()	Is the player's king protecting one of his pieces?

**Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Table 5**  
 Complex terminals for evolving chess endgame players. *Opp*: opponent, *My*: player. Some of these terminals perform lookahead, while others compare with the original board

Terminal	Description
F=EvaluateMaterial()	The material value of the board
B=IsMaterialIncrease()	Did the player capture a piece?
B=IsMate()	Is this a mate position?
B=IsMateInOne()	Can the opponent mate the player after this move?
B=OppPieceCanBeCaptured()	Is it possible to capture one of the opponent's pieces without retaliation?
B=MyPieceCannotBeCaptured()	Is it not possible to capture one of the player's pieces without retaliation?
B=IsOppKingStuck()	Do all legal moves for the opponent's king advance it closer to the edges?
B=IsMyKingNotStuck()	Is there a legal move for the player's king that advances it away from the edges?
B=IsOppKingBehindPiece()	Is the opponent's king two or more squares behind one of his pieces?
B=IsMyKingNotBehindPiece()	Is the player's king not two or more squares behind one of my pieces?
B=IsOppPiecePinned()	Is one or more of the opponent's pieces pinned?
B=IsMyPieceNotPinned()	Are all the player's pieces not pinned?

Since some of these terminals are hard to compute, and most appear more than once in the individual's trees, we used a memoization scheme to save time [1]: After the first calculation of each terminal, the result is stored, so that further calls to the same terminal (on the same board) do not repeat the calculation. Memoization greatly reduced the evolutionary run-time.

**The Function Set** The function set used included the If function, and simple Boolean functions. Although our tree returns a real number, we omitted arithmetic functions, for several reasons. First, a large part of contemporary research in the field of machine learning and game theory (in particular for perfect-information games) revolves around inducing logical rules for learning games

### Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Table 6

Function set of GP chess player individual.  $B$ : Boolean,  $F$ : Float

Function	Description
$F=If3(B_1, F_1, F_2)$	If $B_1$ is non-zero, return $F_1$ , else return $F_2$
$B=Or2(B_1, B_2)$	Return 1 if at least one of $B_1, B_2$ is non-zero, 0 otherwise
$B=Or3(B_1, B_2, B_3)$	Return 1 if at least one of $B_1, B_2, B_3$ is non-zero, 0 otherwise
$B=And2(B_1, B_2)$	Return 1 only if $B_1$ and $B_2$ are non-zero, 0 otherwise
$B=And3(B_1, B_2, B_3)$	Return 1 only if $B_1, B_2$ , and $B_3$ are non-zero, 0 otherwise
$B=Smaller(B_1, B_2)$	Return 1 if $B_1$ is smaller than $B_2$ , 0 otherwise
$B=Not(B_1)$	Return 0 if $B_1$ is non-zero, 1 otherwise

(for example, see [4,5,11]). Second, according to the players we consulted, while evaluating positions involves considering various aspects of the board, some more important than others, performing logical operations on these aspects seems natural, while mathematical operations does not. Third, we observed that numeric functions sometimes returned extremely large values, which interfered with subtle calculations. Therefore the scheme we used was a (carefully ordered) series of Boolean queries, each returning a fixed value (either an ERC or a numeric terminal, see below). See Table 6 for the complete list of functions.

**Fitness Evaluation** As we used a competitive evaluation scheme, the fitness of an individual was determined by its success against its peers. We used the random-two-ways method, in which each individual plays against a fixed number of randomly selected peers. Each of these encounters entailed a fixed number of games, each starting from a randomly generated position in which no piece was attacked.

The score for each game was derived from the outcome of the game. Players that managed to mate their opponents received more points than those that achieved only a material advantage. Draws were rewarded by a score of low value and losses entailed no points at all.

The final fitness for each player was the sum of all points earned in the entire tournament for that generation.

**Control Parameters and Run Termination** We used the standard reproduction, crossover, and mutation operators. The major parameters were: population size – 80, generation count – between 150 and 250, reproduction probability – 0.35, crossover probability – 0.5, and mutation probability – 0.15 (including ERC).

**Results** We pitted our top evolved chess-endgame players against two very strong external opponents: 1) A pro-

### Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Table 7

Percent of wins, advantages, and draws for best GP-EndChess player in tournament against two top competitors

	%Wins	%Advs	%Draws
Master	6.00	2.00	68.00
CRAFTY	2.00	4.00	72.00

gram we wrote ('Master') based upon consultation with several high-ranking chess players (the highest being Boris Gutkin, ELO 2400, International Master); 2) CRAFTY – a world-class chess program, which finished second in the 2004 World Computer Speed Chess Championship ([www.cs.biu.ac.il/games/](http://www.cs.biu.ac.il/games/)). Speed chess (*blitz*) involves a time-limit per move, which we imposed both on CRAFTY and on our players. Not only did we thus seek to evolve good players, but ones that play well *and fast*. Results are shown in Table 7. As can be seen, GP-EndChess manages to hold its own, and even win, against these top players. For more details on GP-EndChess see [13,29].

Deeper analysis of the strategies developed [12] revealed several important shortcomings, most of which stemmed from the fact that they used deep knowledge and little search (typically, they developed only *one* level of the search tree). Simply increasing the search depth would not solve the problem, since the evolved programs examine each board very thoroughly, and scanning many boards would increase time requirements prohibitively. And so we turned to evolution to find an optimal way to overcome this problem: How to add more search at the expense of less knowledgeable (and thus less time-consuming) node evaluators, while attaining better performance. In [15] we evolved the search algorithm itself, focusing on the *Mate-In-N* problem: find a key move such that even with the best possible counter-plays, the opponent cannot avoid being mated in (or before) move  $N$ . We showed that our evolved search algorithms successfully solve sev-

eral instances of the Mate-In-N problem, for the hardest ones developing 47% less game-tree nodes than CRAFTY. Improvement is thus not over the basic alpha-beta algorithm, but over a world-class program using all standard enhancements [15].

Finally, in [14], we examined a strong evolved chess-endgame player, focusing on the player's emergent capabilities and tactics in the context of a chess match. Using a number of methods we analyzed the evolved player's building blocks and their effect on play level. We concluded that evolution has found combinations of building blocks that are far from trivial and cannot be explained through simple combination – thereby indicating the possible emergence of complex strategies.

### Example: Robocode

**Program Architecture** A Robocode player is written as an event-driven Java program. A main loop controls the tank activities, which can be interrupted on various occasions, called *events*. The program is limited to four lines of code, as we were aiming for the HaikuBot category, one of the divisions of the international league with a four-line code limit. The main loop contains one line of code that directs the robot to start turning the gun (and the mounted radar) to the right. This insures that within the first gun cycle, an enemy tank will be spotted by the radar, triggering a *ScannedRobotEvent*. Within the code for this event, three additional lines of code were added, each controlling a single actuator, and using a single numerical input that was supplied by a genetic programming-evolved sub-program. The first line instructs the tank to move to a distance specified by the first evolved argument. The second line instructs the tank to turn to an azimuth specified by the second evolved argument. The third line instructs the gun (and radar) to turn to an azimuth specified by the third evolved argument (Fig. 3).

```
Robocode Player
while (true)
    TurnGunRight(INFINITY); //main code loop
    ...
OnScannedRobot() {
    MoveTank(<GP#1>);
    TurnTankRight(<GP#2>);
    TurnGunRight(<GP#3>);
}
```

**Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Figure 3**  
Robocode player's code layout (HaikuBot division)

### Genetic and Evolutionary Algorithms and Programming: General Introduction and Appl. to Game Playing, Table 8

**Robocode representation.** **a** Terminal set. **b** Function set ( $F$ : Float)

Terminal	Description
Energy()	Returns the remaining energy of the player
Heading()	Returns the current heading of the player
X()	Returns the current horizontal position of the player
Y()	Returns the current vertical position of the player
MaxX()	Returns the horizontal battlefield dimension
MaxY()	Returns the vertical battlefield dimension
EnemyBearing()	Returns the current enemy bearing, relative to the current player's heading
EnemyDistance()	Returns the current distance to the enemy
EnemyVelocity()	Returns the current enemy's velocity
EnemyHeading()	Returns the current enemy heading, relative to the current player's heading
EnemyEnergy()	Returns the remaining energy of the enemy
Constant()	An ERC (Ephemeral Random Constant) in the range $[-1, 1]$
Random()	Returns a random real number in the range $[-1, 1]$
Zero()	Returns the constant 0

Function	Description
Add( $F, F$ )	Add two real numbers
Sub( $F, F$ )	Subtract two real numbers
Mul( $F, F$ )	Multiply two real numbers
Div( $F, F$ )	Divide first argument by second, if denominator non-zero, otherwise return zero
Abs( $F$ )	Absolute value
Neg( $F$ )	Negative value
Sin( $F$ )	Sine function
Cos( $F$ )	Cosine function
ArcSin( $F$ )	Arcsine function
ArcCos( $F$ )	Arccosine function
IfGreater( $F, F, F$ )	If first argument greater than second, return value of third argument, else return value of fourth argument
IfPositive( $F, F, F$ )	If first argument is positive, return value of second argument, else return value of third argument
Fire( $F$ )	If argument is positive, execute fire command with argument as firepower and return 1; otherwise, do nothing and return 0

**Terminal and Function Sets** We divided the terminals into three groups according to their functionality [27], as shown in Table 8:

1. Game-status indicators: A set of terminals that provide real-time information on the game status, such as

last enemy azimuth, current tank position, and energy levels.

2. Numerical constants: Two terminals, one providing the constant 0, the other being an ERC (ephemeral random constant). This latter terminal is initialized to a random real numerical value in the range  $[-1, 1]$ , and does not change during evolution.
3. Fire command: This special function is used to curtail one line of code by not implementing the fire actuator in a dedicated line.

**Fitness Measure** We explored two different modes of learning: using a fixed external opponent as teacher, and coevolution – letting the individuals play against each other; the former proved better. However, not one external opponent was used to measure performance but three, these adversaries downloaded from the HaikuBot league ([robocode.yajags.com](http://robocode.yajags.com)). The fitness value of an individual equals its average fractional score (over three battles).

**Control Parameters and Run Termination** The major evolutionary parameters [19] were: population size – 256, generation count – between 100 and 200, selection method – tournament, reproduction probability – 0, crossover probability – 0.95, and mutation probability – 0.05. An evolutionary run terminates when fitness is observed to level off. Since the game is highly nondeterministic a *lucky* individual might attain a higher fitness value than better overall individuals. In order to obtain a more accurate measure for the evolved players we let each of them do battle for 100 rounds against 12 different adversaries (one at a time). The results were used to extract the top player – to be submitted to the international league.

**Results** We submitted our top player to the HaikuBot division of the international league. At its very first tournament it came in third, later climbing to first place of 28 ([robocode.yajags.com/20050625/haiku-1v1.html](http://robocode.yajags.com/20050625/haiku-1v1.html)). All other 27 programs, defeated by our evolved strategy, were written by humans. For more details on GP-Robocode see [27,29].

### Backgammon: Major Results

We pitted our top evolved backgammon players against *Pubeval*, a free, public-domain board evaluation function written by Tesauro. The program – which plays well – has become the de facto yardstick used by the growing community of backgammon-playing program developers. Our top evolved player was able to attain a win percentage of 62.4% in a tournament against Pubeval, about 10%

higher (!) than the previous top method. Moreover, several evolved strategies were able to surpass the 60% mark, and most of them outdid all previous works. For more details on GP-Gammon see [2,3,29].

### Future Directions

Evolutionary computation is a fast growing field. As shown above, difficult, real-world problems are being tackled on a daily basis, both in academia and in industry. In the future we expect major developments in the underlying theory. Partly spurred by this we also expect major new application areas to succumb to evolutionary algorithms, and many more human-competitive results. Expecting such pivotal breakthroughs may seem perhaps a bit of overreaching, but one must always keep in mind Evolutionary Computation's success in Nature.

### Bibliography

1. Abelson H, Sussman GJ, Sussman J (1996) Structure and Interpretation of Computer Programs, 2nd edn. MIT Press, Cambridge
2. Azaria Y, Sipper M (2005) GP-Gammon: Genetically programming backgammon players. *Genet Program Evolvable Mach* 6(3):283–300. doi:[10.1007/s10710-005-2990-0](https://doi.org/10.1007/s10710-005-2990-0)
3. Azaria Y, Sipper M (2005) GP-Gammon: Using genetic programming to evolve backgammon players. In: Keijzer M, Tetamanzi A, Collet P, van Hemert J, Tomassini M (eds) Proceedings of 8th European Conference on Genetic Programming (EuroGP2005). Lecture Notes in Computer Science, vol 3447. Springer, Heidelberg, pp 132–142. doi:[10.1007/b107383](https://doi.org/10.1007/b107383)
4. Bain M (1994) Learning logical exceptions in chess. Ph D thesis, University of Strathclyde, Glasgow, Scotland. [citeseer.ist.psu.edu/bain94learning.html](http://citeseer.ist.psu.edu/bain94learning.html)
5. Bonanno G (1989) The logic of rational play in games of perfect information. Papers 347, California Davis – Institute of Governmental Affairs. <http://ideas.repec.org/p/fth/caldav/347.html>
6. Charness N (1991) Expertise in chess: The balance between knowledge and search. In: Ericsson KA, Smith J (eds) Toward a general theory of Expertise: Prospects and limits. Cambridge University Press, Cambridge
7. Cramer NL (1985) A representation for the adaptive generation of simple sequential programs. In: Grefenstette JJ (ed) Proceedings of the 1st International Conference on Genetic Algorithms. Lawrence Erlbaum Associates, Mahwah, pp 183–187
8. Epstein SL (1999) Game playing: The next moves. In: Proceedings of the Sixteenth National Conference on Artificial Intelligence. AAAI Press, Menlo Park, pp 987–993
9. Fogel DB (2006) Evolutionary Computation: Toward a New Philosophy of Machine Intelligence, 3rd edn. Wiley-IEEE Press, Hoboken
10. Fogel LJ, Owens AJ, Walsh MJ (1966) Artificial Intelligence Through Simulated Evolution. Wiley, New York
11. Fürnkranz J (1996) Machine learning in computer chess: The next generation. *Int Comput Chess Assoc J* 19(3):147–161. [citeseer.ist.psu.edu/furnkranz96machine.html](http://citeseer.ist.psu.edu/furnkranz96machine.html)

12. Hauptman A, Sipper M (2005) Analyzing the intelligence of a genetically programmed chess player. In: Late Breaking Papers at the 2005 Genetic and Evolutionary Computation Conference, distributed on CD-ROM at GECCO-2005, Washington DC
13. Hauptman A, Sipper M (2005) GP-EndChess: Using genetic programming to evolve chess endgame players. In: Keijzer M, Tettamanzi A, Collet P, van Hemert J, Tomassini M (eds) Proceedings of 8th European Conference on Genetic Programming (EuroGP2005). Lecture Notes in Computer Science, vol 3447. Springer, Heidelberg, pp 120–131. doi:[10.1007/b107383](https://doi.org/10.1007/b107383)
14. Hauptman A, Sipper M (2007) Emergence of complex strategies in the evolution of chess endgame players. *Adv Complex Syst* 10(1):35–59. doi:[10.1142/s0219525907001082](https://doi.org/10.1142/s0219525907001082)
15. Hauptman A, Sipper M (2007) Evolution of an efficient search algorithm for the mate-in- $n$  problem in chess. In: Ebner M, O'Neill M, Ekárt A, Vanneschi L, Esparcia-Alcázar AI (eds) Proceedings of 10th European Conference on Genetic Programming (EuroGP2007). Lecture Notes in Computer Science vol. 4455. Springer, Heidelberg, pp 78–89. doi:[10.1007/978-3-540-71605-1\\_8](https://doi.org/10.1007/978-3-540-71605-1_8)
16. Holland JH (1975) Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. University of Michigan Press, Ann Arbor (2nd edn. MIT Press, Cambridge, 1992)
17. Holland JH (1992) Adaptation in Natural and Artificial Systems, 2nd edn. MIT Press, Cambridge
18. Kilinç S, Jain V, Aggarwal V, Cam U (2006) Catalogue of variable frequency and single-resistance-controlled oscillators employing a single differential difference complementary current conveyor. *Frequenz: J RF-Eng Telecommun* 60(7–8):142–146
19. Koza JR (1992) Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press, Cambridge
20. Koza JR, Keane MA, Streeter MJ, Mydlowec W, Yu J, Lanza G (2003) Genetic Programming IV: Routine Human-Competitive Machine Intelligence. Kluwer, Norwell
21. Laird JE, van Lent M (2000) Human-level AI's killer application: Interactive computer games. In: AAAI-00: Proceedings of the 17th National Conference on Artificial Intelligence. MIT Press, Cambridge, pp 1171–1178
22. Lohn JD, Hornby GS, Linden DS (2005) An evolved antenna for deployment on NASA's Space Technology 5 mission. In: O'Reilly UM, Yu T, Riolo R, Worzel B (eds) Genetic Programming Theory and Practice II, Genetic Programming, vol 8, chap 18. Springer, pp 301–315. doi:[10.1007/0-387-23254-0\\_18](https://doi.org/10.1007/0-387-23254-0_18)
23. Montana DJ (1995) Strongly typed genetic programming. *Evol Comput* 3(2):199–230. doi:[10.1162/evco.1995.3.2.199](https://doi.org/10.1162/evco.1995.3.2.199)
24. Peña-Reyes CA, Sipper M (2001) Fuzzy CoCo: A cooperative-coevolutionary approach to fuzzy modeling. *IEEE Trans Fuzzy Syst* 9(5):727–737. doi:[10.1009/91.963759](https://doi.org/10.1009/91.963759)
25. Preble S, Lipson M, Lipson H (2005) Two-dimensional photonic crystals designed by evolutionary algorithms. *Appl Phys Lett* 86(6):061111. doi:[10.1063/1.1862783](https://doi.org/10.1063/1.1862783)
26. Schwefel HP (1995) Evolution and Optimum Seeking. Wiley, New York
27. Shichel Y, Ziserman E, Sipper M (2005) GP-Robocode: Using genetic programming to evolve robocode players. In: Keijzer M, Tettamanzi A, Collet P, van Hemert J, Tomassini M (eds) Genetic Programming: 8th European Conference, EuroGP 2005, Lausanne, Switzerland, March 30–April 1, 2005. Lecture Notes in Computer Science, vol 3447. Springer, Berlin, pp 143–154. doi:[10.1007/b107383](https://doi.org/10.1007/b107383)
28. Sipper M (2002) Machine Nature: The Coming Age of Bio-Inspired Computing. McGraw-Hill, New York
29. Sipper M, Azaria Y, Hauptman A, Shichel Y (2007) Designing an evolutionary strategizing machine for game playing and beyond. *IEEE Trans Syst, Man, Cybern, Part C: Appl Rev* 37(4):583–593
30. Stanley KO, Miikkulainen R (2002) Evolving neural networks through augmenting topologies. *Evol Comput* 10(2):99–127. doi:[10.1162/106365602320169811](https://doi.org/10.1162/106365602320169811)
31. Turing AM (1950) Computing machinery and intelligence. *Mind* 59(236):433–460. [http://links.jstor.org/sici? sici=0026-4423\(195010\)2:59:236<433:CMAI>2.0.CO;2-5](http://links.jstor.org/sici? sici=0026-4423(195010)2:59:236<433:CMAI>2.0.CO;2-5)
32. Wolpert DH, Macready WG (1997) No free lunch theorems for optimization. *IEEE Trans Evol Comput* 1(1):67–82. doi:[10.1109/4235.585893](https://doi.org/10.1109/4235.585893)
33. Yao X (1999) Evolving artificial neural networks. *Proc IEEE* 87(9):1423–1447. doi:[10.1009/5.784219](https://doi.org/10.1009/5.784219)

## Genetic-Fuzzy Data Mining Techniques

TZUNG-PEI HONG<sup>1</sup>, CHUN-HAO CHEN<sup>2</sup>, VINCENT S. TSENG<sup>2</sup>

<sup>1</sup> Department of Computer Science and Information Engineering, National University of Kaohsiung, Kaohsiung, Taiwan

<sup>2</sup> Department of Computer Science and Information Engineering, National Cheng-Kung University, Tainan, Taiwan

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Data Mining](#)

[Fuzzy Sets](#)

[Fuzzy Data Mining](#)

[Genetic Algorithms](#)

[Genetic-Fuzzy Data Mining Techniques](#)

[Future Directions](#)

[Bibliography](#)

### Glossary

**Data mining** Data mining is the process of extracting desirable knowledge or interesting patterns from existing databases for specific purposes. The common techniques include mining association rules, mining sequential patterns, clustering, and classification, among others.

**Fuzzy set theory** The fuzzy set theory was first proposed by Zadeh in 1965. It is primarily concerned with quantifying and reasoning using natural language in which words can have ambiguous meanings. It is widely used in a variety of fields because of its simplicity and similarity to human reasoning.

**Fuzzy data mining** The concept of fuzzy sets can be used in data mining to handle quantitative or linguistic data. Basically, fuzzy data mining first uses membership functions to transform each quantitative value into a fuzzy set in linguistic terms and then uses a fuzzy mining process to find fuzzy association rules.

**Genetic algorithms** Genetic Algorithms (GAs) were first proposed by Holland in 1975. They have become increasingly important for researchers in solving difficult problems since they could provide feasible solutions in a limited amount of time. Each possible solution is encoded as a chromosome (individual) in a population. According to the principle of survival of the fittest, GAs generate the next population by several genetic operations such as crossover, mutation, and reproductions.

**Genetic-fuzzy data mining** Genetic algorithms have been widely used for solving optimization problems. If the fuzzy mining problem can be converted into an optimization problem, then the GA techniques can easily be adopted to solve it. They are thus called genetic-fuzzy data-mining techniques. They are usually used to automatically mine both appropriate membership functions and fuzzy association rules from a set of transaction data.

### Definition of the Subject

Data mining is the process of extracting desirable knowledge or interesting patterns from existing databases for specific purposes. Most conventional data-mining algorithms identify the relationships among transactions using binary values. However, transactions with quantitative values are commonly seen in real-world applications.

Fuzzy data-mining algorithms are thus proposed for extracting interesting linguistic knowledge from transactions stored as quantitative values. They usually integrate fuzzy-set concepts and mining algorithms to find interesting fuzzy knowledge from a given transaction data set. Most of them mine fuzzy knowledge under the assumption that a set of membership functions [8,23,24,35,36,50] is known in advance for the problem to be solved. The given membership functions may, however, have a critical influence on the final mining results. Different membership functions may infer different knowledge. Automatically deriv-

ing an appropriate set of membership functions for a fuzzy mining problem is thus very important. There are at least two reasons for it. The first one is that a set of appropriate membership functions may not be defined by experts because lots of money and time are needed and experts are not always available. The second one is that data and concepts are always changing along with time. Some mechanisms are thus needed to automatically adapt the membership functions to the changes if needed. The fuzzy mining problem can thus be extended to finding both appropriate membership functions and fuzzy association rules from a set of transaction data.

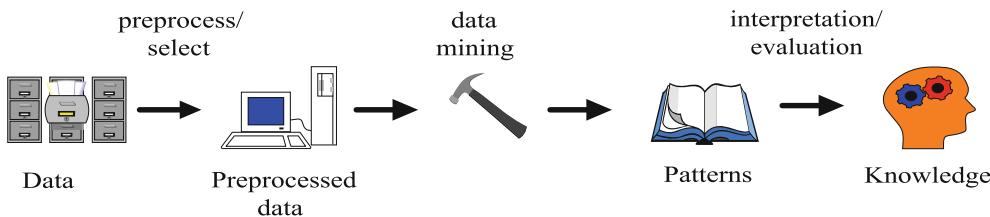
Recently, genetic algorithms have been widely used for solving optimization problems. If the fuzzy mining problem can be converted into an optimization problem, then the GA techniques can easily be adopted to solve it. They are thus called genetic-fuzzy data-mining techniques. They are usually used to automatically mine both appropriate membership functions and fuzzy association rules from a set of transaction data. Some existing approaches are introduced here. These techniques can dynamically adapt membership functions by genetic algorithms according to some criteria, use them to fuzzify the quantitative transactions, and find fuzzy association rules by fuzzy mining approaches.

### Introduction

Most enterprises have databases that contain a wealth of potentially accessible information. The unlimited growth of data, however, inevitably leads to a situation in which accessing desired information from a database becomes difficult. Knowledge discovery in databases (KDD) has thus become a process of considerable interest in recent years, as the amounts of data in many databases have grown tremendously large. KDD means the application of nontrivial procedures for identifying effective, coherent, potentially useful, and previously unknown patterns in large databases [16]. The KDD process [16] is shown in Fig. 1.

In Fig. 1, data are first collected from a single or multiple sources. These data are then preprocessed, including methods such as sampling, feature selection or reduction, data transformation, among others. After that, data-mining techniques are then used to find useful patterns, which are then interpreted and evaluated to form human knowledge.

Especially, data mining plays a critical role to the KDD process. It involves applying specific algorithms for extracting patterns or rules from data sets in a particular representation. Because of its importance, many researchers



**Genetic-Fuzzy Data Mining Techniques, Figure 1**

A KDD Process

in the database and machine learning fields are primarily interested in this topic because it offers opportunities to discover useful information and important relevant patterns in large databases, thus helping decision-makers easily analyze the data and make good decisions regarding the domains concerned. For example, there may exist some implicitly useful knowledge in a large database containing millions of records of customers' purchase orders over the recent years. This knowledge can be found by appropriate data-mining approaches. Questions such as "what are the most important trends in customers' purchase behavior?" can thus be easily answered.

Most of the mining approaches were proposed for binary transaction data. However, in real applications, quantitative data exists and should also be considered. Fuzzy set theory is being used more and more frequently in intelligent systems because of its simplicity and similarity to human reasoning [51]. The theory has been applied in fields such as manufacturing, engineering, diagnosis, economics, among others [45,52]. Several fuzzy learning algorithms for inducing rules from given sets of data have been designed and used to good effect within specific domains [7,22]. As to fuzzy data mining, many algorithms are also proposed [8,23,24,35,36,50].

Most of these fuzzy data-mining algorithms assume the membership functions are already known. In fuzzy mining problems, the given membership functions may, however, have a critical influence on the final mining results. Developing effective and efficient approaches to derive both the appropriate membership functions and fuzzy association rules automatically are thus worth being studied. Genetic algorithms are widely used for finding membership functions in different fuzzy applications. In this article, we discuss the genetic-fuzzy data-mining approaches which can mine both appropriate membership functions and fuzzy association rules [9,10,11,12,25,26,27,31,32,33]. The genetic-fuzzy mining problems can be divided into four kinds according to the types of fuzzy mining problems and the ways of processing items. The types of fuzzy mining problems include Single-minimum-Support Fuzzy

Mining (SSFM) and Multiple-minimum-Support Fuzzy Mining (MSFM). The ways of processing items include processing all the items together (integrated approach) and processing them individually (divide-and-conquer approach). Each of them will be described in details in the following sections. But first of all, the basic concepts of data mining will be described below.

### Data Mining

Data mining techniques have been used in different fields to discover interesting information from databases in recent years. Depending on the type of databases processed, mining approaches may be classified as working on transaction databases, temporal databases, relational databases, multimedia databases, and data streams, among others. On the other hand, depending on the classes of knowledge derived, mining approaches may be classified as finding association rules, classification rules, clustering rules, and sequential patterns, among others.

Finding association rules in transaction databases is most commonly seen in data mining. It is initially applied to market basket analysis for getting relationships of purchased items. An association rule can be expressed as the form  $A \rightarrow B$ , where  $A$  and  $B$  are sets of items, such that the presence of  $A$  in a transaction will imply the presence of  $B$ . Two measures, support and confidence, are evaluated to determine whether a rule should be kept. The support of a rule is the fraction of the transactions that contain all the items in  $A$  and  $B$ . The confidence of a rule is the conditional probability of the occurrences of items in  $A$  and  $B$  over the occurrences of items in  $A$ . The support and the confidence of an interesting rule must be larger than or equal to a user-specified minimum support and a minimum confidence, respectively.

To achieve this purpose, Agrawal and his coworkers proposed several mining algorithms based on the concept of large item sets to find association rules in transaction data [1,2,3,4]. They divided the mining process into two phases. In the first phase, candidate item sets were gener-

ated and counted by scanning the transaction data. If the number of an itemset appearing in the transactions was larger than a predefined threshold value (called minimum support), the itemset was considered a large itemset. Itemsets containing only one item were processed first. Large itemsets containing only single items were then combined to form candidate itemsets containing two items. This process was repeated until all large itemsets had been found. In the second phase, association rules were induced from the large itemsets found in the first phase. All possible association combinations for each large itemset were formed, and those with calculated confidence values larger than a predefined threshold (called minimum confidence) were output as association rules. In addition to the above approach, there are many other ones proposed for finding association rules.

Most mining approaches focus on binary valued transaction data. Transaction data in real-world applications, however, usually consist of quantitative values. Many sophisticated data-mining approaches have thus been proposed to deal with various types of data [6,46,53]. This also presents a challenge to workers in this research field.

In addition to proposing methods for mining association rules from transactions of binary values, Srikant et al. also proposed a method [46] for mining association rules from those with quantitative attributes. Their method first determines the number of partitions for each quantitative attribute, and then maps all possible values of each attribute into a set of consecutive integers. It then finds large itemsets whose support values are greater than the user-specified minimum-support levels. For example, the following is a quantitative association rule “If *Age* is [20, ..., 29], then *Number of Car* is [0, 1]” with a support value (60%) and a confidence value (66.6%). This means that if the age of a person is between 20 to 29-years old, then he/she has zero or one car with 66.6%. Of course, different partition approaches for discretizing the quantitative values may influence the final quantitative association rules. Some researches have thus been proposed for discussing and solving this problem [6,53]. Recently, fuzzy sets have also been used in data mining to handle quantitative data due to its ability to deal with the interval boundary problem. The theory of fuzzy sets will be introduced below.

## Fuzzy Sets

Fuzzy set theory was first proposed by Zadeh in 1965 [51]. It is primarily concerned with quantifying and reasoning using natural language in which words can have ambiguous meanings. It is widely used in a variety of fields be-

cause of its simplicity and similarity to human reasoning [13,45,52]. For example, the theory has been applied in fields such as manufacturing, engineering, diagnosis, economics, among others [19,29,37].

Fuzzy set theory can be thought of as an extension of traditional crisp sets, in which each element must either be in or not in a set. Formally, the process by which individuals from a universal set  $X$  are determined to be either members or nonmembers of a crisp set can be defined by a *characteristic or discrimination function* [51]. For a given crisp set  $A$ , this function assigns a value  $\mu_A(x)$  to every  $x \in X$  such that

$$\mu_A(x) = \begin{cases} 1 & \text{if and only if } x \in A \\ 0 & \text{if and only if } x \notin A \end{cases}.$$

The function thus maps elements of the universal set to the set containing 0 and 1. This kind of function can be generalized such that the values assigned to the elements of the universal set fall within specified ranges, referred to as the membership grades of these elements in the set. Larger values denote higher degrees of set membership. Such a function is called a membership function,  $\mu_A(x)$ , by which a fuzzy set  $A$  is usually defined. This function is represented by

$$\mu_A: X \rightarrow [0, 1],$$

where  $[0, 1]$  denotes the interval of real numbers from 0 to 1, inclusive. The function can also be generalized to any real interval instead of  $[0, 1]$ .

A special notation is often used in the literature to represent fuzzy sets. Assume that  $x_1$  to  $x_n$  are the elements in fuzzy set  $A$ , and  $\mu_1$  to  $\mu_n$  are, respectively, their grades of membership in  $A$ .  $A$  is then represented as follows:

$$A = \mu_1/x_1 + \mu_2/x_2 + \cdots + \mu_n/x_n.$$

An  $\alpha$ -cut of a fuzzy set  $A$  is a crisp set  $A_\alpha$  that contains all the elements in the universal set  $X$  with their membership grades in  $A$  greater than or equal to a specified value of  $\alpha$ . This definition can be written as

$$A_\alpha = \{x \in X \mid \mu_A(x) \geq \alpha\}.$$

The scalar cardinality of a fuzzy set  $A$  defined on a finite universal set  $X$  is the summation of the membership grades of all the elements of  $X$  in  $A$ . Thus,

$$|A| = \sum_{x \in X} \mu_A(x).$$

Three basic and commonly used operations on fuzzy sets are complementation, union and intersection, as proposed by Zadeh. They are described as follows.

1. The complementation of a fuzzy set  $A$  is denoted by  $\neg A$ , and the membership function of  $\neg A$  is given by

$$\mu_{\neg A}(x) = 1 - \mu_A(x) \quad \forall x \in X.$$

2. The intersection of two fuzzy sets  $A$  and  $B$  is denoted by  $A \cap B$ , and the membership function of  $A \cap B$  is given by

$$\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} \quad \forall x \in X.$$

3. The union of two fuzzy sets  $A$  and  $B$  is denoted by  $A \cup B$ , and the membership function of  $A \cup B$  is given by

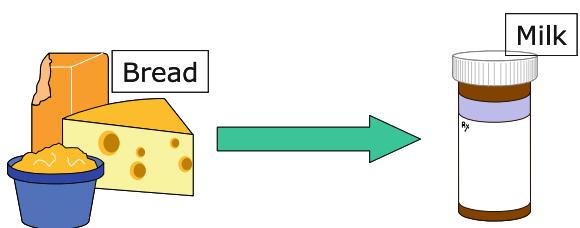
$$\mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\} \quad \forall x \in X.$$

Note that there are other calculation formula for the complementation, union and intersection, but the above are the most popular.

## Fuzzy Data Mining

As mentioned above, the fuzzy set theory is a natural way to process quantitative data. Several fuzzy learning algorithms for inducing rules from given sets of data have thus been designed and used to good effect with specific domains [7,22,41]. Fuzzy data mining approaches have also been developed to find knowledge with linguistic terms from quantitative transaction data. The knowledge obtained is expected to be easy to understand. A fuzzy association rule is shown in Fig. 2.

In Fig. 2, instead of quantitative intervals used in quantitative association rules, linguistic terms are used to represent the knowledge. As we can observe from the rule “If *middle* amount of *bread* is bought, then *high* amount of *milk* is bought”, *bread* and *milk* are items, and *middle* and *high* are linguistic terms. The rule means that if the quan-



Genetic-Fuzzy Data Mining Techniques, Figure 2  
A fuzzy association rule

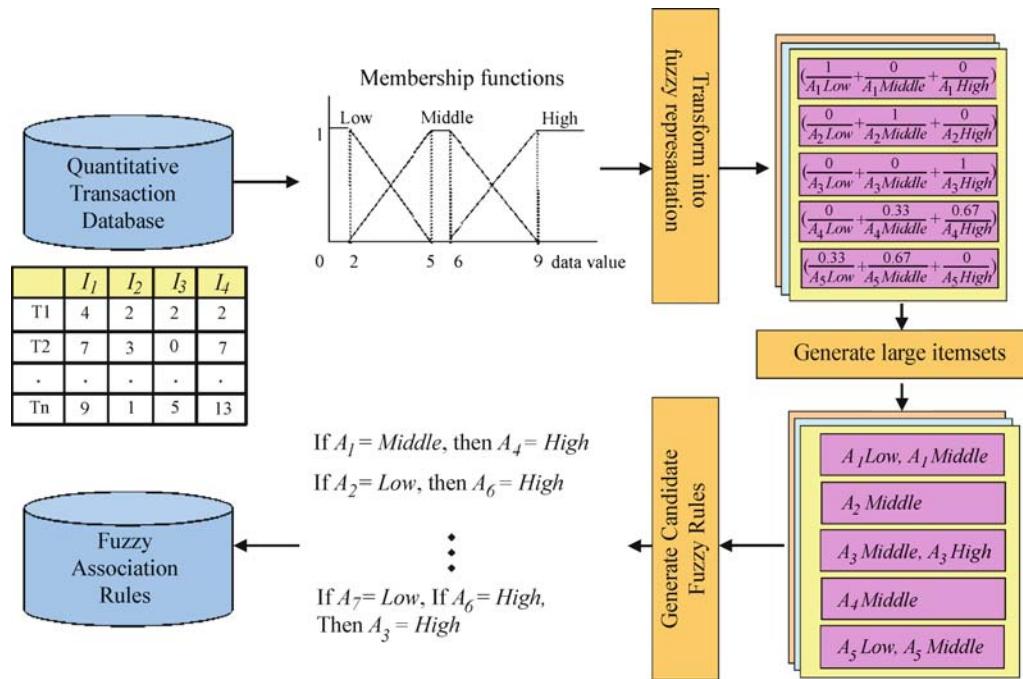
ity of the purchased item *bread* is *middle*, then there is a high possibility that the associated purchased item is *milk* with *high* quantity.

Many approaches have been proposed for mining fuzzy association rules [8,23,24,35,36,50]. Most of the approaches set a single minimum support threshold for all the items or itemsets and identify the association relationships among transactions. In real applications, different items may have different criteria to judge their importance. Multiple minimum support thresholds are thus proposed for this purpose. We can thus divide the fuzzy data mining approaches into two types, namely Single-minimum-Support Fuzzy Mining (SSFM) [8,23,24,35,50] and Multiple-minimum-Support Fuzzy Mining (MSFM) problems [36].

In the SSFM problem, Chan and Au proposed an F-APACS algorithm to mine fuzzy association rules [8]. They first transformed quantitative attribute values into linguistic terms and then used the adjusted difference analysis to find interesting associations among attributes. Kuok et al. proposed a mining approach for fuzzy association rules. Instead of minimum supports and minimum confidences used in most mining approaches, significance factors and certainty factors were used to derive large itemsets and fuzzy association rules [35]. At nearly the same time, Hong et al. proposed a fuzzy mining algorithm to mine fuzzy rules from quantitative transaction data [23]. Basically, these fuzzy mining algorithms first used membership functions to transform each quantitative value into a fuzzy set in linguistic terms and then used a fuzzy mining process to find fuzzy association rules. Yue et al. then extended the above concept to find fuzzy association rules with weighted items from transaction data [50]. They adopted Kohonen self-organized mapping to derive fuzzy sets for numerical attributes. In general, the basic concept of fuzzy data mining for the SSFM problem is shown in Fig. 3.

In Fig. 3, the process for fuzzy data mining first transforms quantity transactions into a fuzzy representation according to the predefined membership functions. The transformed data are then calculated to generate large itemsets. Finally, the generated large itemsets are used to derive fuzzy association rules.

As to the MSFM problem, Lee et al. proposed a mining algorithm which used multiple minimum supports to mine fuzzy association rules [36]. They assumed that items had different minimum supports and the maximum constraint was used. That is, the minimum support for an itemset was set as the maximum of the minimum supports of the items contained in the itemset. Under the constraint, the characteristic of level-by-level processing was



**Genetic-Fuzzy Data Mining Techniques, Figure 3**  
The concept of fuzzy data mining for the SSFM problem

kept, such that the original a priori algorithm could easily be extended to finding large itemsets. In addition to the maximum constraint, other constraints such as the minimum constraint can also be used with different rationale.

In addition to the above fuzzy mining approaches, fuzzy data mining with taxonomy or fuzzy taxonomy has been developed. Fuzzy web mining is another application of it. Besides, fuzzy data mining is strongly related to fuzzy control, fuzzy clustering, and fuzzy learning.

In most fuzzy mining approaches, the membership functions are usually predefined in advance. Membership functions are, however, very crucial to the final mined results. Below, we will describe how the genetic algorithm can be combined with fuzzy data mining to make the entire process more complete. The concept of the genetic algorithm will first be briefly introduced in the next section.

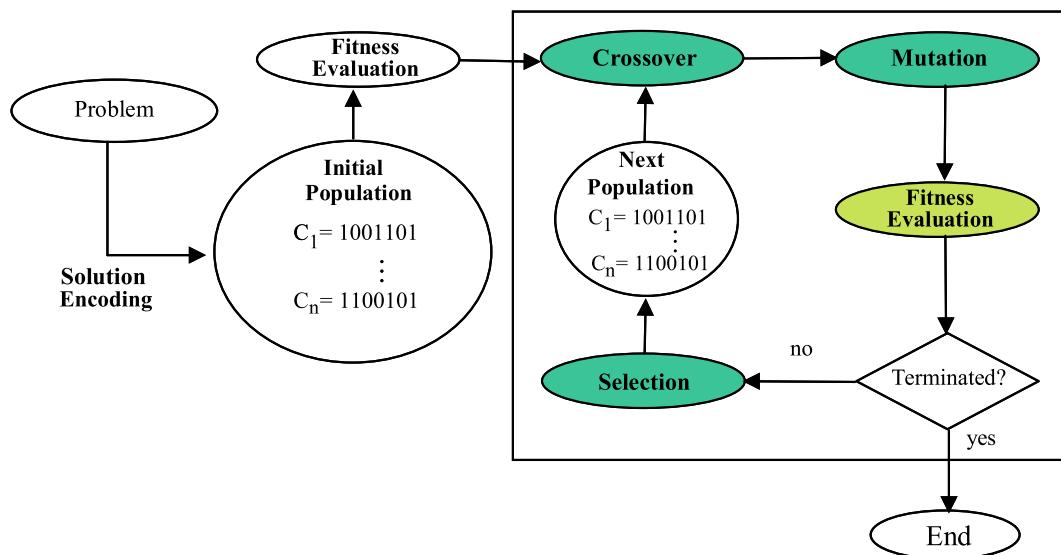
### Genetic Algorithms

Genetic Algorithms (GAs) [17,20] have become increasingly important for researchers in solving difficult problems since they could provide feasible solutions in a limited amount of time [21]. They were first proposed by Holland in 1975 [20] and have been successfully applied to the fields of optimization [17,39,40,43], machine learning [17,39], neural networks [40], fuzzy logic con-

trollers [43], and so on. GAs are developed mainly based on the ideas and techniques from genetic and evolutionary theory [20]. According to the principle of survival of the fittest, they generate the next population by several operations, with each individual in the population representing a possible solution. There are three principal operations in a genetic algorithm.

1. The *crossover* operation: it generates offspring from two chosen individuals in the population by exchanging some bits in the two individuals. The offspring thus inherit some characteristics from each parent.
2. The *mutation* operation: it generates offspring by randomly changing one or several bits in an individual. The offspring may thus possess different characteristics from their parents. Mutation prevents local searches of the search space and increases the probability of finding global optima.
3. The *selection* operation: it chooses some offspring for survival according to predefined rules. This keeps the population size within a fixed constant and puts good offspring into the next generation with a high probability.

On applying genetic algorithms to solving a problem, the first step is to define a representation that describes the



**Genetic-Fuzzy Data Mining Techniques, Figure 4**

The entire GA process

problem states. The most common way used is the bit string representation. An initial population of *individuals*, called *chromosomes*, is then defined and the three genetic operations (crossover, mutation, and selection) are performed to generate the next generation. Each chromosome in the population is evaluated by a *fitness function* to determine its goodness. This procedure is repeated until a user-specified termination criterion is satisfied. The entire GA process is shown in Fig. 4.

### Genetic-Fuzzy Data Mining Techniques

In the previous section for fuzzy data mining, several approaches were introduced, in which the membership functions were assumed to be known in advance. The given membership functions may, however, have a critical influence on the final mining results. Although many approaches for learning membership functions were proposed [14,42,44,47,48], most of them were usually used for classification or control problems. There were several strategies proposed for learning membership functions in classification or control problems by genetic algorithms. Below are some of them:

1. Learning membership functions first, then rules;
2. Learning rules first, then membership functions;
3. Simultaneously learning rules and membership functions;
4. Iteratively learning rules and membership functions.

For fuzzy mining problems, many researches have also been done by combining the genetic algorithm and the fuzzy concepts to discover both suitable membership functions and useful fuzzy association rules from quantitative values. However, most of them adopt the first strategy. That is, the membership functions are first learned and then the fuzzy association rules are derived based on the obtained membership functions. It is done in this way because the number of association rules is often large in mining problems and can not easily be coded in a chromosome.

In this article, we introduce several genetic-fuzzy data mining algorithms that can mine both appropriate membership functions and fuzzy association rules [9,10,11,12,25,26,27,31,32,33]. The genetic-fuzzy mining problems can be divided into four kinds according to the types of fuzzy mining problems and the ways of processing items. The types of fuzzy mining problems include Single-minimum-Support Fuzzy-Mining (SSFM) and Multiple-Minimum-Support Fuzzy-Mining (MSFM) as mentioned above. The ways of processing items include processing all the items together (integrated approach) and processing them individually (divide-and-conquer approach). The integrated genetic-fuzzy approaches encode all membership functions of all items (or attributes) into a chromosome (also called an individual). The genetic algorithms are then used to derive a set of appropriate membership functions according to the designed fitness function. Finally, the best set of membership functions are then used to mine fuzzy

**Genetic-Fuzzy Data Mining Techniques, Table 1**  
The four different genetic-fuzzy data mining problems

	Integrated approach	Divide-and-conquer approach
Single minimum support	IGFSMS Problem	DGFMS Problem
Multiple minimum supports	IGFMMS Problem	DGFMMMS Problem

association rules. On the other hand, the divide-and-conquer genetic-fuzzy approaches encode membership functions of each item into a chromosome. In other words, chromosomes in a population were maintained just for only one item. The membership functions can thus be found for one item after another or at the same time via parallel processing. In general, the chromosomes in the divide-and-conquer genetic-fuzzy approaches are much shorter than those in the integrated approaches since the former only focus on individual items. But there are more application limitations on the former than on the latter. This will be explained later.

The four kinds of problems are thus the Integrated Genetic-Fuzzy problem for items with a Single Minimum Support (IGFSMS) [9,11,26,31,32,33], the Integrated Genetic-Fuzzy problem for items with Multiple Minimum Supports (IGFMMS) [12], the Divide-and-Conquer Genetic-Fuzzy problem for items with a Single Minimum Support (DGFMS) [10,25,27] and the Divide-and-Conquer Genetic-Fuzzy problem for items with Multiple Minimum Supports (DGFMMMS). The classification is shown in Table 1.

Each of the four kinds of genetic-fuzzy data mining problems will be introduced in the following sections.

### The Integrated Genetic-Fuzzy Problem for Items with a Single Minimum Support (IGFSMS)

Many approaches have been published for solving the IGFSMS problem [9,11,26,31,32,33]. For example, Hong et al. proposed a genetic-fuzzy data-mining algorithm for extracting both association rules and membership functions from quantitative transactions [26]. They proposed a GA-based framework for searching membership functions suitable for given mining problems and then use the final best set of membership functions to mine fuzzy association rules. The proposed framework is shown in Fig. 5.

The proposed framework consists of two phases, namely mining membership functions and mining fuzzy association rules. In the first phase, the proposed framework maintains a population of sets of membership func-

tions, and uses the genetic algorithm to automatically derive the resulting one. It first transforms each set of membership functions into a fixed-length string. The chromosome is then evaluated by the number of large 1-itemsets and the suitability of membership functions. The fitness value of a chromosome  $C_q$  is then defined as

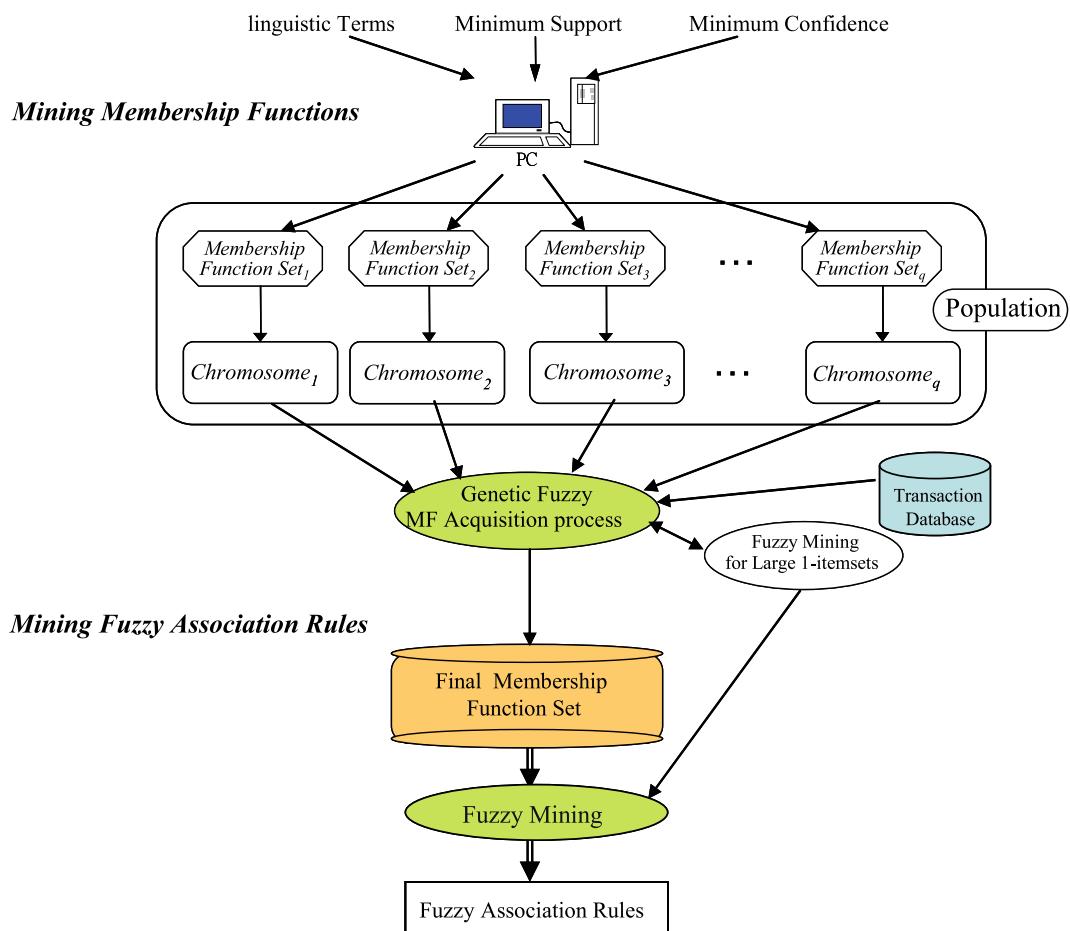
$$f(C_q) = \frac{|L_1|}{\text{suitability}(C_q)},$$

where  $|L_1|$  is the number of large 1-itemsets obtained by using the set of membership functions in  $C_q$ . Using the number of large 1-itemsets can achieve a trade-off between execution time and rule interestingness. Usually, a larger number of 1-itemsets will result in a larger number of all itemsets with a higher probability, which will thus usually imply more interesting association rules. The evaluation by 1-itemsets is, however, faster than that by all itemsets or interesting association rules. Of course, the number of all itemsets or interesting association rules can also be used in the fitness function. A discussion for different choices of fitness functions can be found in [9].

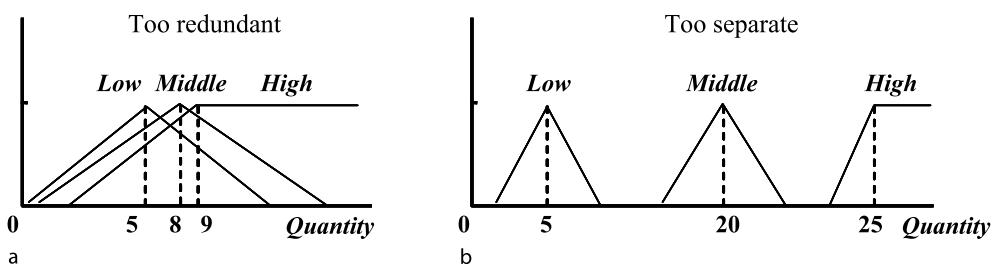
The suitability measure is used to reduce the occurrence of bad types of membership functions. The two bad types of membership functions are shown in Fig. 6, where the first one is too redundant, and the second one is too separate.

Two factors, called the overlap factor and the coverage factor, are used to avoid the bad shapes. The overlap factor is designed for avoiding the first bad case (too redundant), and the coverage factor is for the second one (too separate). Each factor has its formula for evaluating a value from a chromosome. After fitness evaluation, the approach then chooses appropriate chromosomes for mating, gradually creating good offspring membership function sets. The offspring membership function sets then undergo recursive evolution until a good set of membership functions has been obtained. In the second phase, the final best membership functions are gathered to mine fuzzy association rules. The fuzzy mining algorithm proposed in [24] is adopted to achieve this purpose.

The calculation for large 1-itemsets, however, will still take a lot of time, especially when the database can not totally be fed into the main memory. An enhanced approach, called the cluster-based fuzzy-genetic mining algorithm was thus proposed [11] to speed up the evaluation process and keep nearly the same quality of solutions as that in [26]. That approach also maintains a population of sets of membership functions and uses the genetic algorithm to derive the best one. Before fitness evaluation, the clustering technique is first used to cluster chromosomes. It uses the  $k$ -means clustering approach to gather similar



**Genetic-Fuzzy Data Mining Techniques, Figure 5**  
A genetic-fuzzy framework for the IGFSMS problem



**Genetic-Fuzzy Data Mining Techniques, Figure 6**  
The two bad types of membership functions

chromosomes into groups. The two factors, overlap factor and coverage factor, are used as two attributes for clustering. For example, coverage and overlap factors for ten chromosomes are shown in Table 2, where the column “Suitability” represents the pair (coverage factor, overlap factor).

The  $k$ -means clustering approach is then executed to divide the ten chromosomes into  $k$  clusters. In this example, assume the parameter  $k$  is set at 3. The three clusters found are shown in Table 3. The representative chromosomes in the three clusters are  $C_5$  (4.37, 0.33),  $C_4$  (4.66, 0) and  $C_9$  (4.09, 8.33).

**Genetic-Fuzzy Data Mining Techniques, Table 2**  
The coverage and the overlap factors of ten chromosomes

Chromosome	Suitability	Chromosome	Suitability
$C_1$	(4, 0)	$C_6$	(4.5, 0)
$C_2$	(4.24, 0.5)	$C_7$	(4.45, 0)
$C_3$	(4.37, 0)	$C_8$	(4.37, 0.53)
$C_4$	(4.66, 0)	$C_9$	(4.09, 8.33)
$C_5$	(4.37, 0.33)	$C_{10}$	(4.87, 0)

**Genetic-Fuzzy Data Mining Techniques, Table 3**  
The three clusters found in the example

Cluster <sub>i</sub>	Chromosomes	Representative chromosome
Cluster <sub>1</sub>	$C_1, C_2, C_5, C_8$	$C_5$
Cluster <sub>2</sub>	$C_3, C_4, C_6, C_7, C_{10}$	$C_4$
Cluster <sub>3</sub>	$C_9$	$C_9$

All the chromosomes in a cluster use the number of large 1-itemsets derived from the representative chromosome in the cluster and their own suitability of membership functions to calculate their fitness values. Since the number for scanning a database decreases, the evaluation cost can thus be reduced. In this example, the representative chromosomes are chromosomes  $C_4$ ,  $C_5$ ,  $C_9$  and it only needs to calculate the number of large 1-itemsets three times. The evaluation results are utilized to choose appropriate chromosomes for mating in the next generation. The offspring membership function sets then undergo recursive evolution until a good set of membership functions has been obtained. Finally, the derived membership functions are used to mine fuzzy association rules.

Kaya and Alhaji also proposed several genetic-fuzzy data mining approaches to derive membership functions and fuzzy association rules [31,32,33]. In [31], the proposed approach tries to derive membership functions, which can get a maximum profit within an interval of user-specified minimum support values. It then uses the derived membership functions to mine fuzzy association rules. The concept of their approaches is shown in Fig. 7.

As shown in Fig. 7a, the approach first derives membership functions from the given quantitative transaction database by genetic algorithms. The final membership functions are then used to mine fuzzy association rules. Figure 7b shows the concept of maximizing the large itemsets of the given minimum support interval. It is used as the fitness function. Kaya and Alhaji also extended the approach to mine fuzzy weighted association rules [32]. Furthermore, fitness functions are not easily defined for GA applications, such that multiobjective genetic algorithms have also been developed [28,30]. In other words, more

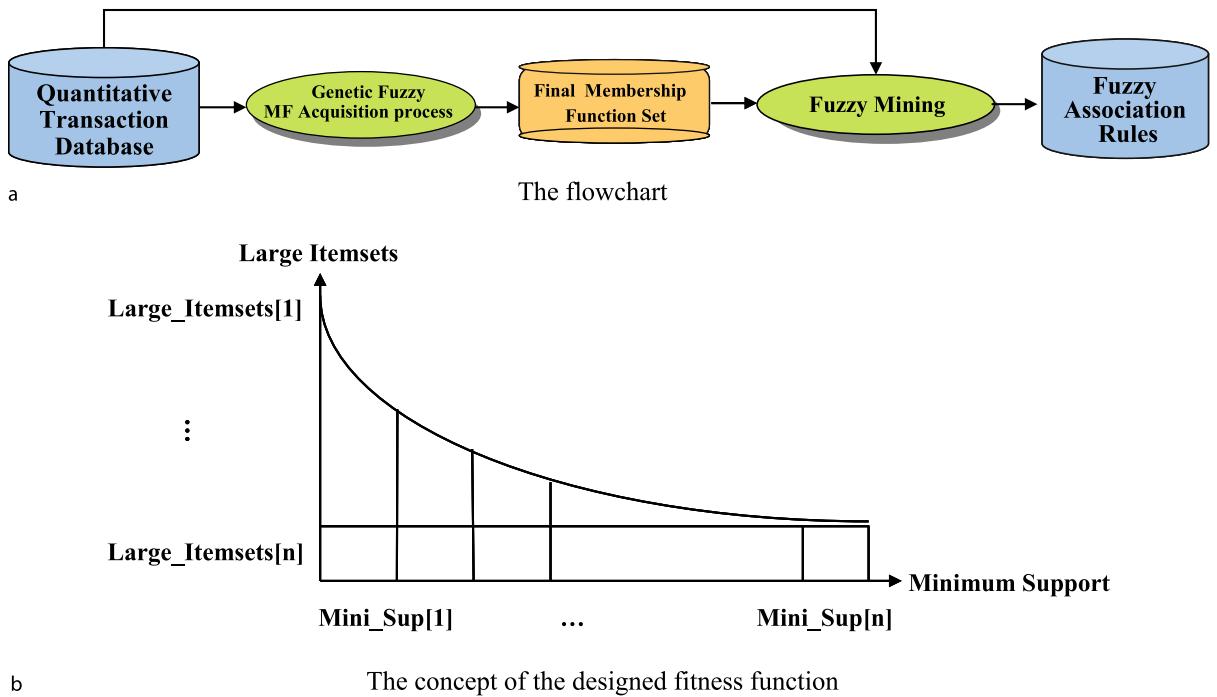
than one criterion is used in the evaluation. A set of solutions, namely nondominated points (also called Pareto-Optimal Surface), is derived and given to users, instead of only the one best solution obtained by genetic algorithms. Kaya and Alhaji thus proposed an approach based on multiobjective genetic algorithms to learn membership functions, which were then used to generate interesting fuzzy association rules [33]. Three objective functions, namely strength, interestingness and comprehensibility, were used in their approach to find the Pareto-Optimal Surface. In addition to the above approaches for the IGFSMS problem, some others are still in progress.

### The Integrated Genetic-Fuzzy Problem for Items with Multiple Minimum Supports (IGFMMS)

In the above subsection, it can be seen that lots of researches focus on integrated genetic-fuzzy approaches for items with a single minimum support. However, different items may have different criteria to judge their importance. For example, assume among a set of items there are some which are expensive. They are thus seldom bought because of their high cost. Besides, the support values of these items are low. A manager may, however, still be interested in these products due to their high profits. In such cases, the above approaches may not be suitable for this problem. Chen et al. thus proposed another genetic-fuzzy data mining approach [12], which was an extension of the approach proposed in [26], to solve it. The approach combines the clustering, fuzzy and genetic concepts to derive minimum support values and membership functions for items. The final minimum support values and membership functions are then used to mine fuzzy association rules. The genetic-fuzzy mining framework for the IGFMMS problem is shown in Fig. 8.

As shown in Fig. 8, the framework can be divided into two phases. The first phase searches for suitable minimum support values and membership functions of items and the second phase uses the final best set of minimum support values and membership functions to mine fuzzy association rules. The proposed framework maintains a population of sets of minimum support values and membership functions, and uses the genetic algorithm to automatically derive the resulting one.

A genetic algorithm requires a population of feasible solutions to be initialized and updated during the evolution process. As mentioned above, each individual within the population is a set of minimum support values and isosceles-triangular membership functions. Each membership function corresponds to a linguistic term of a certain item. In this approach, the initial set of chromosomes



**Genetic-Fuzzy Data Mining Techniques, Figure 7**  
The concept of Kaya and Alhaji's approaches

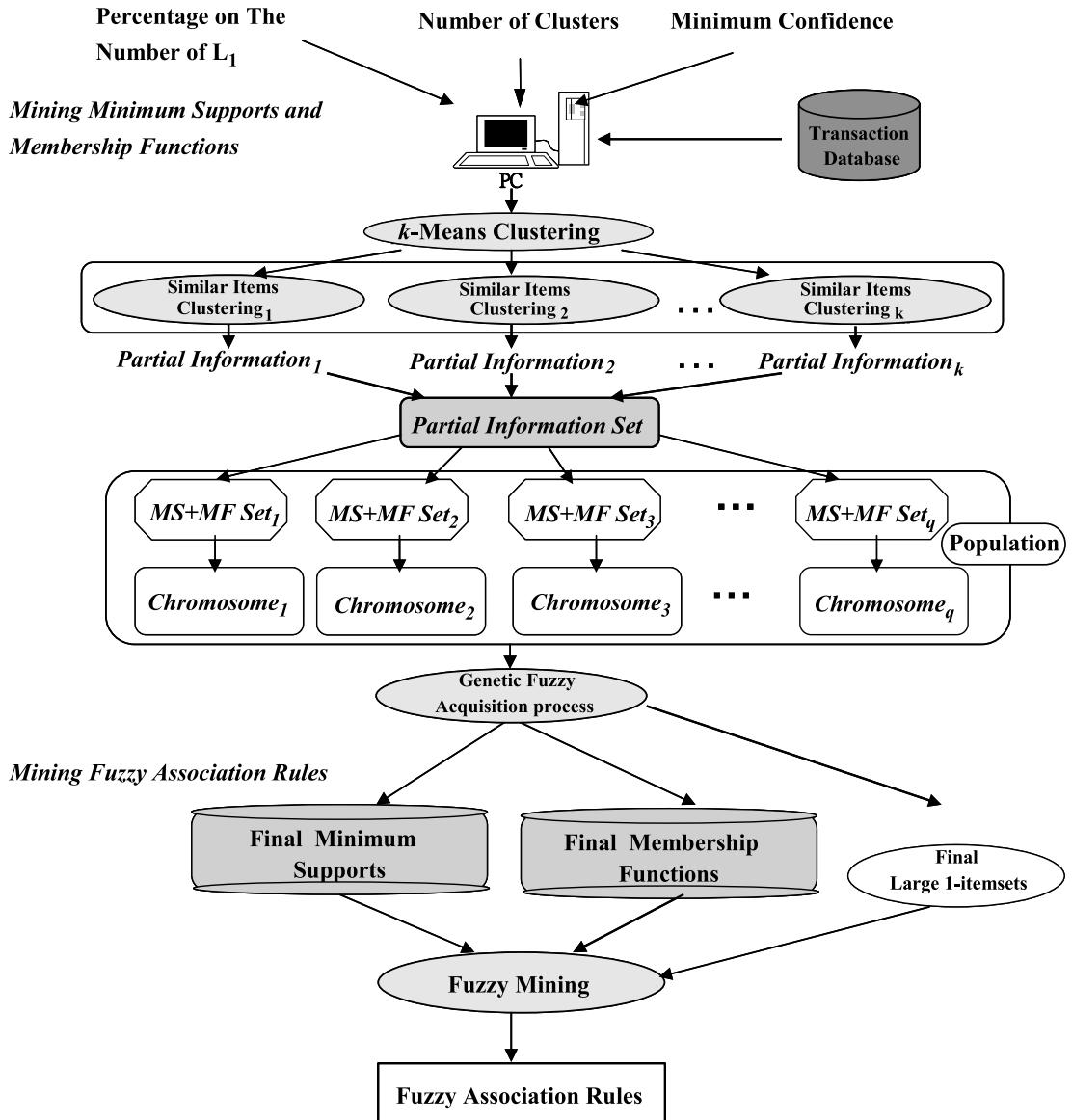
is generated based on the initialization information derived by the  $k$ -means clustering approach on the transactions. The frequencies and quantitative values of items in the transactions are the two main factors to gather similar items into groups. The initialization information includes an appropriate number of linguistic terms, the range of possible minimum support values and membership functions of each item. All the items in the same cluster are considered to have similar characteristics and are assigned similar initialization values when a population is initialized. The approach then generates and encodes each set of minimum support values and membership functions into a fixed-length string according to the initialization information.

In this approach, the minimum support values of items may be different. It is hard to assign the values. As an alternative, the values can be determined according to the required number of rules. It is, however, very time-consuming to obtain the rules for each chromosome. As mentioned above, a larger number of 1-itemsets will usually result in a larger number of all itemsets with a higher probability, which will thus usually imply more interesting association rules. The evaluation by 1-itemsets is faster than that by all itemsets or interesting association rules. Using the number of large 1-itemsets can thus achieve a trade-off between execution time and rule interestingness [26].

A criterion should thus be specified to reflect the user preference on the derived knowledge. In the approach, the required number of large 1-itemsets  $RNL$  is used for this purpose. It is the number of linguistic large 1-itemsets that a user wants to get from an item. It can be defined as the number of linguistic terms of an item multiplied by the predefined percentage which reflects users' preference on the number of large 1-itemsets. It is used to reflect the closeness degree between the number of derived large 1-itemsets and the required number of large 1-itemsets. For example, assume there are three linguistic terms for an item and the predefined percentage  $p$  is set at 80%. The  $RNL$  value is then set as  $\lfloor 3 \times 0.8 \rfloor$ , which is 2. The fitness function is then composed of the suitability of membership functions and the closeness to the  $RNL$  value. The minimum support values and membership functions can thus be derived by GA and are then used to mine fuzzy association rules by a fuzzy mining approach for multiple minimum supports such as the one in [36].

#### The Divide-and-Conquer Genetic-Fuzzy Problem for Items with a Single Minimum Support (DGFSMS)

The advantages of the integrated genetic-fuzzy approaches lie in that they are simple, easy to use, and with few constraints in the fitness functions. In addition to the num-

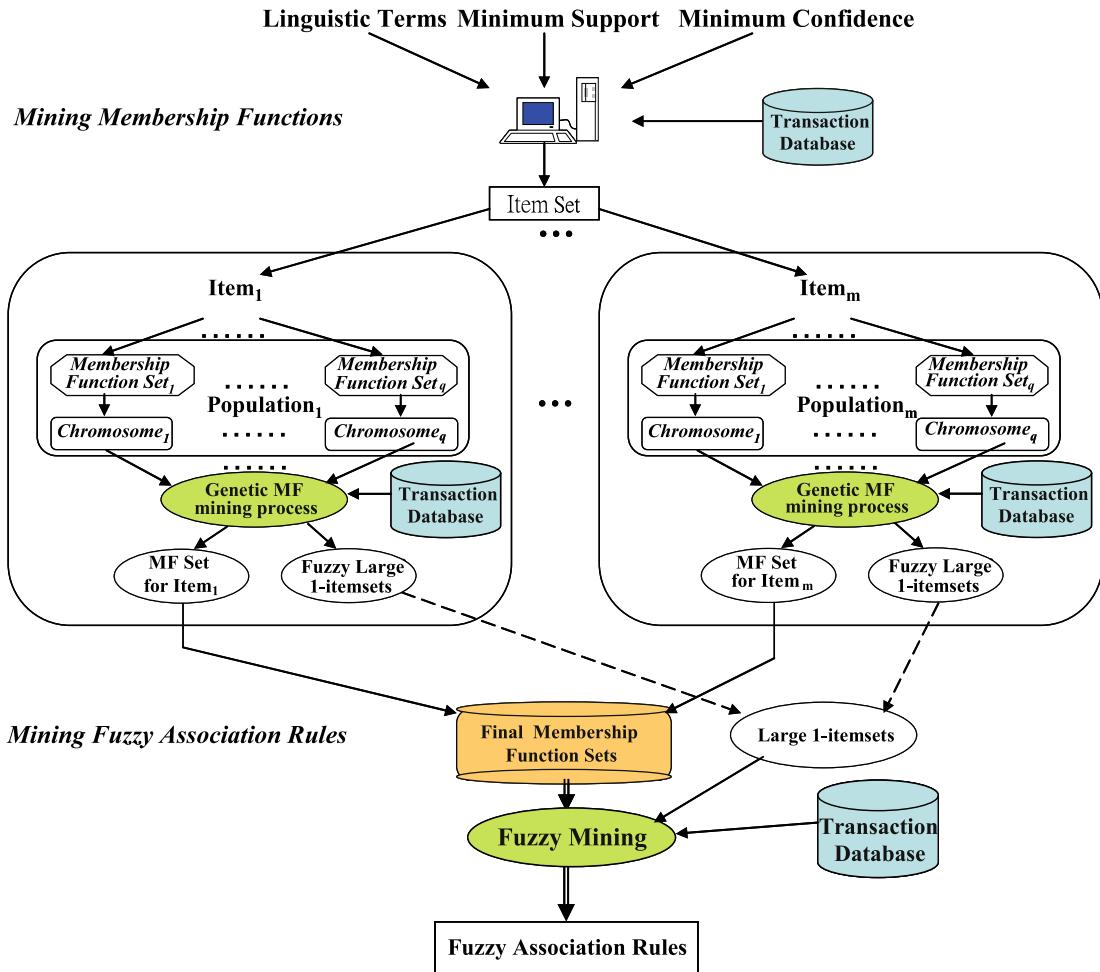


**Genetic-Fuzzy Data Mining Techniques, Figure 8**  
A genetic-fuzzy framework for the IGFMMMS problem

ber of large 1-itemsets, the other criteria can also be used. However, if the number of items is large, the integrated genetic-fuzzy approaches may need lots of time to find a near-optimal solution because the length of a chromosome is very long. Recently, the divide-and-conquer strategy has been used in the evolutionary computation community to very good effect. Many algorithms based on it have also been proposed in different applications [5,15, 34,49]. When the number of large 1-itemsets is used in fitness evaluation, the divide-and-conquer strategy becomes

a good choice to deal with it since each item can be individually processed in this situation. Hong et al. thus used a GA-based framework with the divide-and-conquer strategy to search for membership functions suitable for the mining problem [27]. The framework is shown in Fig. 9.

The proposed framework in Fig. 9 is divided into two phases: mining membership functions and mining fuzzy association rules. Assume the number of items is *m*. In the phase of mining membership functions, it maintains *m* populations of membership functions, with each popula-



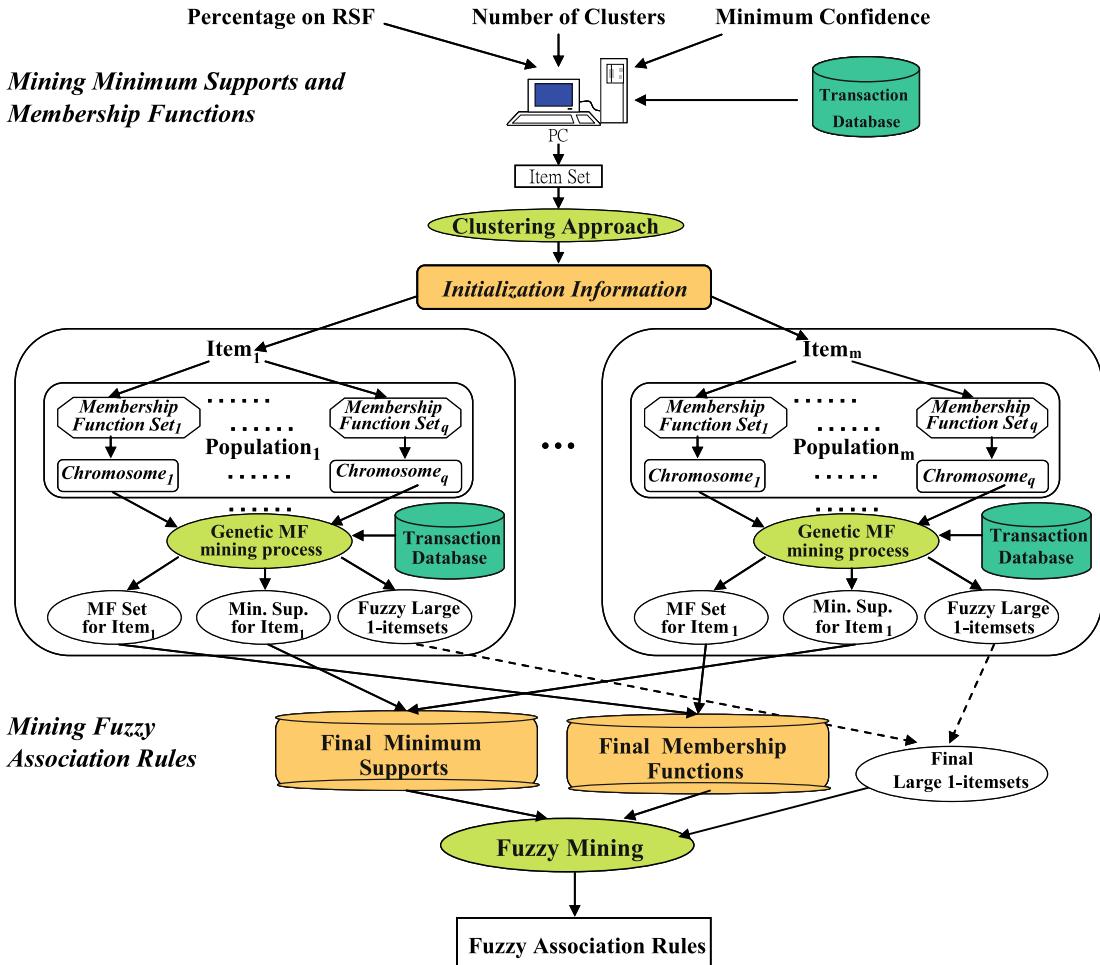
**Genetic-Fuzzy Data Mining Techniques, Figure 9**  
A genetic-fuzzy framework for the DGFSMS problem

tion for an item. Each chromosome in a population represents a possible set of membership functions for that item. The chromosomes in the same population are of the same length. The fitness of each set of membership functions is evaluated by the fuzzy-supports of the linguistic terms in large 1-itemsets and by the suitability of the derived membership functions. The offspring sets of membership functions undergo recursive evolution until a good set of membership functions has been obtained. Next, in the phase of mining fuzzy association rules, the sets of membership function for all the items are gathered together and used to mine the fuzzy association rules from the given quantitative database. An enhanced approach [10], which combines the clustering and the divide-and-conquer techniques, was also proposed to speed up the evaluation process. The clustering idea is similar to that in IGFSMS [11] except that the center value of each membership function

is also used as an attribute to cluster chromosomes. The clustering process is thus executed according to the coverage factors, the overlap factors and the center values of chromosomes. For example, assume each item has three membership functions. In total five attributes including one coverage factor, one overlap factor and three center values, are used to form appropriate clusters. Note that the number of linguistic terms for each item is predefined in the mentioned approaches. It may also be automatically and dynamically adjusted [25].

#### The Divide-and-Conquer Genetic-Fuzzy Problem for Items with Multiple Minimum Supports (DGFMS)

The problem may be thought of as the combination of the IGFMMS and the DGFSMS problems. The framework for the DGFMS problem can thus be easily designed from



**Genetic-Fuzzy Data Mining Techniques, Figure 10**  
A genetic-fuzzy framework for the DGFMMMS problem

the previous frameworks for IGFMMMS and DGFSMS. It is shown in Fig. 10.

The proposed framework in Fig. 10 is divided into two phases: mining minimum supports and membership functions, and mining fuzzy association rules. In the first phase, the clustering approach is first used for deriving initialization information which is then used for obtaining better initial populations as used for IGFMMMS. It then maintains  $m$  populations of minimum supports and membership functions, with each population for an item. Next, in the phase of mining fuzzy association rules, the minimum support values and membership functions for all the items are gathered together and are used to mine fuzzy interesting association rules from the given quantitative database.

### Future Directions

In this article, we have introduced some genetic-fuzzy data mining techniques and their classification. The concept of fuzzy sets is used to handle quantitative transactions and the process of genetic calculation is executed to find appropriate membership functions. The genetic-fuzzy mining problems are divided into four kinds according to the types of fuzzy mining problems and the ways of processing items. The types of fuzzy mining problems include Single-minimum-Support Fuzzy-Mining (SSFM) and Multiple-minimum-Support Fuzzy-Mining (MSFM). The methods of processing items include processing all the items together (integrated approach) and processing them individually (divide-and-conquer approach). Each of the

four kinds of problems has been described with some approaches given.

Data mining is very important especially because the data amounts in the information era are extremely large. The topic will continuously grow but with a variety of forms. Some possible research directions in the future about genetic-fuzzy data mining are listed as follows.

- **Applying multiobjective genetic algorithms to the genetic-fuzzy mining problems:** In the article, mined knowledge (number of large itemsets or number of rules) and suitability of membership functions are two important factors used in genetic-fuzzy data mining. Analyzing the relationship between the two factors is thus an interesting and important task. Besides, multiobjective genetic algorithms can also be used to consider the factors at the same time.
- **Analyzing effects of different shapes of membership functions and different genetic operators:** Different shapes of membership functions may have different results on genetic-fuzzy data mining. They may be evaluated in the future. Different genetic operations may also be tried to obtain better results than the ones used in the above approaches.
- **Enhancing performance of the fuzzy-rule mining phase:** The final goal of genetic-fuzzy mining techniques introduced in this article is to mine appropriate fuzzy association rules. However, the phase of mining fuzzy association rules is very time-consuming. How to improve the process of mining interesting fuzzy rules is thus worth studying. Some possible approaches include modifying existing approaches, combining the existing ones with other techniques, or defining new evaluation criteria.
- **Developing visual tools for these genetic-fuzzy mining approaches:** Another interesting aspect of future work is to develop visual tools for demonstrating the genetic-fuzzy mining results. It can help a decision maker easily understand or get useful information quickly. The visual tools may include, for example, how to show the derived membership functions and to illustrate the interesting fuzzy association rules.

## Bibliography

1. Agrawal R, Srikant R (1994) Fast algorithm for mining association rules. In: Proceedings of the international conference on very large data bases, pp 487–499
2. Agrawal R, Imielinski T, Swami A (1993) Database mining: a performance perspective. *Trans IEEE Knowl Data Eng* 5(6):914–925
3. Agrawal R, Imielinski T, Swami A (1993) Mining association rules between sets of items in large database. In: Proceedings of the conference ACMSIGMOD, Washington DC, USA
4. Agrawal R, Srikant R, Vu Q (1997) Mining association rules with item constraints. In: Proceedings of the third international conference on knowledge discovery in databases and data mining, Newport Beach, California, August 1997
5. Au WH, Chan KCC, Yao X (2003) A novel evolutionary data mining algorithm with applications to churn prediction. *Trans IEEE Evol Comput* 7(6):532–545
6. Aumann Y, Lindell Y (1999) A statistical theory for quantitative association rules. In: Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining, pp 261–270
7. Casillas J, Cordón O, del Jesus MJ, Herrera F (2005) Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction. *Trans IEEE Fuzzy Syst* 13(1):13–29
8. Chan CC, Au WH (1997) Mining fuzzy association rules. In: Proceedings of the conference on information and knowledge management, Las Vegas, pp 209–215
9. Chen CH, Hong TP, Vincent Tseng S (2007) A comparison of different fitness functions for extracting membership functions used in fuzzy data mining. In: Proceedings of the symposium IEEE on foundations of computational intelligence, pp 550–555
10. Chen CH, Hong TP, Tseng VS (2007) A modified approach to speed up genetic-fuzzy data mining with divide-and-conquer strategy. In: Proceedings of the congress IEEE on evolutionary computation (CEC), pp 1–6
11. Chen CH, Tseng VS, Hong TP (2008) Cluster-based evaluation in fuzzy-genetic data mining. *Trans IEEE Fuzzy Syst* 16(1):249–262
12. Chen CH, Hong TP, Tseng VS, Lee CS (2008) A genetic-fuzzy mining approach for items with multiple minimum supports. Accepted and to appear in Soft Computing (SCI)
13. Chen J, Mikulcic A, Kraft DH (2000) An integrated approach to information retrieval with fuzzy clustering and fuzzy inferencing. In: Pons O, Vila MA, Kacprzyk J (eds) Knowledge management in fuzzy databases. Physica, Heidelberg
14. Cordón O, Herrera F, Villar P (2001) Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base. *Trans IEEE Fuzzy Syst* 9(4):667–674
15. Darwen PJ, Yao X (1997) Speciation as automatic categorical modularization. *Trans IEEE Evol Comput* 1(2):101–108
16. Frawley WJ, Piatetsky-Shapiro G, Matheus CJ (1991) Knowledge discovery in databases: an overview. In: Proceedings of the workshop AAAI on knowledge discovery in databases, pp 1–27
17. Goldberg DE (1989) Genetic algorithms in search, optimization and machine learning. Addison Wesley, Boston
18. Grefenstette JJ (1986) Optimization of control parameters for genetic algorithms. *Trans IEEE Syst Man Cybern* 16(1):122–128
19. Heng PA, Wong TT, Rong Y, Chui YP, Xie YM, Leung KS, Leung PC (2006) Intelligent inferencing and haptic simulation for Chinese acupuncture learning and training. *Trans IEEE Inf Technol Biomed* 10(1):28–41
20. Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan Press, Michigan
21. Homaifar A, Guan S, Liepins GE (1993) A new approach on the traveling salesman problem by genetic algorithms. In: Proceedings of the fifth international conference on genetic algorithms

22. Hong TP, Lee YC (2001) Mining coverage-based fuzzy rules by evolutionary computation. In: Proceedings of the international IEEE conference on data mining, pp 218–224
23. Hong TP, Kuo CS, Chi SC (1999) Mining association rules from quantitative data. *Intell Data Anal* 3(5):363–376
24. Hong TP, Kuo CS, Chi SC (2001) Trade-off between time complexity and number of rules for fuzzy mining from quantitative data. *Int J Uncertain Fuzziness Knowledge-Based Syst* 9(5):587–604
25. Hong TP, Chen CH, Wu YL, Tseng VS (2004) Finding active membership functions in fuzzy data mining. In: Proceedings of the workshop on foundations of data mining in the fourth international IEEE conference on data mining
26. Hong TP, Chen CH, Wu YL, Lee YC (2006) AGA-based fuzzy mining approach to achieve a trade-off between number of rules and suitability of membership functions. *Soft Comput* 10(11):1091–1101
27. Hong TP, Chen CH, Wu YL, Lee YC (2008) Genetic-fuzzy data mining with divide-and-conquer strategy. *Trans IEEE Evol Comput* 12(2):252–265
28. Ishibuchi H, Yamamoto T (2004) Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining. *Fuzzy Sets Syst* 141:59–88
29. Ishibuchi H, Yamamoto T (2005) Rule weight specification in fuzzy rule-based classification systems. *Trans IEEE Fuzzy Syst* 13(4):428–435
30. Jin Y (2006) Multi-objective machine learning. Springer, Berlin
31. Kaya M, Alhajj R (2003) A clustering algorithm with genetically optimized membership functions for fuzzy association rules mining. In: Proceedings of the international IEEE conference on fuzzy systems, pp 881–886
32. Kaya M, Alhajj R (2004) Genetic algorithms based optimization of membership functions for fuzzy weighted association rules mining. In: Proceedings of the international symposium on computers and communications, vol 1, pp 110–115
33. Kaya M, Alhajj R (2004) Integrating multi-objective genetic algorithms into clustering for fuzzy association rules mining. In: Proceedings of the fourth international IEEE conference on data mining, pp 431–434
34. Khare VR, Yao X, Sendhoff B, Jin Y, Wersing H (2005) Co-evolutionary modular neural networks for automatic problem decomposition. In: Proceedings of the (2005) congress IEEE on evolutionary computation, vol 3, pp 2691–2698
35. Kuok C, Fu A, Wong M (1998) Mining fuzzy association rules in databases, *Record SIGMOD* 27(1):41–46
36. Lee YC, Hong TP, Lin WY (2004) Mining fuzzy association rules with multiple minimum supports using maximum constraints. In: Lecture notes in computer science, vol 3214. Springer, Heidelberg, pp 1283–1290
37. Liang H, Wu Z, Wu Q (2002) A fuzzy based supply chain management decision support system. In: Proceedings of the world congress on intelligent control and automation, vol 4, pp 2617–2621
38. Mamdani EH (1974) Applications of fuzzy algorithms for control of simple dynamic plants. *Proc IEEE* 121(12):1585–1588
39. Michalewicz Z (1994) Genetic algorithms + data structures = evolution programs. Springer, New York
40. Mitchell M (1996) An introduction to genetic algorithms. MIT Press, Cambridge MA
41. Rasmani KA, Shen Q (2004) Modifying weighted fuzzy subset-hood-based rule models with fuzzy quantifiers. In: Proceedings of the international IEEE conference on fuzzy systems, vol 3, pp 1679–1684
42. Roubos H, Setnes M (2001) Compact and transparent fuzzy models and classifiers through iterative complexity reduction. *Trans IEEE Fuzzy Syst* 9(4):516–524
43. Sanchez E et al (1997) Genetic algorithms and fuzzy logic systems: soft computing perspectives (advances in fuzzy systems – applications and theory, vol 7). World-Scientific, River Edge
44. Setnes M, Roubos H (2000) GA-fuzzy modeling and classification: complexity and performance. *Trans IEEE Fuzzy Syst* 8(5):509–522
45. Siler W, James J (2004) Fuzzy expert systems and fuzzy reasoning. Wiley, New York
46. Srikant R, Agrawal R (1996) Mining quantitative association rules in large relational tables. In: Proceedings of the (1996) international ACM SIGMOD conference on management of data, Montreal, Canada, June 1996, pp 1–12
47. Wang CH, Hong TP, Tseng SS (1998) Integrating fuzzy knowledge by genetic algorithms. *Trans IEEE Evol Comput* 2(4):138–149
48. Wang CH, Hong TP, Tseng SS (2000) Integrating membership functions and fuzzy rule sets from multiple knowledge sources. *Fuzzy Sets Syst* 112:141–154
49. Yao X (2003) Adaptive divide-and-conquer using populations and ensembles. In: Proceedings of the (2003) international conference on machine learning and application, pp 13–20
50. Yue S, Tsang E, Yeung D, Shi D (2000) Mining fuzzy association rules with weighted items. In: Proceedings of the international IEEE conference on systems, man and cybernetics, pp 1906–1911
51. Zadeh LA (1965) Fuzzy set. *Inf Control* 8(3):338–353
52. Zhang H, Liu D (2006) Fuzzy modeling and fuzzy control. Springer, New York
53. Zhang Z, Lu Y, Zhang B (1997) An effective partitioning-combining algorithm for discovering quantitative association rules. In: Proceedings of the Pacific-Asia conference on knowledge discovery and data mining, pp 261–270

## Genome Organization

GUSTAVO GLUSMAN, ARIAN F. A. SMIT  
Institute for Systems Biology, Seattle, USA

### Article Outline

- [Glossary](#)
- [Definition](#)
- [Introduction](#)
- [Molecular Structure of the Genome](#)
- [Genome Size vs. Organismal Complexity](#)
- [The Genome as Habitat for Transposable Elements](#)
- [Genomes in Constant Flux](#)
- [Transcriptome Complexity](#)
- [Transcript Networks](#)
- [Future Directions](#)
- [Bibliography](#)

## Glossary

**Genome** The entire genetic complement of a species.

**Transcript** An RNA molecule derived by copying from a DNA template.

**Splicing** The process of removing introns from a transcript and thereby joining the exons into a mature RNA molecule.

**Alternative splicing** The ability to process a transcript by splicing in more than one way, leading to the production of alternative RNA products.

**Gene** In the simplest cases, a segment of DNA that specifies the structure of a protein or an RNA molecule. The definition becomes less straightforward in complex genomic loci able to produce several overlapping transcripts with potentially disparate functions.

**Recombination** A process by which two chromosomes exchange segments, typically during the formation of sperm and eggs.

**Gene conversion** A process by which a segment of one chromosome is copied to another chromosome, replacing the original sequence.

## Definition

With rare exceptions, all known living organisms encode their genetic material in the form of double-stranded DNA, in one or more chromosomes, collectively referred to as the “genome”. A cell lacking its genome cannot survive for long, since it cannot produce new transcripts in response to environmental challenges. For example, red blood cells lose their DNA and can only function as oxygen shuttles until they break down. The genome includes most of the information needed by the cells to stay alive, to differentiate into new cell types, and to perform their functions in the context of the organism. As such, it is the ultimate resource for identifying the full set of components in the living system. Eukaryotic genomes are much larger than strictly needed to encode the relatively modest set of genes in them, but several mechanisms give rise to a very complex transcriptome.

## Introduction

In 1920, the German botanist Hans Winkler wrote: “I propose the expression Genom for the haploid chromosome set, which, together with the pertinent protoplasm, specifies the material foundations of the species ...” (*Verbreitung und Ursache der Parthenogenesis im Pflanzen- und Tierreiche*, Verlag Fischer, Jena). Since then, the word “genome” has evolved in meaning somewhat and gave rise

to words like “genomics”, defining whole new fields of research.

The word “genome” currently has two main meanings, depending on the context:

1. The genome of a species: The corpus of genetic material that characterizes a species, including its most frequent genetic variants. This meaning of “the genetic endowment of a species” is essentially that intended by Winkler’s original definition. For species for which the genome has been sequenced, the “reference genome” denotes the specific sequence that was obtained. Even though the reference genome is frequently a patchwork of sequences derived from several sources, it is taken to be the representative sequence for the species.
2. The genome of an organism or an individual, also known as the “personal genome”. This refers to the specific instance of the genetic material present in most cells of a given organism, stressing the individual’s variations or polymorphisms, and including both parental copies.

A genome is sometimes taken to represent “the set of genes of an organism”. Indeed, some proposed approaches to sequencing personal genomes involve re-sequencing only the exons of known genes. It is known though, that the vast expanses of “junk” DNA within and between genes harbor scattered signals that are important for the regulation of gene expression: A clean-cut subdivision between functional and non-functional DNA is not quite possible. The simplest and most convenient definition of the genome includes in it all the genetic material, regardless of our current state of knowledge about its potential function.

Finally, the word “genome” has also been borrowed to other contexts. For example in Computer Science, when discussing genetic algorithms as a tool for discovering solutions hidden in a complex space that cannot be searched exhaustively, a genome is the set of parameters that define a possible solution to a problem.

The genome has been likened, among other things, to “a blueprint” and “a parts list”. Neither of these is entirely appropriate: A blueprint specifies a detailed construction plan, without reference to the origin of the components, while the parts list simply enumerates the components without reference to their interaction. The genome incorporates both aspects.

The genome is an instruction book, providing all the information needed to maintain a cell’s structural integrity and to perform its functions, to usher it through the complicated ritual of cell division, and to direct its development into more specialized structures. The genome in a cell includes the instructions for building cellular com-

ponents, many of which are in turn used as tools and as raw materials to make other cellular components, in a complex cascade that requires extensive regulation and control. The information required for that control is also encoded in the genome in many different ways. The construction plan (the “blueprint”) is for the most part not explicitly specified in the genome, but rather it emerges from the complex interactions between the cellular components and their spatiotemporal patterns of expression.

### Molecular Structure of the Genome

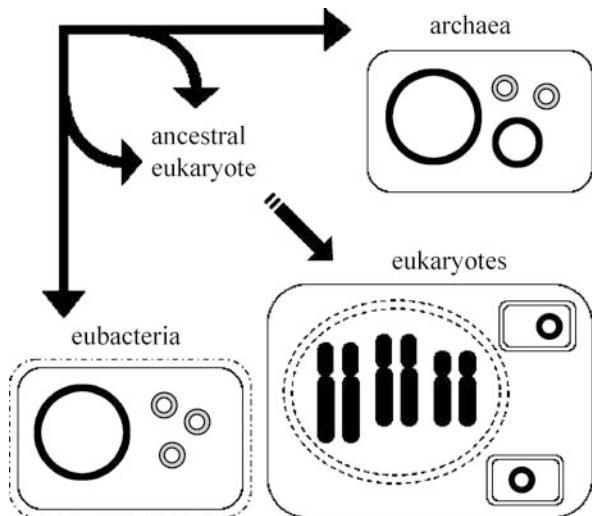
#### Molecular Topology

Nucleic acids are linear polymers with clearly defined polarity, conventionally denoted as running from 5' to 3' end of the molecule. The linearity and polarity severely limit the range of possible molecular topologies: A single-stranded molecule of DNA or RNA can remain linear or form a circle. The specific sequence of nucleotide bases confers variability: Depending on the sequence, local structures can form by base complementarity between different parts of the molecule. A central property of nucleic acids, crucial to the transmission of hereditary information, is the ability to form double-stranded structures, in which two molecules are intertwined and held together by extensive base complementarity. A double-stranded nucleic acid molecule can, again, remain linear or be closed into a circle. Depending on the sequence and the torsional forces induced by various processes, a variety of local structures can emerge and some are used for regulation [69].

Even though a linear molecule would appear to have a simpler topology than a circular structure, the linearity introduces some mechanistic complications. DNA replication requires priming and proceeds linearly along the DNA molecule from 5' to 3'. While one strand of DNA can be replicated continuously, the complementary strand needs to be synthesized in pieces (called Okazaki fragments), which are then ligated. A consequence of this is that the replication of linear DNA molecules requires special structures (called telomeres) and mechanisms to avoid becoming shorter each generation due to the inability to prime the 3' ends. Changes in telomere length have been implicated in disease and in senescence [13].

### Evolution of Genome Structure

When discussing genome organization, it is important to keep in mind that there are three main lineages of living systems: Eukaryotes, bacteria and archaea, the last two collectively called prokaryotes. There are essentially two main



**Genome Organization, Figure 1**

**Evolution of gross genome structure.** Two main lineages, leading to modern archaea and modern eubacteria, merged to produce the ancestral eukaryotic genome. This third lineage adopted a large number of innovations including linear chromosomes, homologous chromosome pairs (diploid genome), sequestration in the nucleus (*double dashed lines*), multicellularity, etc. Eukaryotic cells include double-membraned mitochondria, which are the remnants of the endosymbiotic bacteria from the ancestral genomic merger. Gray circles indicate plasmids (non-essential genetic material) while black represents chromosomes (essential genetic material)

types of chromosomal organization, emphasizing circular and linear chromosomes; these are characteristic of prokaryotes and eukaryotes, respectively (Fig. 1). The ancestral eukaryotic genome arose from repeated fusion of archaeal and bacterial genomes, followed by a large number of structural and functional innovations [52,93,120].

The typical prokaryotic genome is organized into a small number of circular DNA molecules. Some of these (and usually the largest) harbor genes that are essential for the survival of the organism, and are considered to be proper chromosomes: These need to be segregated accurately to ensure proper cell division [33]. Prokaryotic genomes typically include also “plasmids”, which are much smaller and include only non-essential genes. The prokaryotic genome is located in a region of the cell called the nucleoid, which is not separated from the rest of the cytoplasm by a membrane.

In contrast, the typical eukaryotic genome includes several linear chromosomes, and sometimes a large number of them. The eukaryotic chromosomes are sequestered in the cell’s nucleus, a complex structure including a double membrane perforated by pores, which mediate the transport of materials between the nucleus and the cyto-

plasm. During cellular division, the chromosomes adopt a highly condensed “rod” chromatin structure (as depicted in Fig. 1), but during the interphase (between cell divisions) chromosomes expand to fill the entire volume of the nucleus, largely separated into chromosomal territories [15]. The precise organization of the chromosomes within the nucleus has not yet been fully characterized, but there is clear indication of the existence of self-organizing microenvironments within the nucleus [98], which may support the integration of regulatory signals controlling gene expression [132].

In addition to the linear chromosomes in the nucleus, eukaryotic cells contain several copies of very small, circular DNA molecules located in the mitochondria; plants have additional such chromosomes in the chloroplasts. These organellar chromosomes include genes that are essential for the maintenance and function of their respective organelle, and therefore they are essential to the organism. The mitochondrial and chloroplast chromosomes are therefore considered an integral part of the eukaryotic genome, despite their prokaryotic characteristics.

A further, profound difference in genome organization between prokaryotes and eukaryotes relates to the number of copies of the entire set of chromosomes found in the cell. Prokaryotes normally have one copy of the main chromosome per cell, though in fast growing conditions the chromosome may be continually in an intermediate state of replication, yielding a fractional chromosome copy number. In eukaryotes, the entire set of chromosomes (called the haploid set) is frequently doubled (the diploid state) with each chromosome being represented by two homologous copies. Cells in diploid state can divide into two cells retaining the diploid state through the carefully orchestrated process of mitosis, or they can produce haploid cells (which have a single copy of each chromosome) via meiosis. Haploid cells can fuse to generate diploid cells. There is significant variation among species in the life cycle: For example, vertebrate cells spend most of the time in diploid state and generate haploid gametes only for reproduction, while most fungi are normally haploid organisms, and plants switch between both states. Changes in ploidy have profound implications on gene redundancy and on the regulation of gene dosage, and often leads to sterility in animals, while offspring are likely to be isolated from the rest of the population. Polyploidy through whole-genome duplication is therefore a significant disruption requiring many further adaptations, and often leads to speciation with an increase in chromosome number [105]. Polyploidy can arise from mistakes during meiosis, either within one species (autopolyploidy) or through hybridization of two closely related species (allopolyploidy); the latter

may be a rare event, but has the advantage of basically resulting in a diploid with twice the number of chromosomes without the chromosomal segregation issues associated with autoploidy. Relatively recent events are abundant in plants and are easily recognized. Quite a few clear examples exist in animals as well, like the polyploidizations in the ancestors of the salmon family (Salmonidae) and of the *Xenopus* genus of frogs. The only case in mammals is found among the South American octodont rodents, where two of the thirteen species have double the number of chromosomes and double the amount of DNA relative to their sister species.

### Genome Size vs. Organismal Complexity

The first fully sequenced genomes were those of viruses, starting with the 3569 nucleotide long bacteriophage MS2, an RNA virus [35] and the 5386 nucleotide long Phage  $\Phi$ -X174, a DNA virus [102]. While viral genomes are not representative of the genomes of free-living organisms, their organization gave first insights into how compact and “optimized” a genome can be.

The first bacterial genome to be completed was that of *Haemophilus influenzae* [38]. To date, several hundred bacterial genomes have been sequenced, and their full sequences are available through online databases (Table 1). Hundreds of eukaryotic genomes have been sequenced or are currently in draft form, providing an astounding resource for the research community (Table 1). Landmark papers include preliminary analyzes [1,2,19,54,126].

One of the most obvious characteristics of bacterial genomes is the high density of genes, with compact open reading frames that are not interrupted by introns, and that are separated by very little intergenic sequence. The use of DNA by prokaryotes gives the impression of being highly optimized, with 80%–95% of the genome coding for proteins. In extreme cases, the same DNA sequence can be part of more than one ORF, in different frames or strands. Bacterial genes are typically organized into operons, which are sets of consecutive genes on the same strand that are transcribed as a unit.

By extrapolating from the observed characteristics of prokaryotic genomes, it would only be natural to expect the much larger eukaryotic genomes to include many more genes. For a sense of proportion, a prokaryotic genome expanded to the size of a vertebrate genome could encode a few million genes. Such expanded gene numbers would appear to be a simple explanation for the higher organismal complexity observed in eukaryotes. As it turns out, though, vertebrate genomes are much less gene dense than prokaryotic genomes: The coding fraction in eukary-

**Genome Organization, Table 1****Selected internet resources**

Resource description	URL
Integrated Microbial Genomes system at JGI	<a href="http://img.jgi.doe.gov/">http://img.jgi.doe.gov/</a>
TIGR Comprehensive Microbial Resource	<a href="http://cmr.tigr.org/tigr-scripts/CMR/shared/Genomes.cgi">http://cmr.tigr.org/tigr-scripts/CMR/shared/Genomes.cgi</a>
NCBI Entrez Genome Database	<a href="http://www.ncbi.nlm.nih.gov/sites/entrez?db=Genome">http://www.ncbi.nlm.nih.gov/sites/entrez?db=Genome</a>
UCSC Genome Bioinformatics (the Genome Browser)	<a href="http://genome.ucsc.edu/">http://genome.ucsc.edu/</a>
miRBase, a database of miRNA sequences and targets	<a href="http://microrna.sanger.ac.uk">http://microrna.sanger.ac.uk</a>
NHGRI Encyclopedia of DNA Elements (ENCODE)	<a href="http://www.genome.gov/10005107">http://www.genome.gov/10005107</a>
Personal Genome Project	<a href="http://arep.med.harvard.edu/PGP/">http://arep.med.harvard.edu/PGP/</a>

otes decreases with genome size, down to ~1% in humans. In other words, while the human genome is three orders of magnitude larger than that of *E. coli*, it has only one order of magnitude more genes.

There is no simple correlation between genome size and the observed complexity of the organism, and plenty of outliers exist, e.g. the huge genomes of amoebas. Nevertheless, there is a clear ranking from viruses all the way to multicellular eukaryotes, in terms of genome size, gene number, mobile element number, intron number and size, size of intergenic spacer DNA and complexity of regulatory regions [84].

The reduction in coding fraction results from the presence of introns (discussed below) and huge intergenic distances. Two main hypotheses have been raised to explain the growth in genome sizes: The selfish DNA hypothesis, and the bulk DNA hypothesis [48,122]. The selfish DNA hypothesis postulates that the genome size reflects the ecology of the selfish elements multiplying within it. There is a very large number of elements that can multiply within the genome by various mechanisms, giving rise to extensive families of “interspersed repeats” discussed below, and leading to an accumulation of DNA throughout the genome. In contrast, the bulk DNA hypothesis proposes that genome size is actively selected for, since it dictates many cellular features like cell size [23,24], which in turn affects higher level features of the organism [48].

Selection for fast replication [31] and the metabolic cost of DNA [99] have been invoked as possible alternative explanations for the streamlining of microbial genome sizes. Neither appears to be well supported [82], while a deletional bias [86] may be the major force behind variations in bacterial genome size.

### The Genome as Habitat for Transposable Elements

The changes that can befall a genomic sequence are of two kinds. The first can be considered accidental, like mistakes during replication, incomplete repair of envi-

ronmental damage, or sloppy recombination. The second type is due to the selection-driven propagation of transposable elements (TEs). These mutations generally do not take place in a completely random fashion. Many apparent patterns and correlations between genomic features, “genomic complexity” that appears to hint at an underlying organization, are likely due to mutational skews and not to differential adaptive pressures in different regions of the genome. One example may be the so-called isochore structure in land vertebrate (amniote) genomes: The nucleotide composition of large (100 kb-range) genomic regions can range from 30% to 60% G+C nucleotides. This variation correlates with a number of other characteristics, like gene density and intron size, leading initially to suggestions that this nucleotide heterogeneity is functionally maintained. However, it now seems more likely that the variation arises from regional differences in recombination rates; the associated process of gene conversion, in which a short region of one allele is converted to the other, appears to favor resolution of mismatches to G/C nucleotides [32].

Under relatively stable environmental conditions, a genome’s replication, repair and recombination machinery can evolve to reach a balance between stability and flexibility. Given the selective forces on TEs to keep propagating, it is unclear how much control a genome has on the accumulation of mutations of the second type. In fact, the bulk of many eukaryotic genomes is formed by copies of TEs. Being generally without function, these copies have accumulated mutations in a neutral fashion and decayed to various degrees, and are recognized as interspersed repetitive DNA, or “repeats”. The most common repeats are derived from retrotransposons, which reproduce via reverse transcription (RNA to DNA) of their transcripts. These include familiar elements like (endogenous) retroviruses, LINE1 and *Alu* in the human genome. The other common elements are DNA transposons, which move by a cut-and-paste mechanism in which the transposase specifically recognizes the element’s termini. A major conceptual difference between these two groups is that

the proteins in the first mechanism (reverse transcriptase, integrase, etc.) operate in *cis* and those of the second in *trans*: The cytosol-born transposase does not operate on the transcript close at hand but on any DNA copy in the nucleus. Because of this, retrotransposons are able to survive over long periods of evolutionary time in a genome while DNA transposons generally become extinct through the accumulation of defective copies, and depend on frequent horizontal transfer for evolutionary survival [109].

The isolation of the germline in land vertebrates appears to have had a profound effect on the activity of transposable elements in those organisms, by severely limiting the introduction of new elements through horizontal transfer. This is reflected in the nature of interspersed repeats in amniotes (mammals, birds and reptiles) as compared to those in amphibia, fish, and most other organisms: While DNA transposon copies, which are dependent on horizontal transfer for long-term survival, are common in most organisms, they represent only a very small fraction of interspersed repeats in most amniotic genomes. The exceptional mutagenic potential of DNA transposons through the formation of double-strand breaks after excision, may explain the far less dynamic evolution of amniote genomes [18] compared to that of, for example, insects [29] and plants [9].

This is not to say that retrotransposons necessarily are less deleterious for the host genome. Insertions of both DNA transposons and retrotransposons can incapacitate genes, but retrotransposons tend to carry stronger transcriptional regulatory sites. It is likely that the most common form of gene disruption by TEs is through the introduction of transcriptional termination site within introns [46].

From an information science point-of-view, interspersed repetitive DNA decreases the complexity of the genome, as a large fraction can be expressed as derivatives of a limited number of sequences. From a biological point of view, however, interspersed repeats lead to many complications. For one, they provide homologous seeds for “ectopic” recombination between different sites of the genome, resulting in generally deleterious deletions, insertions, and chromosomal crossovers. A less appreciated effect is the concomitant distribution of binding sites for transcription factors, as most TEs carry highly active promoter sites and other transcriptional regulatory sites. Interspersed repeats that have not decayed too far from their original form may account for much spurious transcription in the genome, and even if local transcription is not disturbed, the distribution of thousands of new binding sites for a cellular transcription factor could influence the balance of the regulatory network.

Genomes may not be able to direct the level of activity of transposable elements, but multiple mechanisms have evolved to suppress it. An often exploited, telltale feature of TE activity of any kind is the presence of transcripts in both orientations, as eventually some copies will be co-transcribed with genes in reverse direction. Some fungal genomes inactivate thus detected TEs by directly altering the genomic copies [42]. In placental mammals a cytosine-amidase has evolved into a family of (APOBEC3) proteins that act as an innate immune system inhibiting the replication of retrotransposons and retroviruses by deaminating cytosines (to uracils) in its nascent DNA strands during reverse transcription; the resulting uracil-containing cDNAs are fodder for degradation by cellular uracyl-DNA-glycosylases [34,53]. Much more widespread mechanisms to suppress TE activity involve transcriptional inactivation by methylation or heterochromatization [16], and the post-transcriptional silencing by RNA interference. The Piwi defense system may explain a characteristic pattern in retrotransposon evolution: A retrotransposon seems to be active in bursts, after which the element either goes extinct, or a modified descendant, perhaps unrecognized by RNA interference, takes over. Considering that TEs must have been an overarching problem since the first genome arose, it is not unlikely that methylation, heterochromatization, and RNA interference in general originated as defense mechanisms against parasitic DNA, and only later were adopted for others purposes like cellular gene regulation.

It is clear that, on top of their mutagenic properties and creation of repetitive DNA, TEs have had an enormous impact on genomic organization. Some of the most obvious features of eukaryotic genomes seem to have found their origin in TEs:

- Most ancient introns probably originated in early eukaryotic evolution from mobile elements, possibly resembling present-day prokaryotic group II introns.
- Telomeres are maintained by the enzyme telomerase, which appears to be a reverse transcriptase derived from a retrotransposon [45].
- The major centromere binding protein CENP-B was derived from related transposases, in a process that happened independently in mammals and in fission yeast, while their genomic binding sites (the centromeres) may be derived from their transposon sequences [22,110].
- The genomic rearrangements that underlie the vertebrate adaptive immune response are performed by a domesticated transposase [60].

The cellular adoption of transposable elements proteins, as in the last three examples, is a form of “exaptation” (the

utilization of a structure for a function other than that for which it developed). There are probably over a hundred exapted TE genes, often recombined with cellular genes, in the human genome alone [54]. Non-coding TE sequences have also been exapted, often with many copies at the time [81,89]. An exciting aspect of the adoption of a group of interspersed sequences is that a regulatory network could arise seemingly from scratch. For example, retroviral long terminal repeats carry a battery of transcriptionally regulatory sites, which at opportune locations may become established and can regulate a set of cellular genes in a coordinated fashion [92,124]. Most of the exapted TE peptides in the human genome are DNA binding domains of transposases, which very specifically recognize binding sites at the termini of the dispersed transposon copies. One reason for the frequency of their adoption could be that, after recombination with a cellular transcription factor, of which multiple examples are known [54,109], such proteins and fortunately located transposon copies could create an entirely new regulatory network.

To summarize, the genome is, in a very real sense, the evolutionary environment of a host of selfish mobile genetic elements. The impact of their propagation within the genome is hard to overstate.

## Genomes in Constant Flux

### Gene Duplication

The total number of genes is commonly taken as a measure of genomic complexity, though, as we describe below, the number of different gene products and levels of regulation may be a more relevant measure. Whereas prokaryotic gene number can shift drastically by horizontal transfer, almost all increase in eukaryotic gene number is due to gene duplication: Exceptions are genes acquired from organelle genomes, retroviruses and transposable elements.

The gene content of most eukaryotic genomes is highly dynamic. From a variety of approaches it is estimated that 1 out of 50 genes duplicate every million years in vertebrate genomes. However, since the rate at which new genes are inactivated is far higher (10- to 100-fold) than the rate at which they arise, most duplicated genes are short lived. The total gene number may remain stable for long periods of time, but there is likely to be a continuous turnover with respect to which genes have redundant copies. In the human genome, about 6% of the genes have undergone duplication since our separation from most other mammals, while about 12% of the nucleotides have been substituted. The majority of the resulting gene duplicates have been inactivated.

These numbers exclude the increase in gene number through whole genome duplication (polyploidization); such events in the distant past are referred to as paleopolyploidy. The “2R hypothesis” postulates that two such events happened in early vertebrate evolution, initially almost quadrupling the number of genes evolution could tinker with [63]. Paleopolyploidy has probably been a central mechanism for large-scale genome evolution but is typically hard to prove, mostly due to the inevitable fast decay of many of the redundant genes.

Genes can also duplicate individually or as part of a genomic segment (segmental duplication), rather than along with the whole genome. One type of duplication necessarily restricted to a single gene is retroposition (reverse transcription of the gene’s transcript and reintroduction), which results in an intronless, “processed” gene. This mechanism is common in mammals and over time has given rise to tens of thousands of recognizable copies in the human genome, the great majority of which are inactive processed pseudogenes. This process relies on the LINE1 transposable element, which can retropose any transcript with a poly-A tail. Most other LINE-like transposable elements only interact with transcripts that share the LINE’s terminal sequence and are very unlikely to copy cellular genes, as witnessed by the absence of processed pseudogenes in many organisms. For example, the chicken has none, despite the abundant presence of the LINE-like element CR1.

The low “success rate” of producing functional genes through retroposition is due to the frequent 5’ truncation of the transcript during integration and the inherent inability to copy the transcriptionally regulatory sequences along with the gene. Even when a working copy results, there is only a small chance that it will be functionally conserved, as it is initially redundant. Nevertheless, over a thousand retroposed genes appear to be transcriptionally active in humans, and at least 120 of these retrocopies appear to have evolved into *bona fide* human genes [121].

Transposable elements are central to several other mechanisms of gene duplication. There are three common ways in which a gene or a gene fragment can be co-transposed with a transposable element.

1. Transcription of a retrotransposon can proceed through the poly-A signal into neighboring DNA. Such “transduced” regions are often short, but human functional genes duplicated in this fashion have been reported [129]. Given the usual 5’ truncation of retroposed sequences, the transposable element itself may not be visible at the new site, and the frequency of these events may be underestimated.

2. Rolling circle transposons can similarly transduce 3' flanking sequences, probably by extending the replication process past the original (weakly defined) 3' signal. Elements up to about 20 kb in length have been observed as very common haplotype differences between strains of maize [71]. The longer elements usually contain fragments of multiple cellular genes, indicating frequent but short read-through with each transposition event. In mammals, rolling circle elements were active in a common ancestor over 150 million years ago, while more recent activity has only been reported in several species of bats [95].
3. Genes sandwiched between two copies of the highly specific recognition site of the transposase of a DNA transposon may be duplicated as a unit. This process can be common in genomes with high DNA transposon activity. For example, there are thousands of "Pack-MULES" in the rice genome that contain fragments of cellular genes [56] although, like retroposed genes, a very small fraction of these seem to have resulted in or contributed to functional genes [59].

Most segmental duplications, especially the larger ones that might contain one or more complete genes, do not arise by transposition. Instead, they are the result of recombination events. Even here transposable elements play an indirect but very significant role. The excision of a DNA transposon, and the integration of a retrotransposon, both create temporary double-strand DNA breaks; such breaks frequently trigger recombination. The great increase in both homologous and non-homologous recombination due to DNA transposon activity has been well documented [9]. Furthermore, the interspersed copies of transposable elements are a common cause of misalignment during meiosis, resulting in a deletion in one daughter chromosome, and a (generally tandem) duplication in the other one. One of the first well-characterized gene duplications, that of the  $\gamma$ -globin gene in the ancestor of hominids and old-world monkeys, was mediated by two LINE1 copies flanking the original gene [37]. Comparison of the human and chimp genome revealed that in the short time (6 million years) of separation of these species, around a thousand homologous recombination events between interspersed repeats has deleted about 2 Mb of DNA in each lineage [28,51]. Indeed, it has been postulated that the unusually high number of interspersed segmental duplications in the primate genomes, as compared to that of other sequenced organisms, was due to the extraordinarily high activity of the *Alu* retrotransposon about 30–40 million years ago, which left half a million copies of almost identical 300 bp sequences [5].

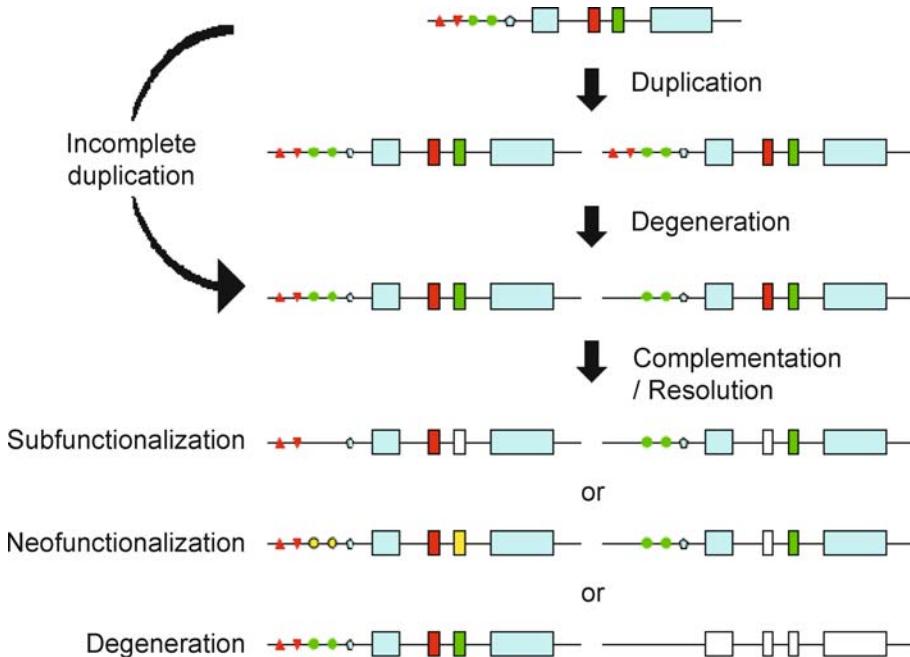
## How Do Duplicated Genes Survive?

The previous section described the many physical mechanisms by which genes may duplicate in a genome, but what forces are there to keep an initially redundant copy of a gene in the genome of a species? We address here two processes by which gene function evolves: Neofunctionalization, and subfunctionalization.

The concept of neofunctionalization is intuitive: Once a new gene copy arises, it is "free" to accumulate mutations and evolve a new, useful function. Given the high rate of accumulation of deleterious mutations, though, it is unlikely that one of the copies will evolve a new function quickly enough. This is especially so in organisms with small population sizes, as the relative strength of adaptive selection over genetic drift is directly related to the population size. Moreover, the presence of a second copy of a gene may even be disadvantageous, as the associated higher level of expression may interfere with proper function. This idea seems corroborated by observations in yeast where genes participating in a protein complex are underrepresented among segmental duplications and members of interacting pairs tend to be co-duplicated [91]. In the case of polyploidy (whole genome duplication), most genes are expressed at an increased level; the loss of individual genes could then lead to imbalances. Indeed, duplicates arising from whole genome duplication have a much longer half-life than those arising from segmental duplications. In vertebrates this difference is about 20 to 25-fold (2 million vs. 45 million years). The longer half-life also reflects another advantage of whole genome duplication, i. e. the preservation of all regulatory sequences, even long distance ones. Considering the ease and frequency with which plants become polyploid, the primary mechanism of chromosomal repatterning in plants may be duplication followed by random loss of component genes.

When a polymorphic gene gives a selective advantage to heterozygotes as compared to the homozygotes of either allele, the gene is said to evolve under balancing selection. Such genes would conceivably be more adept at duplicating, as a genome with both alleles at different loci would have the inherent benefit of "heterozygosity". This mechanism may be at the core of the expansion of gene families greatly benefiting from diversity, like those involved in reproduction, immunity, host defense and chemoreception.

The most feasible route through which duplicated genes may be permanently retained is through the alternative process of subfunctionalization (Fig. 2). When a duplicated gene contains coding or regulating regions that are only important for a subset of its activities (be it in time, space, or function), loss of different functionalities



**Genome Organization, Figure 2**

The Duplication-Degeneration-Complementation (DDC) model for gene duplications. The ancestral gene has two (or more) subfunctions, for each of which independently mutable elements exist. In this particular example, both promoter elements and alternative internal exons exist that are only required for one of two subfunctions. The loss of any one of these regions in one copy of the duplicated gene (which can happen in one or two steps) can be followed by the loss of an independent element in the other copy, after which the copies will be complementary and may both survive. Alternatively though more rarely, the intact copy acquires a new function for the subfunction retained in the damaged copy, after which both copies also may survive. Third, as the damaged copy is still redundant it may simply further decay into a pseudogene. Population genetics predict that the DDC model is both likely, as it depends on common degenerative mutations, and is more apt to occur in small-population species, which tend to have more complex genomes. Figure adapted from Force et al. [39]

in the duplicates can cause the two copies to complement each other, and hence favor the conservation of both in the genome. This Duplication-Degeneration-Complementation (DDC) model [39] is plausible as it invokes common degenerative mutations rather than the exceedingly rare advantageous mutations necessary for neofunctionalization. Subfunctionalization leads to elimination of pleiotropic constraints (i.e. changes cannot interfere with any of the functionalities), so that each duplicate can become optimized to its particular subfunction. This process intuitively adds to increased complexity and flexibility, creating two experts from a single jack-of-all-trades.

An interesting and important prediction is that, through DDC, gene duplications are more likely to become fixated in organisms with small population sizes (which we observe to have more complex genomes), because each step in the process needs to become fixated in the population before the next step occurs, and the time it takes for neutral mutations to reach fixation is propor-

tional to the population size. In addition, the likelihood of this model does not suffer and may even benefit from incomplete gene duplications, which in segmental duplications may be the rule rather than the exception; absence of, for example, a set of upstream regulatory regions combines the duplication and first degeneration step, creating one copy that already misses a subfunction. For genes that arise via transposition, an alternative way in which the transposed copy may lose some of its functionality upon arrival is through differential epigenetic regulation of the locus it inserts into.

The mechanism by which duplications are fixated will depend on the nature of the gene. While the likelihood of duplication through subfunctionalization increases with greater regulatory and functional complexity of a gene, balancing selection-driven neofunctionalization may be prevalent in the expansion of fast evolving genes like those involved in immune response, reproduction, and xenobiotic recognition (e.g. drug detoxification and odorant

recognition). The olfactory receptor gene family offers a striking example of fast expansions and contractions in vertebrates [44].

Other examples of rapidly expanding gene sets can be found in the nearly 6% of the human euchromatin that is derived from recent (< 40 million years old) segmental duplications. While such duplications are overrepresented in pericentromeric and subtelomeric regions, in hominid genomes most are actually interspersed through the genome. These interstitial duplicated units (duplicons) are significantly enriched in genes and gene fragments. Moreover, the gene parts of these duplicons are evolving rapidly (most nucleotide substitutions have affected the amino acid sequences) and are more commonly retained in dupilon copies than other regions, suggesting that positive selection has acted not only to diversify sequences but also to fixate gene duplicates in the population [4]. Since the location of interstitial segmental duplications is strongly correlated with unstable sites, Jiang et al. speculated that the association of a gene with such a site was a “predisposing event”, opening the way for amplification of this gene in different genomic environments [57].

### Transcriptome Complexity

Given the small variability in the number of genes in eukaryotic genomes (and that this number is not that much larger than the respective number in prokaryotes), how does the complex phenotypic diversity of eukaryotes arise? The answer appears to be two-fold: Eukaryotes have the ability of producing many different products from each gene, and have many control mechanisms for regulating and fine-tuning gene expression [83].

The basic mechanism of gene expression involves the recognition, by a set of proteins called the transcriptional complex, of a special DNA sequence called the promoter. The promoter of a gene includes several sequence elements that specify under which conditions the gene should be expressed. The transcriptional complex initiates the copying of the DNA template into an RNA molecule by an enzyme called the RNA polymerase. This process continues until the RNA polymerase encounters a termination signal. The resulting RNA molecule is then further processed, typically leading to the production of a protein molecule.

Prokaryotic genes are easily identified by the presence of an open reading frame (ORF): A contiguous segment of the transcript without any termination codons in (at least) one of the three reading frames in the forward strand. This ORF can be readily translated into a protein sequence, a task carried out *in vivo* by the translation machinery called the ribosome. Sequences outside the ORF are called

untranslated regions (UTRs). A transcript corresponding to a single gene therefore has three well-defined regions: The 5' UTR, the ORF and the 3' UTR, delimited by the initiation codon and the termination codon.

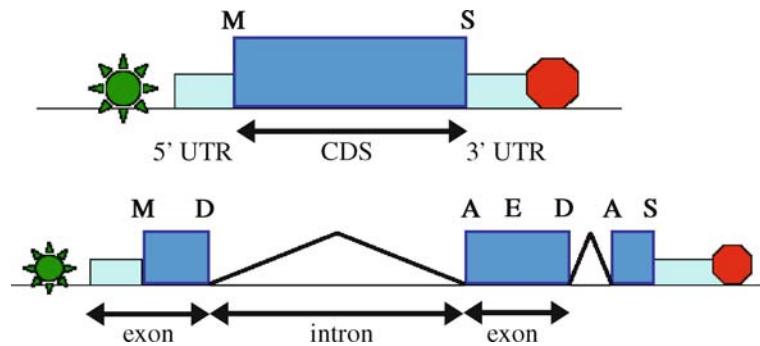
In contrast, three decades ago the surprising discovery was made that eukaryotic genes have discontinuous ORFs [11]. For the vast majority of eukaryotic genes, one or more introns may split the ORF and the 5' UTR into exons (Fig. 3). The boundaries between exons and introns are demarcated by splicing signals: Each intron starts at a “splicing acceptor” and ends at a “splicing donor”.

Most of the information required for splicing resides in the introns, though splicing enhancers have been described, that serve for exon definition. Specialized machinery (called the spliceosome) processes the initial (pre-mRNA, or heterogeneous nuclear RNA, hnRNA) transcripts, removing the introns and splicing the exons together into the mature mRNA. Spliceosomes are complex molecular machines involving five small nuclear RNAs (snRNAs) and hundreds of proteins [119]. A large number of genes encode proteins involved in splicing, amounting to approximately 1% of all human genes. Surprisingly, two different kinds of spliceosomes exist: The major and the minor spliceosomes [107,127], which employ different snRNAs (called U2 and U12, respectively) and have different intron specificities. Based on their wide phylogenetic distribution, it is a safe assumption that both types of introns were present in the “stem eukaryote”, the ancestor of all extant eukaryotic species (Fig. 1). It is not clear how introns originally arose in the genome, though an origin from mobile elements is probable (see above). As with any other elaboration on the gene structure, their presence increases the risk that the host gene will be rendered defective by subsequent mutations [82]. Nevertheless, once established, the splicing machinery conferred significant advantages to the eukaryotic cell, discussed next.

A crucial aspect of pre-mRNA splicing is that it is not deterministic, and it does not always lead to the same final mRNA form. By a variety of pathways, alternative spliced forms may arise. Alternative transcripts can produce protein products with altered cellular localizations, varying substrate affinities, or different activities [133], in many gene families, e. g. transcription factors [118], kinases and phosphatases [40], kallikreins [70], etc.

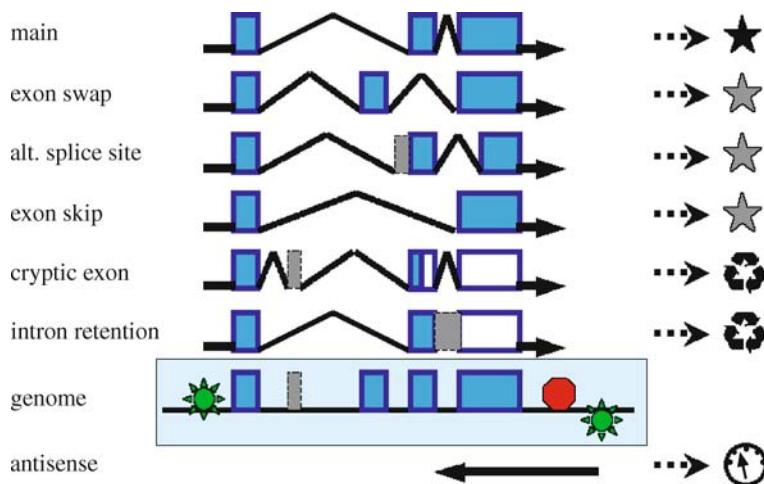
Alternative splicing appears to be the norm, particularly among vertebrates [65]: A genome-wide survey using microarrays showed that over 74% of human multi-exon genes are alternatively spliced [58]. In the plant kingdom, alternative splicing is prevalent in flowering plants [97].

There are several different mechanisms that give rise to alternative splice forms from one promoter (Fig. 4), with



Genome Organization, Figure 3

Schematic of gene structure. **Top:** The basic structure of an intronless gene, with the open reading frame coding region (CDS) flanked by 5' and 3' untranslated regions (UTRs). The star and the stop sign denote the promoter and the poly-adenylation signal, respectively; M and S represent the initiation (methionine) and termination (stop) codons. **Bottom:** The structure of a eukaryotic gene split into exons and introns. D, A and E represent splicing donor, acceptor and enhancer signals, respectively



Genome Organization, Figure 4

The different transcripts arising from a genomic locus. The boxed area represents the genomic sequence. From top to bottom: The main transcript, splicing the main exons and giving rise to the main product (*black star*); a secondary transcript in which one exon was swapped with an alternative exon, giving rise to an alternative product (*gray star*); a secondary transcript employing alternative splice sites, be it by extending into an intron or by losing part of an exon; a secondary transcript skipping an exon; a transcript splicing in a cryptic “cassette” exon; a transcript in which an intron is retained instead of being spliced out; an antisense transcript from a downstream promoter, not coding for a protein but functioning in regulation of the main transcripts (*gauge symbol*). Cryptic exons and intron retention typically (but not always) change the reading frame and give rise to an unusable product (*recycling symbol*). Other graphic elements are as in Fig. 3

very different degrees of prevalence [65]. In exon swapping, an exon is replaced by an alternative exon. Such alternative exons typically arise by local duplication. An exon may have alternative splice sites, delimiting a sequence that is sometimes part of the exon, sometimes spliced out with the neighboring intron. Exon skipping refers to the situation in which an exon is spliced out of the transcript with its flanking introns. Conversely, a “cassette” exon may be included in the transcript. These cryp-

tic exons are typically composed of interspersed repeats and other non-coding intergenic sequence, flanked by sequences similar enough to splice sites to be recognized by the spliceosome [114]. Finally, an intron may be retained, bridging its two flanking exons in the final transcript form (Fig. 4).

In multi-exon genes, more than one exon may be alternatively spliced. This can lead to a combinatorial explosion, potentially giving rise to a very large number of

different transcripts from a single locus. Given its potential to produce very different proteins, or even unusable products [116], it is not surprising that alternative splicing is a highly regulated process, involving many proteins and RNA molecules [111]. Central to this process is the family of the serine- and arginine-rich proteins (SR proteins). SR proteins interact with RNA regulatory sequences, and participate in the spliceosome assembly process [17,96]. Taking advantage of the large amounts of sequence information being produced, as well as *in vivo* splicing reporter systems [125], global analysis approaches are helping elucidate the “cellular code” of splicing [85]. The regulatory sequences include splicing enhancer and silencer elements present in the pre-mRNA, particularly in exons [21,134], but also conserved sequences in introns [113]. There are two main modes of pre-mRNA splicing, involving either the identification of short exons, typically separated by long introns [10,72], or the recognition of short introns [80]. Both modes use the same consensus splice signals [117].

An additional way to achieve a larger variety of alternative transcripts for a gene is by the use of alternative promoters. Many genes are known to have more than one promoter region, each giving rise to its own, unique first exon. These alternative first exons can then be joined with a shared set of exons downstream, adding to the combinatorial capability of transcript production from a single gene locus [79]. Alternatively, the first exon of one gene structure can be an internal exon of another one [26]. Interestingly, transcript elongation and splicing are interrelated in mechanism and regulation, with promoter choice affecting splicing [68]. In turn, the strength of splice sites can affect the choice of transcript termination point, which can lead to the generation of fusion transcripts combining exons from adjacent genes [41].

Several of the mechanisms for alternative splicing produce mature transcripts that include sections of introns, a process called exonization [112]. Such intronic sequences, which were not selected to code for protein sequences, clearly have the potential to interrupt the open reading frame by introducing a premature stop codon. In some cases [77] a proper open reading frame can be restored via a process called RNA editing. The family of adenosine deaminases acting on RNA (ADARs) are able to deaminate adenosines in any long double-stranded RNA, converting them to inosines [64]. Site-specific adenosine deamination can not only correct frame disruptions, but also increase the diversity of transcripts and the resulting proteins [47].

The final quality control of transcripts is effected via a pathway called nonsense-mediated decay (NMD),

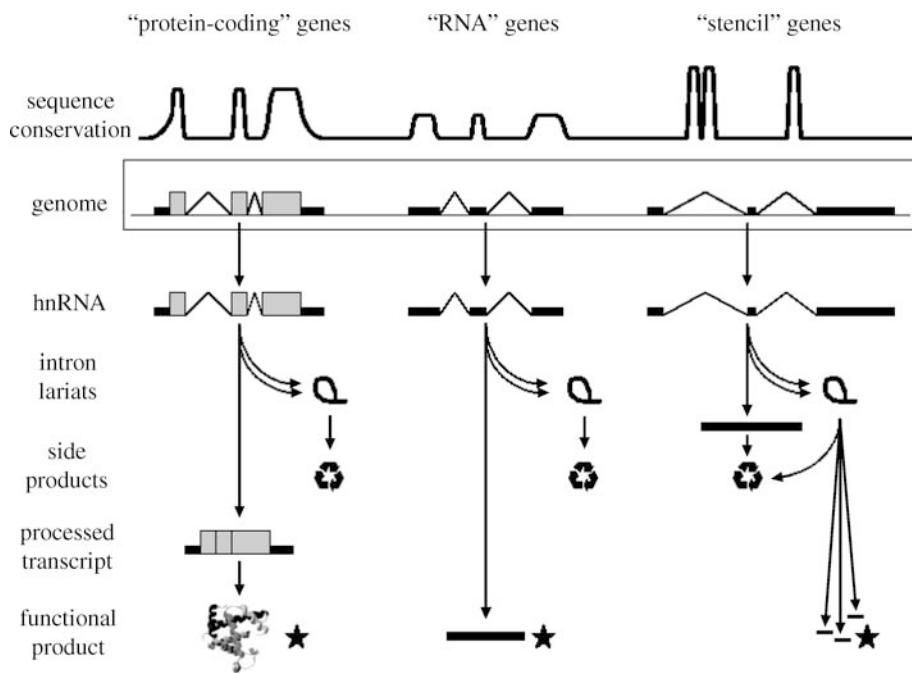
a surveillance system that targets transcripts with premature termination codons [8,103], as well as transcripts with overly extended 3' UTRs [87]. This pathway prevents the production of truncated proteins, which may be detrimental to the cell. Interestingly, NMD is also part of regulatory mechanisms of gene expression [30] particularly of splicing regulators like SR proteins, in which alternatively spliced “poison exons” located in ultraconserved elements introduce unavoidable stop codons [73,88].

Are the products of alternative splicing functional, or are they mostly “noise”? Computational analyses indicate that alternative splicing is regulated in a very tissue-specific manner [130] and suggest that this regulation is strongly modified in cancers [108]. As described above, alternative splicing has profound effects on the structure of the resulting proteome [116]. Taken together, these and other lines of evidence suggest that splicing variants are as a rule functional and actively selected for. In contrast, a statistical comparison of conserved vs. lineage-specific alternatively spliced exons indicates that these are two distinct classes, suggesting that many splice forms observed and cataloged in sequence databases may be aberrant and non-functional [115]. Further evidence from analyses of amino acid frequencies suggest relaxed selection on alternative splice forms [25].

### Transcript Networks

In the previous section we discussed how a variety of different transcripts can arise from a single “gene” and some aspects of how this process is regulated. The classical view of gene function involves the transmission of information from the genome to the protein via a transcript. During the last few years the perception of RNA as a mere information conduit has changed dramatically, and it is now known that this is not the only way an RNA transcript can be functional. RNA molecules can be structural (e.g. ribosomal RNA, transfer RNA) and catalytic, but most importantly, they provide a complex intermediary network of regulation, between DNA-level transcript production and protein-level interaction networks.

The next level of genomic complexity therefore arises from the interaction between transcripts. In fact, we can conceptually subdivide transcripts into three types (Fig. 5), two of which are heavily invested in regulation at the RNA level. Type I describes the classical gene structure in which the mRNA (after splicing out the introns if any were present) serves as the intermediary form leading to the production of the functional product – the protein. In Type II, the (potentially spliced) mRNA transcript itself represents the functional product, without



Genome Organization, Figure 5

Schematic representation of the three types of eukaryotic transcripts, as classified by product type: Type I "protein-coding" genes (left), Type II "RNA" genes (center), and Type III "stencil" genes (right). The boxed area represents the genomic sequence, gray boxes indicate coding exons, thick horizontal lines denote UTRs. In protein-coding genes, the introns are spliced out in lariat form and degraded (recycling symbol), while the processed transcript gives rise to the functional protein product (black star). In RNA genes, the processed transcript is the same as the functional product, without translation to protein. In stencil genes, the processed transcript is meaningless and degraded, while the intron lariats are further processed to produce the functional products – small RNA molecules like miRNAs and snoRNAs. The "sequence conservation" graph reveals the location of the functional elements, extending throughout the exons of protein-coding and RNA genes, but localized in the introns of stencil genes

coding for a protein [50]. Such RNA genes are involved in a large number of cellular processes, including gene regulation [106], and more are being discovered. Type III transcripts differ from the first two types in that, after splicing out of the introns, the resulting spliced mRNA is not the source of the functional product. Instead, specific pathways excise out of the intron lariats short RNA segments including miRNAs (see below) and snoRNAs. We propose to call this third type "stencil genes" because, in similarity to stencils, they are just structural frameworks, while the sections removed from them are those that carry the relevant information. Stencil genes range in size from the very small, e.g. the 5 kb long C6orf48 transcript in the human MHC Class III region [128], to the very large, e.g. the 530 kb long LOC401237 transcript [46]. Interestingly, the three types of transcripts are not mutually exclusive. For example, most snoRNA genes are produced from intron lariats spliced out of transcripts coding for riboproteins [50] and most human miRNA genes are derived from intronic regions [67,101].

The ability of nucleic acids to recognize each other by strand complementarity has afforded several opportunities for the evolution of specific regulatory mechanisms. A large number of genes are now known to be regulated by naturally occurring antisense transcripts [75,131], which can function in *cis* or in *trans* [76]. First identified as specific transcriptional repression by injection of double-stranded RNA in *C. elegans* [36], or RNA interference (RNAi), a more exquisite mechanism for general gene regulation involves the production of short (~ 22 nt long) RNA molecules called micro RNAs, or miRNAs [90] by cleavage from RNA hairpins. miRNAs form complexes with several proteins including Argonaute gene family members, and the resulting miRNP complexes control gene expression by binding to mRNA, frequently in the 3' UTR, recognizing sequence motifs by imperfect sequence complementarity [6]. An mRNA thus recognized by a miRNP complex will be either cleaved and degraded, or its translation to protein will be repressed. The identification of miRNAs and their targets is a very active

area of research, with knowledge being warehoused in the miRBase database (Table 1). Related to miRNAs are small interfering RNAs, or siRNAs, which are produced by cleavage of long double-stranded RNAs. Finally, a novel type of small RNAs was recently described, which interact with Piwi proteins, a subfamily of Argonaute proteins. These Piwi-interacting RNAs, or piRNAs, are expressed mainly in the germline and, as described above, are part of an adaptive defense system against transposable elements [3,66].

Computational analyses suggested that a large fraction of the genome is transcribed [46,104], including many regions previously thought to be devoid of genes (and therefore called “gene deserts”). Experimental analyses demonstrated vast networks of overlapping transcripts on both DNA strands, called “gene forests” [12,20,27,61]. Not surprisingly, some transcripts are only observed in specific cell types and/or conditions [100], suggesting that many additional functional transcript forms are yet to be discovered. The observed complexity of transcription and regulation poses a significant challenge to the classical concept of the gene as the discrete element in the genome, and requires updated definitions [43].

If most of both genomic DNA strands are transcribed in complex loci, with many sequences being multifunctional [12,62], what fraction of the genome is actively selected for functional needs? It has been claimed that everything in the genome is functional [123] to different degrees and in various circumstances, or even in aggregate for “genomic bulk” [135]. On the other hand, it has been proposed that a large fraction of the current genomic features has arisen in a “mostly neutral” fashion [84]. In this context, it is important to keep in mind that the typically small population size of eukaryotic species precludes effective selection over most functional sites, particularly those conferring weak advantageous effects.

## Future Directions

Genomes have evolved over many hundred million years, by a haphazard accumulation of random mutations ranging from the single nucleotide change to the duplication of the entire genome. In prokaryotes, the large population sizes and fast division times afford for very effective selection against deleterious or wasteful mutations, leading to compact and streamlined genomes. In sharp contrast, given the small population sizes of eukaryotic species, the selective advantage of some variants over others has been tempered by the vagaries of random sampling. This ongoing process has led to the accretion of the highly disorganized genomes observed in eukaryotes. One of the

most fundamental and understated results from the analysis of completed genomic sequences is the overwhelming and outstanding lack of evidence for any kind of intelligent design in genome organization (Fig. 6). Even though exceptions exist, in the form of locally coordinated, small clusters of functionally related genes [7], the vast majority of the vertebrate genome is a disordered array of genes in essentially random order and orientation. As genomes evolve, gene order changes following a simple null model of random mutation, with present intergenic distance being the best predictor for future gene linkage [94].

In this context, one major difficulty lies in distinguishing between real biological signals and the background of random variation. We need to understand how the genome is organized and how it evolved, to be able to build suitable null hypotheses for testing whether predictions are likely to be real, and whether specific observations are likely to be biologically meaningful. To surmount this difficulty, the integration of many data types is imperative, as demonstrated by the ENCODE project (Table 1) [12]. Comparative genomics is a powerful tool for distinguishing between evolutionary and functional signal and noise, by identifying what is shared and what is different among related species. In the field of gene prediction, the best results are obtained by integrating several sources of information, reflecting many steps of the process of gene expression, from the genome to the functional product (Fig. 7). Computational and experimental work has shown that many regions of the genome are transcribed in which no classical protein-coding genes can be identified. This suggests that many RNA and stencil genes might remain to be discovered [46].

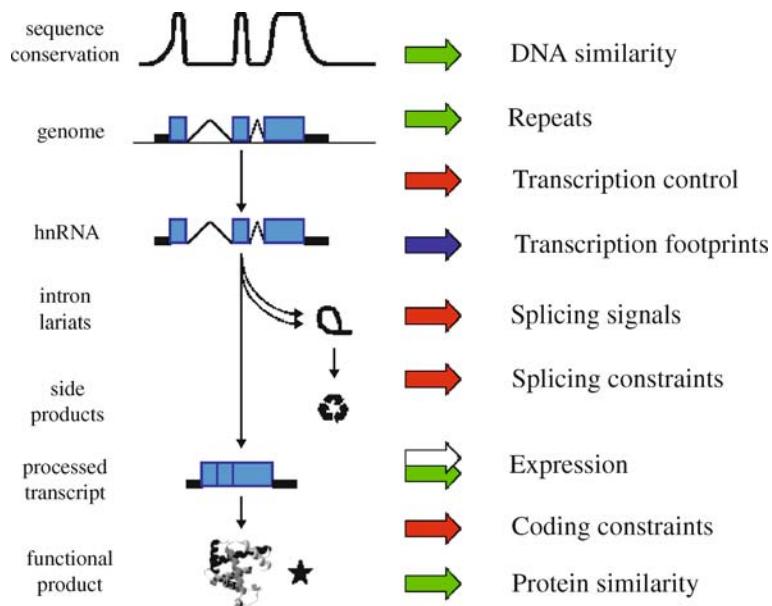
A deep understanding of genome organization and principles of function will enable the new field of genome engineering. Preliminary work has demonstrated the feasibility of combining two bacterial genomes into one [55] and full genome transplantation [74]. Interestingly, the resulting phenotypes observed either reflected the host genome when combined, or the donor genome when replaced. This suggests that researchers will benefit from a wide range of possible phenotypic effects, to be attained by judicious combinations of genomic sources, both natural and synthetic. At a more modest scale, it will be possible to engineer the minimal genetic modifications required to correct the imbalances and deregulations leading to many diseases, including cancer.

The ongoing development of novel sequencing technologies has reduced the cost of sequencing genomes by orders of magnitude. It is expected that within a few years, obtaining an individual’s version of the genome will be extremely affordable, which is one of the essential re-

	non-system	evolved-system	designed-system
<b>purpose</b>	none	survival behavior	purposeful
<b>redundancy</b>	redundant	highly redundant	parsimonious
<b>parts</b>	arbitrary parts	multifunctional parts defunct parts	no extra parts
<b>efficiency</b>	none	wasteful	efficient
<b>response to change</b>	accepts all change	error correction (high turnover) or death	changes outside specification range are catastrophic
<b>examples</b>	mountain river	metazoa bacteria viruses	technology new old
<b>organization</b>	chaotic	messy	organized

Genome Organization, Figure 6

A conceptual comparison between non-systems (random assemblages of objects, e.g. a pile of rocks), living systems (evolved) and technological systems (designed). While designed systems are presumably purposeful and hopefully efficient and parsimonious, non-systems have no purpose – they just are – and as such have no efficiency. The survival behavior of evolved systems gives a functional purpose to their parts, which are typically redundant and multifunctional, or even not used anymore. As such, evolved systems are highly “wasteful”. Designed systems typically have specification ranges within which they function normally, and changes outside those predefined ranges are typically catastrophic. Non-systems accept all change and return to passive equilibrium. Evolved systems typically have many methods of error correction that accumulated over many millions of years of selection in changing environments. Overall, non-systems are chaotic and designed systems are well organized, while evolved systems are very messy and disorganized, typically displaying ancient battle scars, relics or defunct parts, and the minimal level of organization required for survival



Genome Organization, Figure 7

Sources of information used for predicting and recognizing genes in the genome. **Red arrows:** Ab initio gene prediction methods. **Green arrows:** Similarity-based gene prediction methods. **Blue arrow:** Footprint-based transcript prediction. Observed transcripts are used dually to identify genes (white arrow) and to predict similar genes (green arrow). Other graphical elements as in Figs. 3 and 4

quirements for personalized medicine [14]. Already, the elucidation of an individual's genome [78] provided insight into the level of polymorphism arising from different types of mutations, and indicated that alterations involving more than a single base pair represent a sizable fraction of the variability among alleles in the human population. As more personal genome sequences become available through the Personal Genome Project (Table 1) and other efforts, it will be possible to obtain a detailed characterization of mutation frequencies, on which to base proper models of neutral evolution. Based on the cumulative results of population sequencing, it will be possible to derive robust "reference" genomes for human populations, leading to a highly detailed characterization of the history of human evolution. The genetic analysis of complex diseases will be strongly impacted by the cumulative genomic knowledge, leading to a comprehensive identification of the genes associated with most major diseases, and a detailed characterization of their interactions.

## Bibliography

### Primary Literature

1. Adams MD, Celtniker SE, Holt RA, Evans CA, Gocayne JD et al (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287:2185–2195
2. Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM et al (2002) Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* 297:1301–1310
3. Aravin AA, Hannon GJ, Brennecke J (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 318:761–764
4. Bailey JA, Eichler EE (2006) Primate segmental duplications: Crucibles of evolution, diversity and disease. *Nat Rev Genet* 7:552–564
5. Bailey JA, Liu G, Eichler EE (2003) An Alu transposition model for the origin and expansion of human segmental duplications. *Am J Hum Genet* 73:823–834
6. Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116:281–297
7. Batada NN, Hurst LD (2007) Evolution of chromosome organization driven by selection for reduced gene expression noise. *Nat Genet* 39:945–949
8. Behm-Ansmant I, Kashima I, Rehwinkel J, Sauliere J, Wittkopp N et al (2007) mRNA quality control: An ancient machinery recognizes and degrades mRNAs with nonsense codons. *FEBS Lett* 581:2845–2853
9. Bennetzen JL (2007) Patterns in grass genome evolution. *Curr Opin Plant Biol* 10:176–181
10. Berget SM (1995) Exon recognition in vertebrate splicing. *J Biol Chem* 270:2411–2414
11. Berget SM, Moore C, Sharp PA (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci USA* 74:3171–3175
12. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR et al (2007) Identification and analysis of functional ele-  
ments in 1% of the human genome by the ENCODE pilot project. *Nature* 447:799–816
13. Blasco MA (2007) Telomere length, stem cells and aging. *Nat Chem Biol* 3:640–649
14. Blow N (2007) Genomics: The personal side of genomics. *Nature* 449:627–630
15. Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K et al (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol* 3:e157
16. Bourc'his D, Bestor T (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 431:96–99
17. Bourgeois CF, Lejeune F, Stevenin J (2004) Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. *Prog Nucleic Acid Res Mol Biol* 78:37–88
18. Bourque G, Zdobnov EM, Bork P, Pevzner PA, Tesler G (2005) Comparative architectures of mammalian and chicken genomes reveal highly variable rates of genomic rearrangements across different lineages. *Genome Res* 15:98–110
19. *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* 282:2012–2018
20. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC et al (2005) The transcriptional landscape of the mammalian genome. *Science* 309:1559–1563
21. Cartegni L, Chew SL, Krainer AR (2002) Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nat Rev Genet* 3:285–298
22. Casola C, Hucks D, Feschotte C (2008) Convergent domestication of pogo-Like Transposases into centromere-binding proteins in fission yeast and mammals. *Mol Biol Evol* 25(1):29–41
23. Cavalier-Smith T (1978) Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNAC-value paradox. *J Cell Sci* 34:247–278
24. Cavalier-Smith T, Beaton MJ (1999) The skeletal function of non-genic nuclear DNA: New evidence from ancient cell chimaeras. *Genetica* 106:3–13
25. Chen FC, Wang SS, Chen CJ, Li WH, Chuang TJ (2006) Alternatively and constitutively spliced exons are subject to different evolutionary forces. *Mol Biol Evol* 23:675–682
26. Chen WH, Lv G, Lv C, Zeng C, Hu S (2007) Systematic analysis of alternative first exons in plant genomes. *BMC Plant Biol* 7:55
27. Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S et al (2005) Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* 308:1149–1154
28. Chimpanzee Genome Sequencing Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437:69–87
29. Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B et al (2007) Evolution of genes and genomes on the drosophila phylogeny. *Nature* 450:203–218
30. Conti E, Izaurralde E (2005) Nonsense-mediated mRNA decay: Molecular insights and mechanistic variations across species. *Curr Opin Cell Biol* 17:316–325
31. Doolittle WF, Sapienza C (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284:601–603

32. Duret L, Eyre-Walker A, Galtier N (2006) A new perspective on isochore evolution. *Gene* 385:71–74
33. Egan ES, Fogel MA, Waldor MK (2005) Divided genomes: Negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol Microbiol* 56:1129–1138
34. Esnault C, Heidmann O, Delebecque F, Dewannieux M, Ribet D et al (2005) APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. *Nature* 433:430–433
35. Fiers W, Contreras R, Duerinck F, Haegeman G, Iserentant D et al (1976) Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene. *Nature* 260:500–507
36. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE et al (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806–811
37. Fitch DH, Bailey WJ, Tagle DA, Goodman M, Sieu L et al (1991) Duplication of the gamma-globin gene mediated by L1 long interspersed repetitive elements in an early ancestor of simian primates. *Proc Natl Acad Sci USA* 88:7396–7400
38. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF et al (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512
39. Force A, Lynch M, Pickett FB, Amores A, Yan YL et al (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151:1531–1545
40. Forrest AR, Taylor DF, Crowe ML, Chalk AM, Waddell NJ et al (2006) Genome-wide review of transcriptional complexity in mouse protein kinases and phosphatases. *Genome Biol* 7:R5
41. Frith MC, Carninci P, Kai C, Kawaji J, Bailey TL et al (2007) Splicing bypasses 3' end formation signals to allow complex gene architectures. *Gene* 403:188–193
42. Galagan J, Selker E (2004) RIP: The evolutionary cost of genome defense. *Trends Genet* 20:417–423
43. Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J et al (2007) What is a gene, post-ENCODE? History and updated definition. *Genome Res* 17:669–681
44. Gilad Y, Man O, Glusman G (2005) A comparison of the human and chimpanzee olfactory receptor gene repertoires. *Genome Res* 15:224–230
45. Gladyshev EA, Arkhipova IR (2007) Telomere-associated endonuclease-deficient Penelope-like retroelements in diverse eukaryotes. *Proc Natl Acad Sci USA* 104:9352–9357
46. Glusman G, Qin S, El-Gewely MR, Siegel AF, Roach JC et al (2006) A third approach to gene prediction suggests thousands of additional human transcribed regions. *PLoS Comput Biol* 2:e18
47. Gott JM (2003) Expanding genome capacity via RNA editing. *C R Biol* 326:901–908
48. Gregory TR (2001) Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biol Rev Camb Philos Soc* 76:65–101
49. Gregory TR (2002) Genome size and developmental complexity. *Genetica* 115:131–146
50. Griffiths-Jones S (2007) Annotating noncoding RNA genes. *Annu Rev Genomics Hum Genet* 8:279–298
51. Han K, Lee J, Meyer TJ, Wang J, Sen SK et al (2007) Alu recombination-mediated structural deletions in the chimpanzee genome. *PLoS Genet* 3:1939–1949
52. Hedges SB, Chen H, Kumar S, Wang DY, Thompson AS et al (2001) A genomic timescale for the origin of eukaryotes. *BMC Evol Biol* 1:4
53. Holmes R, Malim M, Bishop K (2007) APOBEC-mediated viral restriction: Not simply editing? *Trends Biochem Sci* 32:118–128
54. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921
55. Itaya M, Tsuge K, Koizumi M, Fujita K (2005) Combining two genomes in one cell: Stable cloning of the *Synechocystis* PCC6803 genome in the *bacillus subtilis* 168 genome. *Proc Natl Acad Sci USA* 102:15971–15976
56. Jiang N, Bao Z, Zhang X, Eddy SR, Wessler SR (2004) Pack-MULE transposable elements mediate gene evolution in plants. *Nature* 431:569–573
57. Jiang Z, Tang H, Ventura M, Cardone MF, Marques-Bonet T et al (2007) Ancestral reconstruction of segmental duplications reveals punctuated cores of human genome evolution. *Nat Genet* 39:1361–1368
58. Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM et al (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* 302:2141–2144
59. Juretic N, Hoen DR, Huynh ML, Harrison PM, Bureau TE (2005) The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. *Genome Res* 15:1292–1297
60. Kapitonov VV, Jurka J (2005) RAG1 core and V(D)J recombination signal sequences were derived from transib transposons. *PLoS Biol* 3:e181
61. Kapranov P, Drenkow J, Cheng J, Long J, Helt G et al (2005) Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. *Genome Res* 15:987–997
62. Kapranov P, Willingham AT, Gingeras TR (2007) Genome-wide transcription and the implications for genomic organization. *Nat Rev Genet* 8:413–423
63. Kasahara M (2007) The 2R hypothesis: An update. *Curr Opin Immunol* 19:547–552
64. Keegan LP, Leroy A, Sproul D, O'Connell MA (2004) Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes. *Genome Biol* 5:209
65. Kim E, Magen A, Ast G (2007) Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res* 35:125–131
66. Kim VN (2006) Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. *Genes Dev* 20:1993–197
67. Kim YK, Kim VN (2007) Processing of intronic microRNAs. *Embo J* 26:775–783
68. Kornblith AR (2005) Promoter usage and alternative splicing. *Curr Opin Cell Biol* 17:262–268
69. Kouzine F, Levens D (2007) Supercoil-driven DNA structures regulate genetic transactions. *Front Biosci* 12:4409–4423
70. Kurlender L, Borgono C, Michael IP, Obiezu C, Elliott MB et al (2005) A survey of alternative transcripts of human tissue kallikrein genes. *Biochim Biophys Acta* 1755:1–14
71. Lai J, Li Y, Messing J, Dooner HK (2005) Gene movement by Helitron transposons contributes to the haplotype variability of maize. *Proc Natl Acad Sci USA* 102:9068–9073
72. Lam BJ, Hertel KJ (2002) A general role for splicing enhancers in exon definition. *RNA* 8:1233–1241
73. Larea LF, Inada M, Green RE, Wengrod JC, Brenner SE (2007) Unproductive splicing of SR genes associated with

- highly conserved and ultraconserved DNA elements. *Nature* 446:926–929
74. Lartigue C, Glass JI, Alperovich N, Pieper R, Parmar PP et al (2007) Genome transplantation in bacteria: Changing one species to another. *Science* 317:632–638
  75. Lavorgna G, Dahary D, Lehner B, Sorek R, Sanderson CM et al (2004) In search of antisense. *Trends Biochem Sci* 29:88–94
  76. Lehner B, Williams G, Campbell RD, Sanderson CM (2002) Antisense transcripts in the human genome. *Trends Genet* 18:63–65
  77. Lev-Maor G, Sorek R, Levanon EY, Paz N, Eisenberg E et al (2007) RNA-editing-mediated exon evolution. *Genome Biol* 8:R29
  78. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL et al (2007) The diploid genome sequence of an individual human. *PLoS Biol* 5:e254
  79. Li Q, Lee JA, Black DL (2007) Neuronal regulation of alternative pre-mRNA splicing. *Nat Rev Neurosci* 8:819–831
  80. Lim LP, Burge CB (2001) A computational analysis of sequence features involved in recognition of short introns. *Proc Natl Acad Sci USA* 98:11193–11198
  81. Lowe CB, Bejerano G, Haussler D (2007) Thousands of human mobile element fragments undergo strong purifying selection near developmental genes. *Proc Natl Acad Sci USA* 104:8005–8010
  82. Lynch M (2006) Streamlining and simplification of microbial genome architecture. *Annu Rev Microbiol* 60:327–349
  83. Lynch M (2006) The origins of eukaryotic gene structure. *Mol Biol Evol* 23:450–468
  84. Lynch M, Conery JS (2003) The origins of genome complexity. *Science* 302:1401–1404
  85. Matlin AJ, Clark F, Smith CW (2005) Understanding alternative splicing: Towards a cellular code. *Nat Rev Mol Cell Biol* 6:386–398
  86. Mira A, Ochman H, Moran NA (2001) Deletional bias and the evolution of bacterial genomes. *Trends Genet* 17:589–596
  87. Muhlrad D, Parker R (1999) Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. *RNA* 5:1299–1307
  88. Ni JZ, Grate L, Donohue JP, Preston C, Nobida N et al (2007) Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev* 21:708–718
  89. Nishihara H, Smit AF, Okada N (2006) Functional noncoding sequences derived from SINEs in the mammalian genome. *Genome Res* 16:864–874
  90. Ouellet DL, Perron MP, Gobeil LA, Plante P, Provost P (2006) MicroRNAs in gene regulation: When the smallest governs it all. *J Biomed Biotechnol* 2006:69616
  91. Papp B, Pál C, Hurst LD (2003) Dosage sensitivity and the evolution of gene families in yeast. *Nature* 424:194–197
  92. Peaston AE, Esvikov AV, Gruber JH, de Vries WN, Holbrook AE et al (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* 7:597–606
  93. Pisani D, Cotton JA, McLnerney JO (2007) Supertrees disentangle the chimerical origin of eukaryotic genomes. *Mol Biol Evol* 24:1752–1760
  94. Poyatos JF, Hurst LD (2007) The determinants of gene order conservation in yeasts. *Genome Biol* 8:R233
  95. Pritham EJ, Feschotte C (2007) Massive amplification of rolling-circle transposons in the lineage of the bat *Myotis lucifugus*. *Proc Natl Acad Sci USA* 104:1895–1900
  96. Reddy AS (2004) Plant serine/arginine-rich proteins and their role in pre-mRNA splicing. *Trends Plant Sci* 9:541–547
  97. Reddy AS (2007) Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu Rev Plant Biol* 58:267–294
  98. Rippe K (2007) Dynamic organization of the cell nucleus. *Curr Opin Genet Dev* 17:373–380
  99. Rogozin IB, Makarova KS, Natale DA, Spiridonov AN, Tatusov RL et al (2002) Congruent evolution of different classes of non-coding DNA in prokaryotic genomes. *Nucleic Acids Res* 30:4264–4271
  100. Rozowsky J, Wu J, Lian Z, Nagalakshmi U, Korbel JO et al (2006) Novel transcribed regions in the human genome. *Cold Spring Harb Symp Quant Biol* 71:111–116
  101. Ruby JG, Jan CH, Bartel DP (2007) Intronic microRNA precursors that bypass Drosha processing. *Nature* 448:83–86
  102. Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR et al (1977) Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* 265:687–695
  103. Schell T, Kulozik AE, Hentze MW (2002) Integration of splicing, transport and translation to achieve mRNA quality control by the nonsense-mediated decay pathway. *Genome Biol* 3:REVIEWS1006
  104. Sémon M, Duret L (2004) Evidence that functional transcription units cover at least half of the human genome. *Trends Genet* 20:229–232
  105. Sémon M, Wolfe K (2007) Consequences of genome duplication. *Curr Opin Genet Dev* 17:505–512
  106. Shamovsky I, Nudler E (2006) Gene control by large noncoding RNAs. *Sci STKE* 2006:pe40
  107. Sharp PA, Burge CB (1997) Classification of introns: U2-type or U12-type. *Cell* 91:875–879
  108. Skotheim RI, Nees M (2007) Alternative splicing in cancer: Noise, functional, or systematic? *Int J Biochem Cell Biol* 39:1432–1449
  109. Smit AF (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Dev* 9:657–663
  110. Smit AF, Riggs AD (1996) Tiggers and DNA transposon fossils in the human genome. *Proc Natl Acad Sci USA* 93:1443–1448
  111. Soller M (2006) Pre-messenger RNA processing and its regulation: A genomic perspective. *Cell Mol Life Sci* 63:796–819
  112. Sorek R (2007) The birth of new exons: Mechanisms and evolutionary consequences. *RNA* 13:1603–1608
  113. Sorek R, Ast G (2003) Intronic sequences flanking alternatively spliced exons are conserved between human and mouse. *Genome Res* 13:1631–1637
  114. Sorek R, Lev-Maor G, Reznik M, Dagan T, Belinky F et al (2004) Minimal conditions for exonization of intronic sequences: 5' splice site formation in alu exons. *Mol Cell* 14:221–231
  115. Sorek R, Shamir R, Ast G (2004) How prevalent is functional alternative splicing in the human genome? *Trends Genet* 20:68–71
  116. Stamm S, Ben-Ari S, Rafalska I, Tang Y, Zhang Z et al (2005) Function of alternative splicing. *Gene* 344:1–20
  117. Sterner DA, Carlo T, Berget SM (1996) Architectural limits on split genes. *Proc Natl Acad Sci USA* 93:15081–15085
  118. Taneri B, Snyder B, Novoradovsky A, Gaasterland T (2004) Alternative splicing of mouse transcription factors affects

- their DNA-binding domain architecture and is tissue specific. *Genome Biol* 5:R75
119. Valadkhan S (2005) snRNAs as the catalysts of pre-mRNA splicing. *Curr Opin Chem Biol* 9:603–608
120. Vellai T, Vida G (1999) The origin of eukaryotes: The difference between prokaryotic and eukaryotic cells. *Proc Biol Sci* 266:1571–1577
121. Vinckenbosch N, Dupanloup I, Kaessmann H (2006) Evolutionary fate of retroposed gene copies in the human genome. *Proc Natl Acad Sci USA* 103:3220–3225
122. Vinogradov AE (2004) Evolution of genome size: Multilevel selection, mutation bias or dynamical chaos? *Curr Opin Genet Dev* 14:620–626
123. von Sternberg R (2002) On the roles of repetitive DNA elements in the context of a unified genomic-epigenetic system. *Ann N Y Acad Sci* 981:154–188
124. Wang T, Zeng J, Lowe CB, Sellers RG, Salama SR et al (2007) Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proc Natl Acad Sci USA* 104:18613–18618
125. Wang Z, Rolish ME, Yeo G, Tung V, Mawson M et al (2004) Systematic identification and analysis of exonic splicing silencers. *Cell* 119:831–845
126. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF et al (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520–562
127. Will CL, Luhrmann R (2005) Splicing of a rare class of introns by the U12-dependent spliceosome. *Biol Chem* 386:713–724
128. Xie T, Rowen L, Aguado B, Ahearn ME, Madan A et al (2003) Analysis of the gene-dense major histocompatibility complex class III region and its comparison to mouse. *Genome Res* 13:2621–2636
129. Xing J, Wang H, Belancio VP, Cordaux R, Deininger PL et al (2006) Emergence of primate genes by retrotransposon-mediated sequence transduction. *Proc Natl Acad Sci USA* 103:17608–17613
130. Xu Q, Modrek B, Lee C (2002) Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res* 30:3754–3766
131. Yelin R, Dahary D, Sorek R, Levanon EY, Goldstein O et al (2003) Widespread occurrence of antisense transcription in the human genome. *Nat Biotechnol* 21:379–386
132. Zaidi SK, Young DW, Javed A, Pratap J, Montecino M et al (2007) Nuclear microenvironments in biological control and cancer. *Nat Rev Cancer* 7:454–463
133. Zavolan M, Kondo S, Schonbach C, Adachi J, Hume DA et al (2003) Impact of alternative initiation, splicing, and termination on the diversity of the mRNA transcripts encoded by the mouse transcriptome. *Genome Res* 13:1290–1300
134. Zheng ZM (2004) Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *J Biomed Sci* 11:278–294
135. Zuckerkandl E (2002) Why so many noncoding nucleotides? The eukaryote genome as an epigenetic machine. *Genetica* 115:105–129
- Hartwell L et al (2006) *Genetics: From Genes to Genomes*. McGraw-Hill, Boston
- Lynch M (2007) *The Origins of Genome Architecture*. Sinauer, Sunderland
- Margulis L, Sagan D (2003) *Acquiring Genomes*. Basic Books, New York
- Watson JD et al (2003) *Molecular Biology of the Gene*. Cummings

## Geo-complexity and Earthquake Prediction

VLADIMIR KEILIS-BOROK<sup>1,2</sup>, ANDREI GABRIELOV<sup>3</sup>, ALEXANDRE SOLOVIEV<sup>2,4</sup>

<sup>1</sup> Institute of Geophysics and Planetary Physics and Department of Earth and Space Sciences, University of California, Los Angeles, USA

<sup>2</sup> International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Moscow, Russia

<sup>3</sup> Departments of Mathematics and Earth and Atmospheric Sciences, Purdue University, West Lafayette, USA

<sup>4</sup> The Abdus Salam International Center for Theoretical Physics, Trieste, Italy

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Lithosphere as a Hierarchical Complex System](#)

[General Scheme of Prediction](#)

[Four Paradigms](#)

[Earthquake Prediction and Earthquake Preparedness](#)

[Further Goals](#)

[Acknowledgment](#)

[Bibliography](#)

### Glossary

**Chaos** Apparently random or unpredictable behavior in systems governed by deterministic laws. The common element in these systems is a very high sensitivity to initial conditions and to the way in which a system is set in motion (Encyclopedia Britannica).

**Complexity** An attribute of nonlinear (chaotic) systems. It comprises instability and complex but not random behavior patterns – “order in chaos”.

**Earthquake** An episode of rupture and discontinuous displacement within the solid Earth. Part of the energy accumulated around the rupture is released by inelas-

### Books and Reviews

Brown TA (2002) *Genomes*. Bios Scientific Publishers, Oxford

Dawkins R (1982) *The Extended Phenotype*. Oxford University Press, Oxford

tic deformation and seismic waves. Both may cause destructive shaking of the ground, if the energy release is sufficiently large.

**Earthquake forecasting** Probabilistic extrapolation of seismic activity comprising many earthquakes.

**Earthquake prediction** Prediction of time interval, geographic area, and magnitude range where an individual future strong earthquake will occur. The prediction is meaningful if it includes an estimated rate of false alarms.

**Earthquake preparedness** A set of actions reducing the damage from the future earthquakes. There are different levels of preparedness.

**Extreme events** Rare events of low probability but high impact on a system where they occur. In different connotations they are also known as critical transitions, disasters, catastrophes, and crises. Over time they persistently recur in both natural and constructed complex systems. In this article the extreme events are the strong earthquakes. An earthquake might be an extreme event in a certain volume of the lithosphere and part of the background seismicity in a larger volume.

**Lithosphere** The earthquake-prone outer shell of the solid Earth. In prediction research it is regarded as a hierarchical complex system.

**Premonitory seismicity patterns** Space-time-magnitude patterns of earthquake occurrences that signal the approach of a strong earthquake.

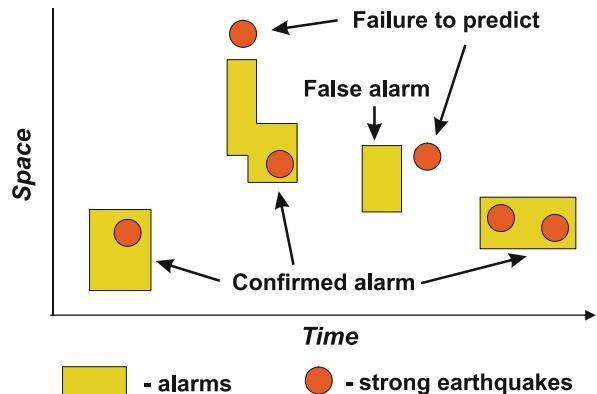
## Definition of the Subject

### Definition

The problem of earthquake prediction is to find when and where a strong earthquake will occur. A prediction is formulated as a discrete sequence of alarms (Fig. 1). The accuracy of a prediction method is captured by probabilities of errors (false alarms and failures to predict) and by the total space-time occupied by alarms. (Sect. “[Error Diagram](#)”).

In terms of prediction studies this is algorithmic prediction of individual extreme events having low probability but large impact. This problem is necessarily intertwined with problems of disaster preparedness, dynamics of solid Earth, and modeling of extreme events in hierarchical complex systems.

*Predictability (“order in chaos”).* Complex systems, lithosphere included, are not predictable with unlimited precision. However, after a coarse-graining (i. e., in a not-too-detailed scale) certain regular behavior patterns emerge and a system becomes predictable, up to certain limits ([[13](#),[20](#),[24](#),[26](#),[36](#),[46](#),[52](#),[83](#)]). Accordingly, earthquake prediction requires a holistic analysis, “from the



**Geo-complexity and Earthquake Prediction, Figure 1**  
Possible outcomes of prediction

whole to details”. Such analysis makes it possible to overcome the geo-complexity itself and the chronic imperfection of observations as well.

*Premonitory patterns.* Certain behavior patterns emerge more frequently as a strong earthquake draws near. Called premonitory patterns, they signal destabilization of the earthquake-prone lithosphere and thus an increase in the probability of a strong earthquake. Premonitory patterns do not necessarily contribute to causing a subsequent strong earthquake; both might be parallel manifestations of the same underlying process – the tectonic development of the Earth in multiple time-, space-, and energy- scales. For that reason premonitory patterns might emerge in a broad variety of observable fields reflecting lithosphere dynamics, and in different scales.

The algorithms considered here, based on premonitory seismicity patterns, provide alarms lasting years to months. There is ample evidence that major findings made in developing these algorithms are applicable to premonitory patterns in other fields, to predicting other geological and geotechnical disasters, and probably to determining shorter and longer alarms (Sect. “[Further Goals](#)”).

### Importance

Algorithmic earthquake prediction provides pivotal constraints for fundamental understanding of the dynamics of the lithosphere and other complex systems. It is also critically important for protecting the global population, economy, and environment. Vulnerability of our world to the earthquakes is rapidly growing, due to proliferation of high-risk construction (nuclear power plants, high dams, radioactive waste disposals, lifelines, etc.), deterioration of ground and infrastructure in megacities, destabilization

of environment, population growth, and escalating socio-economic volatility of the global village. Today a single earthquake with its ripple effects may take up to a million lives; destroy a megacity; trigger a global economic depression (e.g. if it occurs in Tokyo); trigger an ecological catastrophe, rendering a large territory uninhabitable; or destabilize military balance in a region. Regions of low seismicity have become highly vulnerable, e.g. European and Indian platforms, and Central and Eastern parts of the U.S. As a result the earthquakes joined the ranks of the major disasters that, in the words of J. Wisner, have become "a threat to civilization survival, as great as was ever posed by Hitler, Stalin or the atom bomb". Earthquake prediction is necessary to reduce the damage by escalating disaster preparedness. Predictions useful for preparedness should have known, but not necessarily high, accuracy. Such is the standard practice in preparedness for all disasters, wars included.

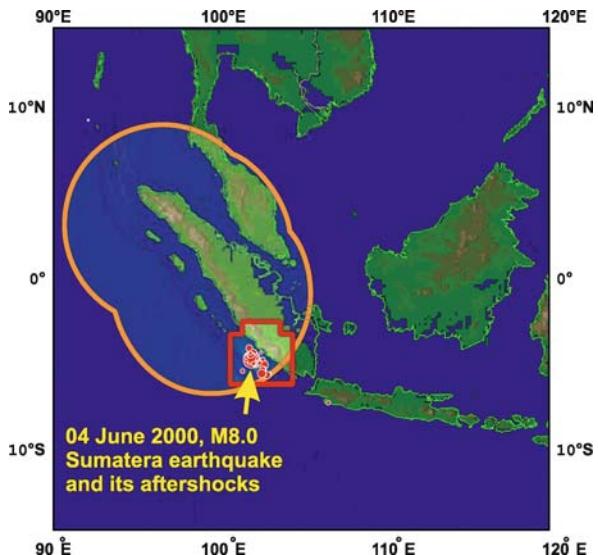
## Introduction

Earthquakes occur in some parts of the outer shell of the solid Earth, called the lithosphere; its thickness ranges from a few kilometers near the mid-ocean ridges to a few hundred kilometers in certain continental regions. At many continental margins the lithosphere bends downward penetrating underlying mantle as seismically active subduction zones. In seismically active regions a significant part of tectonic development is realized through the earthquakes.

About a million earthquakes with magnitude 2 (energy about  $10^{15}$  erg) or more are detected each year worldwide by seismological networks. About a hundred of these cause considerable damage and few times in a decade a catastrophic earthquake occurs.

Catalogs of earthquakes provide the data for detecting premonitory seismicity patterns. Typically for complexity studies we do not have a complete set of fundamental equations that govern dynamics of seismicity and unambiguously define earthquake prediction algorithms. This is due to the multitude of mechanisms controlling seismicity – see Sect. “Generalization: Complexity and Extreme Events”. In lieu of such equations “...we have to rely upon the hypotheses obtained by processing of the experimental data” (A. Kolmogorov on transition to turbulence). Formulating and testing such hypotheses involves exploratory data analysis, numerical and laboratory modeling, and theoretical studies (Sect. “General Scheme of Prediction”).

Diversity of methods and urgency of the problem makes learning by doing a major if not the major form of



**Geo-complexity and Earthquake Prediction, Figure 2**

Prediction of the Sumatra earthquake, June 4th, 2000,  $M = 8.0$  by algorithms M8 and MSc. The orange oval curve bounds the area of alarm determined by algorithm M8, the red rectangle is its reducing made by algorithm MSc. Circles show epicenters of the Sumatra earthquake and its aftershocks. After [43]

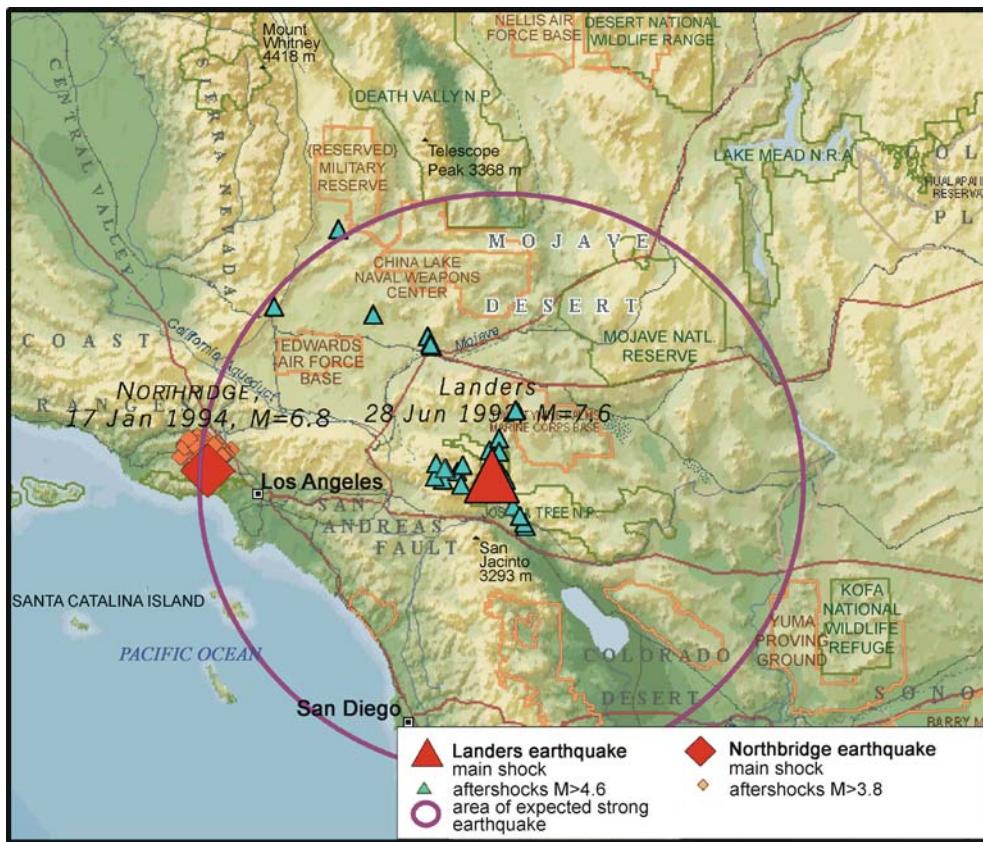
knowledge transfer in prediction of extreme events ([http://cdsagenda5.ictp.it/full\\_display.php?da=a06219](http://cdsagenda5.ictp.it/full_display.php?da=a06219)).

*Reliability of the existing algorithms* has been tested by continuous prediction of future strong earthquakes in numerous regions worldwide. Each algorithm is self-adapting, i.e. applicable without any changes in the regions with different seismic regimes. Predictions are filed in advance at the websites (<http://www.mitp.ru/predictions.html>; <http://www.phys.ualberta.ca/mirrors/mitp/predictions.html>; and <http://www.igpp.ucla.edu/prediction/rtp/>).

Following is the scoring for four different algorithms.

- Algorithms M8 [32] and MSc [44] (MSc stands for the Mendocino Scenario). Algorithm M8 gives alarms with characteristic duration years. MSc gives a second approximation to M8, reducing the area of alarm. An example of their application is shown in Fig. 2.

Continually applied since 1992, algorithm M8 has predicted 10 out of 14 large earthquakes (magnitude 8 or more) which have occurred in the major seismic belts. Alarms occupied altogether about 30% of the time-space considered. Both algorithms applied together reduced the time-space alarms to 15%, but three more target earthquakes were missed by prediction.



Geo-complexity and Earthquake Prediction, Figure 3

Prediction of the Northridge, California earthquake, January 28th, 1994,  $M = 6.8$  by algorithm SSE. The prediction was made by analysis of aftershocks of the Landers earthquake, June 28th, 1992,  $M = 7.6$ . An earthquake with  $M = 6.6$  or larger was expected during the 18 months after the Landers earthquake within the 169-km distance from its epicenter (shown by a circle). The Northridge earthquake occurred on January 28th, 1994, 20 days after the alarm expired. After [43]

- Algorithm SSE or Second Strong Earthquake [43,91]. Its aim is to predict whether or not a second strong earthquake will follow the one that had just occurred. An alarm lasts 18 months after the first strong earthquake. An example of prediction is shown in Fig. 3. Testing by prediction in advance is set up for California, Pamir and Tien Shan, Caucasus, Iberia and Maghreb, the Dead Sea rift, and Italy. Since 1989 this algorithm made 29 predictions; 24 of which were correct and 5 were wrong.

These scores demonstrate predictability of individual earthquakes. A predictions' accuracy is indeed limited, but sufficient to prevent a considerable part of the damage.

- Algorithm RTP or Reverse Tracing of Precursors [37,81]. This algorithm gives alarms with a characteristic duration of months. An example of this prediction is shown in Fig. 4. Testing by prediction in advance started only

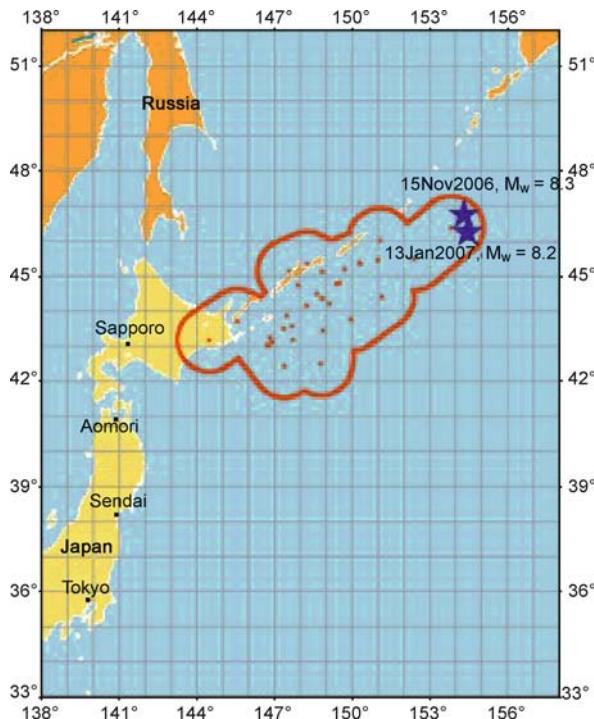
few years ago for California, Japan, the Northern Pacific, Eastern Mediterranean, and Italy with adjacent areas.

*Perspective.* It is encouraging that only a small part of readily available relevant data, models and theories have been used for prediction so far. This suggests a potential for a substantial increase of prediction accuracy.

### Lithosphere as a Hierarchical Complex System

Two major factors turn the lithosphere into a hierarchical dissipative complex system [29,36,87]. The first one is a *hierarchical structure* extending from tectonic plates to grains of rocks. The second factor is *instability* caused by a multitude of nonlinear mechanisms destabilizing the strength and stress fields.

Among extreme events in that system are the strong earthquakes. *An earthquake may be an extreme event in*



**Geo-complexity and Earthquake Prediction, Figure 4**

Prediction of Simushir, Kuril Islands earthquakes, November 15th, 2006,  $M_w = 8.3$  and January 13th, 2007,  $M_w = 8.2$  by Algorithm RTP. An earthquake with magnitude  $M_w \geq 7.2$  is predicted to occur within the time interval from September 30th, 2006, to June 30th, 2007 in the area bordered by the red curve. The red dots show epicenters of an earthquake-forming premonitory chain. The blue stars show epicenters of the predicted earthquakes

a certain volume of the lithosphere and a part of the background seismicity in a larger volume.

## Structure

**Blocks** The structure of the lithosphere presents a hierarchy of volumes, or blocks, which move relative to each other. The largest blocks are the major tectonic plates, of continental size. They are divided into smaller blocks, such as shields or mountain belts. After 15–20 consecutive divisions we come to about  $10^{25}$  grains of rocks of millimeter size.

**Boundary zones** Blocks are separated by relatively thin and less rigid boundary zones. They are called fault zones high in the hierarchy, then faults, sliding surfaces, and, finally, interfaces between grains of rock. Except at the bottom of the hierarchy, a boundary zone presents a similar hierarchical structure with more dense division. Some seg-

ments of the boundary zones, particularly in tectonically young regions, might be less explicitly expressed, presenting a bundle of small ruptures not yet merged into a fault, of a flexure not yet ruptured, etc.

**Nodes** These are even more densely fractured mosaic structures formed around the intersections and junctions of boundary zones. Their origin is due, roughly saying, to collision of the corners of blocks [16,39,40,55]. The nodes play a singular role in the dynamics of the lithosphere. A special type of instability is concentrated within the nodes and strong earthquakes nucleate in nodes. The epicenters of strong earthquakes worldwide are located only within some specific nodes that can be identified by pattern recognition [19,22].

Nodes are well known in the structural geology and geomorphology and play a prominent textbook role in geological prospecting. However their connection with earthquakes is less widely recognized.

The formalized procedure for dividing a territory into blocks  $\Rightarrow$  faults  $\Rightarrow$  nodes is given in [2].

## Fault Network – A Stockpile of Instability

For brevity, the systems of boundary zones and nodes are called here fault networks. They range from the Circum Pacific seismic belt, with the giant triple junctions for the nodes, to interfaces between the grains of rocks, with the corners of grains for the nodes. Their great diversity notwithstanding, fault networks play a similar role in the lithosphere dynamics. Specifically, while tectonic energy is stored in the whole volume of the lithosphere and well beneath, the energy release is to a large extent controlled by the processes in relatively thin fault networks. This contrast is due to the following.

First, the strength of a fault network is smaller than the strength of blocks it separates: fault networks are weakened by denser fragmentation and higher permeability to fluids. For that reason, tectonic deformations are concentrated in fault networks, whereas blocks move essentially as a whole, with a relatively smaller rate of internal deformations. In other words, in the time scale directly relevant to earthquake prediction (hundreds of years or less) the major part of the lithosphere dynamics is realized through deformation of fault networks and relative movement of blocks.

Second, the strength of a fault network is not only smaller, but also highly unstable, sensitive to many processes there. There are two different kinds of such instability. The “physical” one is originated at the micro level by a multitude of physical and chemical mechanisms re-

viewed in the next section. “Geometric” instability is originated at a macro level controlled by the geometry of the fault network (Sect. “[Geometric Instability](#)”). These instabilities largely control dynamics of seismicity, the occurrence of strong earthquakes included.

### “Physical” Instability [23,29]

As in any solid body, deformations and fracturing in the lithosphere are controlled by the relation of the strength field and stress field. The strength is in turn controlled by a great multitude of interdependent mechanisms concentrated in the fault network. We describe, for illustration, several such mechanisms starting with the impact of fluids.

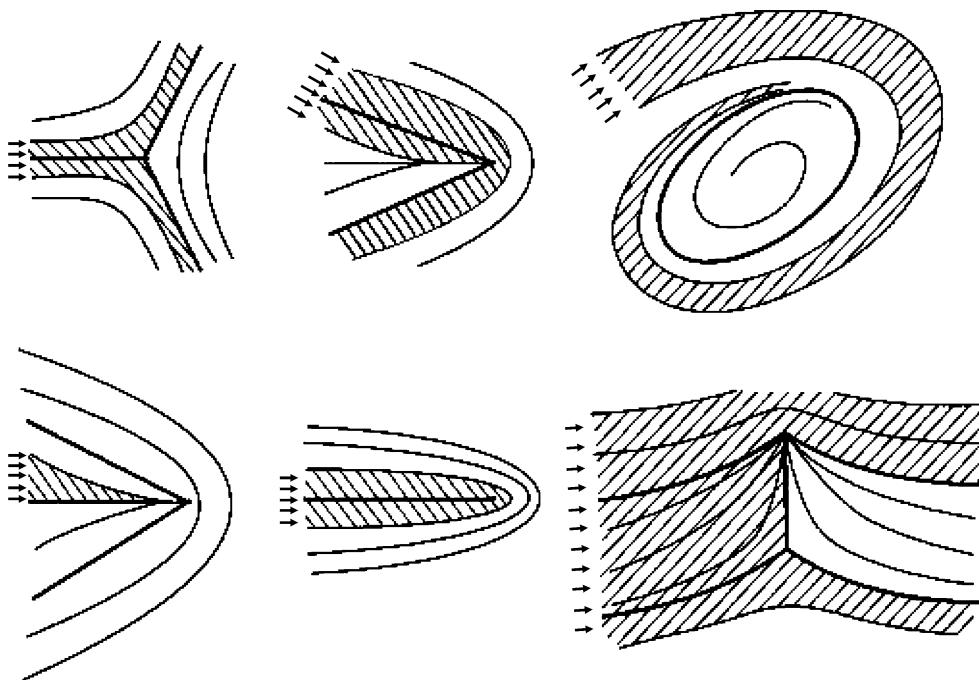
### Rehbinder Effect, or Stress Corrosion [14,85]

**Mechanism** Many solid substances lose their strength when they come in contact with certain surface-active liquids. The liquid diminishes the surface tension  $\mu$  and consequently the strength, which is proportional to  $\sqrt{\mu}$  by the Griffiths criterion. When the strength drops, cracks may emerge under small stress. Then liquid penetrates the cracks and they grow, with drops of liquid propelling for-

ward, until they dissipate. This greatly reduces the stress required to generate the fracturing. Stress corrosion was first discovered for metals and ceramics. Then such combinations of solid substances and surface-active liquids were recognized among the common ingredients of the lithosphere, e. g. basalt and sulphur solutions. When they meet, the basalt is permeated by a grid of cracks and the efficient strength may instantly drop by a factor of 10 or more due to this mechanism alone.

**Geometry of Weakened Areas** Orientation of such cracks at each point is normal to the main tensile stress. The stress field in the lithosphere may be very diverse. However, the shape of weakened areas where the cracks concentrate may be of only a few types, determined by the theory of singularities. Some examples are shown in Fig. 5, where thin lines show the trajectories of cracks; each heavy line is a separatrix, dividing the areas with different patterns of trajectories.

If a liquid infiltrates from a place shown in Fig. 5 by arrows, the cracks concentrate in the shaded area, and its strength plummets. A slight displacement of the source across the separatrix may strongly change the geometry of such fatigue; it may be diverted to quite a different place and take quite a different shape, although not an arbitrary



**Geo-complexity and Earthquake Prediction, Figure 5**

Instability caused by stress corrosion. The geometry of weakened areas depends on the type of singularity and the place where the chemically active fluid comes in. After [14]

one. Furthermore evolution of the stress field may change the type of a singularity, make it disappear or create a new one, and the geometry of fatigue will follow suit.

**Stress Corrosion is Highly Sensitive to Geochemistry of Fluids** For example, gabbro and dolerite are affected only in the presence of iron oxides; Kamchatka ultrabasic rocks are affected by the andesite lava liquids only in the presence of copper oxide, etc. Migration of fluids would cause observable variations of electromagnetic and geochemical fields.

**Summing Up** Stress corrosion brings into lithosphere a strong and specific instability, which may explain many observed premonitory seismicity patterns. However the basic configurations of fatigue, as shown in Fig. 5 might be realizable only in not-too-large areas. This limitation stems from the dissipation of fluids and/or from the inhomogeneity of stress field.

**Other Mechanisms** Boundary zones feature several other mechanisms, potentially as important and certainly as complicated. A few more examples follow.

**Mechanical Lubrication** by fluids migrating through a boundary zone [7]. The ensuing instability will be enhanced by *fingers of fluids* springing out at the front of migration [6].

**Dissolution of Rocks** Its impact is magnified by the *Rikke effect* – an increase of solubility of rocks with pressure. This effect leads to a mass transfer. Solid material is dissolved under high stress and carried out in solution along the stress gradient to areas of lower stress, where it precipitates. The Rikke effect might be easily triggered in a crystalline massif at the corners of rock grains, where stress is likely to concentrate.

**Petrochemical Transitions** Some of them tie up or release fluids, as in the formation or decomposition of serpentines. Other transitions cause a rapid drop of density, such as in the transformation of calcite into aragonite. (This would create a vacuum and unlock the fault; the vacuum will be closed at once by hydrostatic pressure, but a rupture may be triggered.).

Instability is created also by sensitivity of dynamic friction to local physical environment [50], mechanical processes, such as multiple fracturing, buckling, viscous flow, and numerous other mechanisms [49,70].

Most of the above mechanisms are sensitive to variations of pressure and temperature.

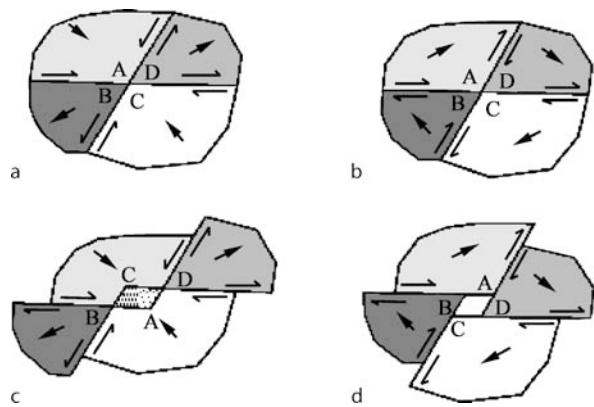
### Geometric Instability [16]

The geometry of fault networks might be, and often is, incompatible with kinematics of tectonic movements, including earthquakes. This leads to stress accumulation, deformation, fracturing, and the change of fault geometry, jointly destabilizing the fault network. Two integral measures of this instability, both concentrated in the nodes, are *geometric* and *kinematic incompatibility* [16].

Each measure estimates the integrated effect of tectonic movements in a wide range of time scales, from seismicity to geodetic movements to neotectonics.

**Geometric Incompatibility** The intersection of two strike-slip faults separating moving blocks. Figure 6 is a simple example of geometric incompatibility. If the movements indicated by arrows in Fig. 6a could occur, the corners A and C would penetrate each other and an intersection point would split into a parallelogram (Fig. 6c). In the general case of a finite number of intersecting faults their intersection point would split into a polygon. Such splitting is not possible in reality; the collision at the corners leads to the accumulation of stress and deformations near the intersection followed by fracturing and changes of fault geometry. The divergence of the corners will be realized by normal faulting.

The expansion of that unrealizable polygon with time,  $S(t) = Gt^2/2$ , measures the intensity of this process. Here,  $S$  is the area of the polygon, determined by the slip rates on intersecting faults;  $t$  is the elapsed time from the collision, and  $G$  is the measure of geometric incompatibility.



**Geo-complexity and Earthquake Prediction, Figure 6**

Geometric incompatibility near a single intersection of faults. a, b initial position of the blocks; c, d extrapolation of the blocks' movement; a, c the locked node: movement is physically unrealizable without fracturing or a change in the fault geometry; b, d the unlocked node. After [16]

Such incompatibility of structure and kinematics was first described in [55] for a triple junction. The study established a condition under which a single junction can retain its geometry as the plates move, so that the stress will not accumulate. It was suggested in [39,40] that the general case, when that condition is not satisfied, the ensuing fracturing would not dissolve the stress accumulation, but only redistribute it among newly formed corners. This triggers further similar fracturing with the result that a hierarchy of progressively smaller and smaller faults is formed about an initial intersection. This is a node, recognizable by the dense mosaic structure, with probably self-similar fractal geometry [39].

A real fault network contains many interacting nodes. Incompatibility  $G$  is additive, and can be estimated for a network as a whole. An analogue of the Stokes theorem connects the total value of  $G$  within a territory with observations on its boundary. This removes the nearly impossible task – to take into account complex internal structure of the nodes. One can instead surround the system of nodes by a contour crossing the less complicated areas. Then the geometric incompatibility can be realistically evaluated from the movements of the fewer faults that cross the contour.

Geometric incompatibility in different nodes is interdependent, because they are connected through the movements of blocks-and-faults system. A strong earthquake in a node would redistribute values  $G$  in other nodes thus affecting the occurrence of earthquakes there. Observations indicating the interaction of nodes have been described by [73,74]. These studies demonstrate phenomenon of long-range aftershocks: a rise of seismic activity in the area, where the next strong earthquake is going to occur within about 10 years.

So far, the theory of geometric incompatibility has been developed for the two-dimensional case, with rigid blocks and horizontal movements.

**Kinematic Incompatibility** Relative movements on the faults would be in equilibrium with the absolute movements of blocks separated by these faults (one could be realized through the other) under the well known Saint-Venant condition of kinematic compatibility [8,56,57]. In the simplest case, shown in Fig. 6, this condition is  $K = \sum v_i = 0$ , where  $v_i$  are slip rates on the faults meeting at the intersection (thin arrows in Fig. 6). The value of  $K$  is the measure of the kinematic incompatibility, causing accumulation of stress and deformation in the blocks. A simple illustration of that phenomenon is the movement of a rectangular block between two pairs of parallel faults. The movement of the block as a whole has to be com-

pensated for by relative movements on all the faults surrounding it: if, for example, the movement takes place on only one fault, the stress will accumulate at other faults and within the block itself thus creating kinematic incompatibility.

Like geometric incompatibility the values of  $K$  are also additive: one may sum up values at different parts of the network. And an analogue of the Stokes theorem links the value of  $K$  for a region with observations on its boundary.

### Generalization: Complexity and Extreme Events

Summing up, dynamics of the lithosphere is controlled by a wide variety of mutually dependent mechanisms concentrated predominantly within fault networks and interacting across and along the hierarchy. Each mechanism creates strong instability of the strength-stress field, particularly of the strength. Except for very special circumstances, none of these mechanisms alone prevails in the sense that the others can be neglected.

Even the primary element of the lithosphere, a grain of rock, may act simultaneously as a material point, a viscoelastic body, an aggregate of crystals, a source or absorber of energy, fluids, volume, with its body and surface involved in different processes.

Assembling the set of governing equations is unrealistic and may be misleading as well: A well-known maxim in nonlinear dynamics tells that *one cannot understand chaotic system by breaking it apart* [12]. One may rather hope for a generalized theory (or at least a model), which directly represents the gross integrated behavior of the lithosphere. That brings us to the concept that *the mechanisms destabilizing the strength of fault networks altogether turn the lithosphere into a nonlinear hierarchical dissipative system, with strong earthquakes among the extreme events*. At the emergence of that concept the lithosphere was called a chaotic system [29,66,87]; the more general term is *complex system* [20,24,31,53,78,83].

### General Scheme of Prediction

Typically for a complex system, the solid Earth exhibits a permanent background activity, a mixture of interacting processes providing the raw data for earthquake prediction. Predictions considered here are based on detecting premonitory patterns of that activity (Sect. “[Definition](#)”).

### Pattern Recognition Approach

Algorithms described here consider prediction as the pattern recognition problem: *Given* the dynamics of relevant fields in a certain area prior to some time  $t$ , *to predict*

whether a strong earthquake will or will not occur within that area during the subsequent time interval  $(t, t + \Delta)$ . Some algorithms also reduce the area where it will occur.

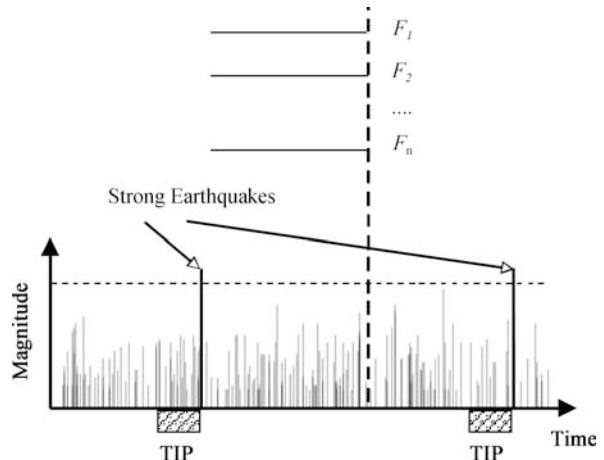
In terms of pattern recognition, the object of recognition is the time  $t$ . The problem is to recognize whether it belongs or not to the time interval  $\Delta$  preceding a strong earthquake. That interval is often called the *TIP* (an acronym for the *time of increased probability* of a strong earthquake). Such prediction is aimed not at the whole dynamics of seismicity but only at the rare extraordinary phenomena, strong earthquakes.

*Pattern recognition of rare events* proves to be very efficient in that approach to prediction. This methodology has been developed by the school of I. Gelfand for the study of rare phenomena of complex origin [9,19,34,71].

## Data Analysis

Prediction algorithms are designed by analysis of the learning material – a sample of past critical events and the time series hypothetically containing premonitory patterns. Analysis comprises four following steps:

1. *Detecting premonitory patterns.* Each time series considered is robustly described by the functionals  $F_k(t)$ ,  $k = 1, 2, \dots$ , capturing hypothetical patterns (Fig. 7). Hypotheses on what these patterns may be are provided by universal modeling of complex systems (Sect. “Fourth Paradigm: Dual Nature of Premonitory Phenomena”), modeling of Earth-specific processes, exploratory data analysis, and practical experience, even if it is intuitive. Pattern recognition of rare events is an efficient common framework for formulating and testing such hypotheses, their diversity notwithstanding.  
With a few exceptions the functionals are defined in sliding time windows; the value of a functional is attributed to the end of the window. In the algorithms described here the time series were earthquake sequences.
2. *Discretization.* Emergence of a premonitory pattern is defined by the condition  $F_k(t) \geq C_k$ . The threshold  $C_k$  is chosen in such a way that a premonitory pattern emerges on one side of the threshold more frequently than on another side. That threshold is usually defined as a certain percentile of the functional  $F_k$ . In such robust representation of the data pattern recognition is akin to exploratory data analysis developed in [86].
3. *Formulating an algorithm.* A prediction algorithm will trigger an alarm when a certain combination of premonitory patterns emerges. This combination is determined by further application of pattern recognition procedures [36,71].



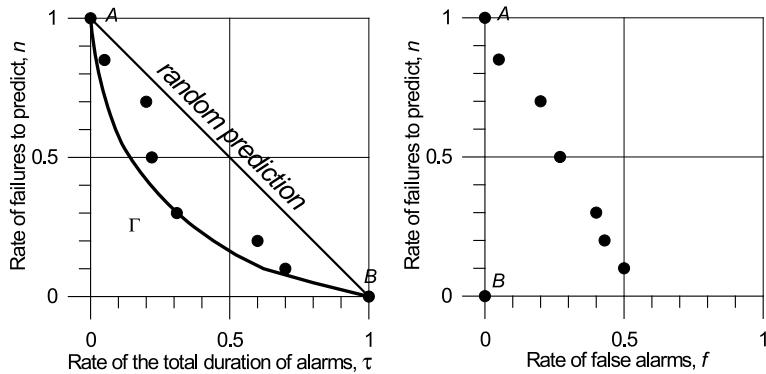
**Geo-complexity and Earthquake Prediction, Figure 7**  
General scheme of prediction. After [29]

4. *Estimating reliability of an algorithm.* This is necessary, since an algorithm inevitably includes many adjustable elements, from selecting the data used for prediction and definition of prediction targets, to the values of numerical parameters. In lieu of the closed theory a priori determining all these elements they have to be adjusted retrospectively, by predicting the past extreme events. That creates the danger of self-deceptive data-fitting: *If you torture the data long enough, it will confess to anything.* Validation of the algorithms requires three consecutive tests.
  - *Sensitivity analysis:* varying adjustable elements of an algorithm.
  - *Out of sample analysis:* applying an algorithm to past data that has not been used in the algorithm’s development.
  - *Predicting in advance* – the only decisive test of a prediction algorithm.

Such tests take a lion’s share of data analysis [17,19,36,93]. A prediction algorithm makes sense only if its performance is (i) sufficiently better than a random guess, and (ii) not too sensitive to variation of adjustable elements. Error diagrams described in the next section show whether these conditions are satisfied.

## Error Diagram

**Definition** An error diagram shows three major characteristics of a prediction’s accuracy. Consider an algorithm applied to a certain territory during the time period  $T$ . During the test  $N$  strong earthquakes have occurred there and  $N_m$  of them have been missed by alarms. Altogether,



**Geo-complexity and Earthquake Prediction, Figure 8**

Scheme of an error diagram. Each point shows the performance of a prediction method: the rate of failures to predict,  $n$ , the relative duration of alarms,  $\tau$ , and the rate of false alarms,  $f$ . Different points correspond to different algorithms. The diagonal in the left plot corresponds to the random guess. Point A corresponds to the trivial optimistic strategy, when an alarm is never declared; point B marks the trivial pessimistic strategy, when an alarm takes place all the time; other points correspond to non-trivial predictions. Best combinations ( $n, \tau$ ) lie on the envelope of these points  $\Gamma$ . After [63]

$A$  alarms have been declared and  $A_f$  of them happened to be false. The total duration of alarms is  $D$ .

Performance of an algorithm is characterized by three dimensionless parameters: the relative duration of alarms,  $\tau = D/T$ ; the rate of failures to predict,  $n = N_m/N$ ; and the rate of false alarms,  $f = A_f/A$ . These three parameters are necessary in any test of a prediction algorithm regardless of a particular methodology. They are juxtaposed on the error diagrams schematically illustrated in Fig. 8. Also called Molchan diagrams, they are used for validation and optimization of prediction algorithms and for joint optimization of prediction and preparedness [59,60,61,62,63]. In many applications parameter  $f$  is not yet considered. In early applications they are called ROC diagrams for relative operating characteristics (e.g., [54]).

## Four Paradigms

Central for determining premonitory patterns is what we know about them a priori. In other words – what are a priori constraints on the functionals  $F_k(t)$  that would capture these patterns (Sect. “Data Analysis”). These constraints are given by the four paradigms described in this section. They have been first found in the quest for premonitory seismicity patterns in the observed and modeled seismicity. There are compelling reasons to apply them also in a wide variety of prediction problems.

*Prehistory.* New fundamental understanding of the earthquake prediction problem was formed during the last 50 or so years, triggering entirely new lines of research. In

hindsight this understanding stems from the following unrelated developments in the early sixties.

- F. Press initiated the installation of the state-of-the-art World-Wide Standardized Seismographic Network (WWSSN) later on succeeded by the Global Seismographic Network (GSN). Thus a uniform data base began to accumulate, augmented by expanding satellite observations.
- E. Lorenz discovered deterministic chaos in an ordinary natural process, thermal convection in the atmosphere [51]. This triggered recognition of deterministic chaos in a multitude of natural and socio-economic processes; however, the turn of seismicity and geodynamics in general came about 30 years later [4,29,66,87]. The phenomenon of deterministic chaos was eventually generalized by less rigorously defined and more widely applicable concept of complexity [20, 24,25].
- I. Gelfand and J. Tukey, working independently, created a new culture of exploratory data analysis that allows coping with the complexity of a process (e.g., [19,86]).
- R. Burridge and L. Knopoff [11] demonstrated that a simple system of interacting elements may reproduce a realistically complex seismicity, fitting many basic heuristic constraints. The models of interacting elements developed in statistical physics extended to seismology.
- L. Malinovskaya found a premonitory seismicity pattern reflecting the rise of seismic activity [33]. This is

the first reported earthquake precursor formally defined and featuring long-range correlations and worldwide similarity.

With broader authorship:

- Plate tectonics established the connection between seismicity and large-scale dynamics of the lithosphere [41].
- Research in experimental mineralogy and rocks mechanics revealed a multitude of mechanisms that may destabilize the strength in the fault zones [70].

### First Paradigm: Basic Types of Premonitory Patterns

*The approach of a strong earthquake is indicated by the following premonitory changes in the basic characteristics of seismicity:*

- *Rising*: Seismic activity, earthquakes clustering in space-time, earthquake correlation range, and irregularity of earthquake sequences. Rise of activity sometimes alternates with seismic quiescence.
- *Transforming*: Magnitude distribution (the Gutenberg–Richter relation). Its right end (at larger magnitudes) bends upward, and left end bends downward.
- *Reversing*: territorial distribution of seismicity.
- *Patterns of two more kinds* yet less explored: Rising response to excitation and decreasing dimensionality of the process considered (i. e. rising correlation between its components).

These patterns resemble asymptotic behavior of a thermodynamical system near the critical point in phase transition. Some patterns have been found first in observations and then in models; other patterns have been found in the opposite order. More specifics are given in [15,17,30,31,35,36,67,79,80,83,84,93].

Patterns capturing rise of intensity and clustering, have been validated by statistically significant predictions of real earthquakes [43,65]; other patterns undergo different stages of testing.

### Second Paradigm: Long-Range Correlations

*The generation of an earthquake is not localized about its future source. A flow of earthquakes is generated by a fault network, rather than each earthquake – by a segment of a single fault. Accordingly, the signals of an approaching earthquake come not from a narrow vicinity of the source but from a much wider area.*

*What is the size of such areas?* Let  $M$  and  $L(M)$  be the earthquake magnitude and the characteristic length of its source, respectively. In the intermediate-term prediction

(on a time scale of years) that size may reach  $10L(M)$ ; it might be reduced down to  $3L$  or even to  $L$  in a second approximation [43]. On a time scale of about 10 years that size reaches about  $100L$ . For example, according to [71], the Parkfield (California) earthquake with  $M$  about 6 and  $L \approx 10$  km “... *is not likely to occur until activity picks up in the Great Basin or the Gulf of California*”, about 800 km away.

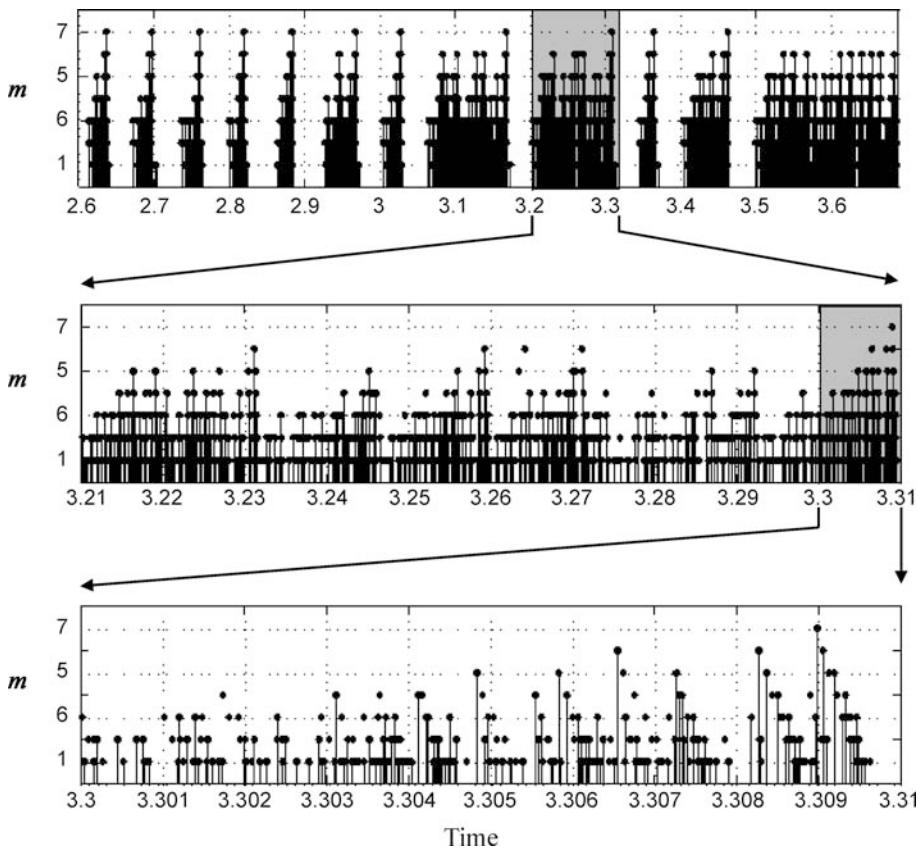
*Historical perspective.* An early estimate of the area where premonitory patterns are formed was obtained in [33] for a premonitory rise of seismic activity. C. Richter, who was sceptical about the feasibility of earthquake prediction, made an exception to that pattern, specifically because it was defined in large areas. He wrote [75]: “... *It is important that (the authors) confirm the necessity of considering a very extensive region including the center of the approaching event. It is very rarely true that the major event is preceded by increasing activity in its immediate vicinity.*”

However, such spreading of premonitory patterns has been often regarded as counterintuitive in earthquake prediction research on the grounds that earthquakes can't trigger each other at such distances. The answer is that earthquakes forming such patterns do not trigger each other but reflect an underlying large-scale dynamics of the lithosphere. Among the indisputable manifestations of that correlation are the following phenomena: migration of earthquakes along fault zones [47,52,58,90]; alternate rise of seismicity in distant areas [71] and even in distant tectonic plates [76]. Global correlations have been found also between major earthquakes and other geophysical phenomena, such as Chandler wobble, variations of magnetic field, and the velocity of Earth's rotation [34,72]. These correlations may be explained by several mechanisms not mutually exclusive. Such mechanisms range from micro-fluctuations of large scale tectonic movements to impact of migrating fluids (e. g., [1,5,7,10,69,71,82,84,89]).

### Third Paradigm: Similarity

*Premonitory phenomena are similar (identical after normalization) in the extremely diverse environments and in a broad energy range (e. g., [1,33,36]). The similarity is not unlimited however and regional variations of premonitory phenomena do emerge.*

Normalized prediction algorithms retain their prediction power in active regions and platforms, with the magnitude of target earthquakes ranging from 8.5 to 4.5. Furthermore, similarity extends to induced seismicity, and to multiple fracturing in engineering constructions and lab-



**Geo-complexity and Earthquake Prediction, Figure 9**

Synthetic earthquake sequence consecutively zoomed. Shaded areas mark zoomed intervals. The model shows the rich variety of behavior on different timescales. Note that the ratio of timescales for the top and bottom panels is  $10^2$ . After [17]

oratory samples (e.g., [3,35,43]). Ultimately, a single but explicit demonstration of similarity was obtained for starquakes – ruptures of the crust of neutron star [45], where the conditions are extremely different than in the Earth. Altogether the corresponding elastic energy release ranges from ergs to  $10^{25}$  ergs (even to  $10^{46}$  ergs if the starquake is counted in).

However, the performance of prediction algorithms does vary from region to region (see [21,35,63]). It is not yet clear whether this is due to imperfect normalization, or to limitations on similarity itself.

#### Fourth Paradigm: Dual Nature of Premonitory Phenomena

Some premonitory patterns are “universal”, common for hierarchical complex systems of different origin; other are specific to geometry of fault networks or to a certain physical mechanism controlling the strength and stress fields in the lithosphere.

**Universal patterns.** These are most of the patterns so far known. They can be reproduced on models not specific to the Earth only, e.g. models of a statistical physics type (direct or inverse cascade, colliding cascades, percolation, dynamical clustering), models of critical phenomena in fluid dynamics, as well as Earth-specific models themselves.

Complete analytical definition of premonitory patterns was obtained recently on the branching diffusion model [18]. Definition includes only three control parameters, thus strongly reducing uncertainty in data analysis (Sect. “Data Analysis”).

Reviews of such models can be found in [15,17,36,66, 83,89,93]. Discussion of particular patterns is given also in [25,42,67,68,88,92].

An example of an earthquake sequence generated by a universal model is shown in Fig. 9 [17]. The modeled seismicity exhibits major features of real seismicity: seismic cycle, switching of seismic regime, the Gutenberg–Richter relation, foreshocks and aftershocks, long-range

correlation, and, finally, the premonitory seismicity patterns.

**Earth-specific patterns** are not yet incorporated in prediction algorithms. We discuss here the patterns reflecting the state of the nodes – structures where the strong earthquakes are nucleated (see Sect. “[Structure](#)”). Quantitative characteristics of that state are geometric incompatibility  $G$  (Sect. “[Geometric Instability](#)”). It shows whether the nodes are locked up or unlocked and quantifies their tendency to fracture and change of the faults geometry. Change of  $G$  might create or dissolve such feature as asperities, relaxation barriers, weak links, and replacement of seismicity by creep or “silent” earthquakes [16]. These features would migrate from node to node with velocity typical of seismicity migration: tens to hundreds km/year [90]. All this makes monitoring of  $G$  highly relevant to detecting premonitory patterns. A simple pattern of that kind is seismic quiescence around the soon-to-break nodes (e.g., [44,58,77]). A simple highly promising possibility is considering separately premonitory phenomena inside and outside of nodes (e.g., [77]).

### **Earthquake Prediction and Earthquake Preparedness**

Given the limited accuracy of predictions, how do we use them for damage reduction? The key to this is to escalate or de-escalate preparedness depending on the following: content of the current alarm (what and where is predicted), probability of a false alarm, and cost/benefit ratio of disaster preparedness measures. Prediction might be useful if its accuracy is *known*, even if it is not high. Such is the standard practice in preparedness for all disasters, war included.

### **Diversity of Damage**

Earthquakes hurt population, economy, and environment in very different ways: destruction of buildings, lifelines, etc; triggering fires; releasing of toxic, radioactive and genetically active materials; triggering other natural disasters, such as floods, avalanches, landslides, tsunamis, etc.

Equally dangerous are the socio-economic and political consequences of earthquakes: disruption of vital services (supply, medical, financial, law enforcement, etc.), epidemics, drop of production, slowdown of economy, unemployment, disruptive anxiety of population, profiteering and crime. The socio-economic consequences may be inflicted also by the undue release of predictions.

Different kinds of damage are developing at different time and space scales, ranging from immediate damage to

chain reaction, lasting tens of years and spreading regionally if not worldwide.

**Diversity of Disaster Preparedness Measures** Such diversity of damage requires a hierarchy of disaster preparedness measures, from building code and insurance to mobilization of post disaster services to red alert. It takes different times, from decades to seconds to undertake different measures; having different cost they can be maintained for different time periods; and they have to be spread over different territories, from selected sites to large regions. No single stage can replace another one for damage reduction and no single measure is sufficient alone.

On the other hand many important measures are inexpensive and do not require high accuracy of prediction. An example is the Northridge, California, earthquake, 1994, which caused economic damage exceeding \$30 billion. Its prediction, published well in advance [48], was not precise – the alarm covered a time period of 18 months and an area 340 km in diameter with dramatically uneven vulnerability. However, low-cost actions, undertaken in response to this prediction (e.g. an out of turn safety inspection) would be well justified if even just a few percent of the damage were prevented.

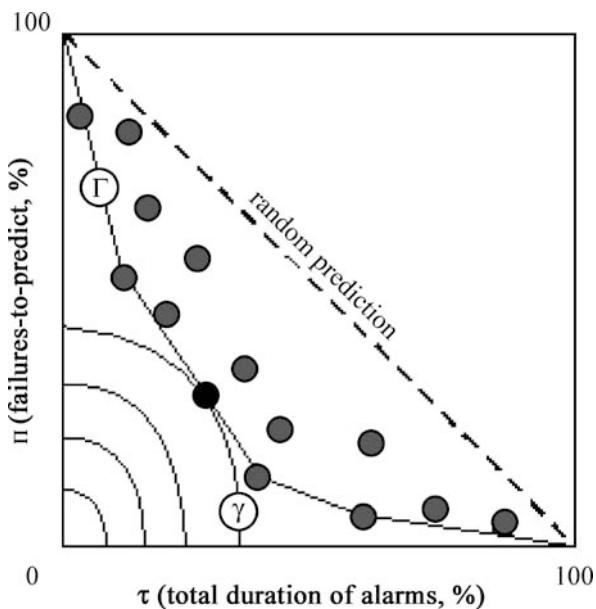
### **Joint Optimization of Prediction and Preparedness**

The choice of preparedness measures is by no means unique. Different measures may supersede or mutually exclude one another, leaving the decision-maker a certain freedom of choice [38]. The definition of the prediction algorithm is not unique either. The designer of the algorithm has certain freedom to choose the tradeoff between different characteristics of its accuracy (rate of failures to predict, duration of alarms, and rate of failures to predict) by varying adjustable elements of the algorithm (Sect. “[General Scheme of Prediction](#)”). That leads to the problem, typical for decision-making with incomplete information: to optimize jointly prediction and preparedness. Figure 10 shows the scheme of such optimization. This figure shows also advantages of a new formulation of prediction: parallel applications of several versions of an algorithm.

Further discussion can be found in [27,28,63,64].

### **Further Goals**

Particularly encouraging for further earthquake prediction research is the wealth of relevant data, models, and theories that are available and yet untapped (the *want amidst plenty* pattern, Conference and School on Predictability of Natural Disasters for our Planet in Danger. A System View: Theory, Models, Data Analysis, 25 June – 6 July



**Geo-complexity and Earthquake Prediction, Figure 10**

Joint optimization of prediction and preparedness based on the theory of optimal control. Dots show points on the error diagram.  $\Gamma$  is their envelope. Thin contours ( $\gamma$ ) show loss curves with constant value of a prevented loss. Optimal strategy is the tangent point of contours  $\Gamma$  and  $\gamma$ . After [63]

2007, Trieste, ICTP, [http://cdsagenda5.ictp.it/full\\_display.php?ida=a06204](http://cdsagenda5.ictp.it/full_display.php?ida=a06204)). Likely within reach is a new generation of prediction algorithms, about five- to ten-fold more accurate than existing ones.

In the general scheme of things, this is a part of wider developments: Emergence of the newly integrated dynamics of the solid Earth, extending from a fundamental concept succeeding plate tectonics to predictive understanding and (with luck) control of geological and geotechnical disasters. And predictive understanding of extreme events (critical phenomena) in the complex systems formed, separately and jointly, by nature and society.

## Acknowledgment

The authors are sincerely grateful for insightful comments of Edo Nyland, William Lee, Michele Caputo, and Antoni Correig.

## Bibliography

### Primary Literature

1. Aki K (1996) Scale dependence in earthquake phenomena and its relevance to earthquake prediction. *Proc Natl Acad Sci USA* 93:3740–3747
2. Alekseevskaya MA, Gabrielov AM, Gvishiani AD, Gelfand IM, Ranzman EY (1977) Formal morphostructural zoning of mountain territories. *J Geophys* 43:227–233
3. Allegre CJ, Le Mouël J-L, Provost V (1982) Scaling rules in rock fracture and possible implications for earthquake prediction. *Nature* 297:47–49
4. Bak P, Chen K, Tang C (1992) A forest-fire model and some thoughts on turbulence. *Phys Lett A* 147:297–300
5. Barenblatt GI (1993) Micromechanics of fracture. In: Bodner ER, Singer J, Solan A, Hashin Z (eds) *Theoretical and Applied Mechanics*. Elsevier, Amsterdam, pp 25–52
6. Barenblatt G (1996) *Scaling, Self-similarity, and Intermediate Asymptotics*. Cambridge University Press, Cambridge
7. Barenblatt GI, Keilis-Borok VI, Monin AS (1983) Filtration model of earthquake sequence. *Trans (Doklady) Acad Sci SSSR* 269:831–834
8. Bird P (1998) Testing hypotheses on plate-driving mechanisms with global lithosphere models including topography, thermal structure, and faults. *J Geophys Res* 103(B5):10115–10129
9. Bongard MM, Vaintsveig MI, Guberman SA, Izvekova ML, Smirnov MS (1966) The use of self-learning programs in the detection of oil containing layers. *Geol Geofiz* 6:96–105 (in Russian)
10. Bowman DD, Ouillon G, Sammis GG, Sornette A, Sornette D (1998) An observational test of the critical earthquake concept. *J Geophys Res* 103:24359–24372
11. Burridge R, Knopoff L (1967) Model and theoretical seismicity. *Bull Seismol Soc Am* 57:341–360
12. Crutchfield JP, Farmer JD, Packard NH, Shaw RS (1986) Chaos. *Sci Am* 255:46–57
13. Farmer JD, Sidorowich J (1987) Predicting chaotic time series. *Phys Rev Lett* 59:845
14. Gabrielov AM, Keilis-Borok VI (1983) Patterns of stress corrosion: Geometry of the principal stresses. *Pure Appl Geophys* 121:477–494
15. Gabrielov A, Dmitrieva OE, Keilis-Borok VI, Kossobokov VG, Kuznetsov IV, Levshina TA, Mirzoev KM, Molchan GM, Negmatullaev SK, Pisarenko VF, Prozoroff AG, Rinehart W, Rotwain IM, Shebalin PN, Shnirman MG, Shreider SY (1986) Algorithm of Long-term Earthquakes' Prediction. Centro Regional de Sismología para America del Sur, Lima
16. Gabrielov AM, Keilis-Borok VI, Jackson DD (1996) Geometric incompatibility in a fault system. *Proc Natl Acad Sci USA* 93:3838–3842
17. Gabrielov AM, Zaliapin IV, Newman WI, Keilis-Borok VI (2000) Colliding cascade model for earthquake prediction. *Geophys J Int* 143(2):427–437
18. Gabrielov A, Keilis-Borok V, Zaliapin I (2007) Predictability of extreme events in a branching diffusion model. arXiv: 0708.1542 [nlin.AO]
19. Gelfand IM, Guberman SA, Keilis-Borok VI, Knopoff L, Press F, Ranzman IY, Rotwain IM, Sadovsky AM (1976) Pattern recognition applied to earthquake epicenters in California. *Phys Earth Planet Inter* 11:227–283
20. Gell-Mann M (1994) *The Quark and the Jaguar: Adventures in the Simple and the Complex*. Freeman and Company, New York
21. Ghil M (1994) Cryothermodynamics: the chaotic dynamics of paleoclimate. *Physica D* 77:130–159
22. Gorshkov A, Kossobokov V, Soloviev A (2003) Recognition of

- earthquake-prone areas. In: Keilis-Borok VI, Soloviev AA (eds) Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer, Berlin-Heidelberg, pp 239–310
23. Grotzinger J, Jordan TH, Press F, Siever R (2007) Understanding Earth, 5th edn. WH Freeman & Co, New York
  24. Holland JH (1995) Hidden Order: How Adaptation Builds Complexity. Addison-Wesley, Reading
  25. Huang Y, Saleur H, Sammis C, Sornette D (1998) Precursors, aftershocks, criticality and self-organized criticality. *Europhys Lett* 41:43–48
  26. Kadanoff LP (1976) Scaling, universality and operator algebras. In: Domb C, Green MS (eds) Phase Transitions and Critical Phenomena, vol 5a. Academic Press, London, pp 1–34
  27. Kantorovich LV, Keilis-Borok VI (1991) Earthquake prediction and decision-making: social, economic and civil protection aspects. In: International Conference on Earthquake Prediction: State-of-the-Art. Scientific-Technical Contributions, CSEM-EMSC, Strasbourg, pp 586–593
  28. Kantorovich LV, Keilis-Borok VI, Molchan GM (1974) Seismic risk and principles of seismic zoning. In: Seismic design decision analysis. Internal Study Report 43, Department of Civil Engineering, MIT, Cambridge (Mass)
  29. Keilis-Borok VI (1990) The lithosphere of the Earth as a nonlinear system with implications for earthquake prediction. *Rev Geophys* 28:19–34
  30. Keilis-Borok VI (ed) (1990) Intermediate-Term Earthquake Prediction: Models, Algorithms, Worldwide Tests. *Phys Earth Planet Inter*, special issue 61(1–2):1–139
  31. Keilis-Borok VI (2002) Earthquake prediction: State-of-the-art and emerging possibilities. *Annu Rev Earth Planet Sci* 30:1–33
  32. Keilis-Borok VI, Kossobokov VG (1990) Premonitory activation of earthquake flow: algorithm M8. *Phys Earth Planet Inter* 61(1–2):73–83
  33. Keilis-Borok VI, Malinovskaya LN (1964) One regularity in the occurrence of strong earthquakes. *J Geophys Res* 69:3019–3024
  34. Keilis-Borok VI, Press F (1980) On seismological applications of pattern recognition. In: Allegre CJ (ed) Source Mechanism and Earthquake Prediction Applications. Editions du Centre national de la recherche scientifique, Paris, pp 51–60
  35. Keilis-Borok VI, Shebalin PN (eds) (1999) Dynamics of Lithosphere and Earthquake Prediction. *Phys Earth Planet Inter*, special issue 111(3–4):179–327
  36. Keilis-Borok VI, Soloviev AA (eds) (2003) Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer, Berlin-Heidelberg
  37. Keilis-Borok V, Shebalin P, Gabrielov A, Turcotte D (2004) Reverse tracing of short-term earthquake precursors. *Phys Earth Planet Inter* 145(1–4):75–85
  38. Keilis-Borok V, Davis C, Molchan G, Shebalin P, Lahr P, Plumb C (2004) Earthquake prediction and disaster preparedness: Interactive algorithms. *EOS Trans AGU* 85(47), Fall Meet Suppl, Abstract S22B-02a
  39. King G (1983) The accommodation of large strains in the upper lithosphere of the earth and other solids by self-similar fault systems: The geometrical origin of b-value. *Pure Appl Geophys* 121:761–815
  40. King G (1986) Speculations on the geometry of the initiation and termination processes of earthquake rupture and its relation to morphology and geological structure. *Pure Appl Geophys* 124:567–583
  41. Knopoff L (1969) The upper mantle of the Earth. *Science* 163:1277–1287
  42. Kossobokov VG, Carlson JM (1995) Active zone size vs. activity: A study of different seismicity patterns in the context of the prediction algorithm M8. *J Geophys Res* 100:6431–6441
  43. Kossobokov V, Shebalin P (2003) Earthquake Prediction. In: Keilis-Borok VI, Soloviev AA (eds) Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer, Berlin-Heidelberg, pp 141–207
  44. Kossobokov VG, Keilis-Borok VI, Smith SW (1990) Localization of intermediate-term earthquake prediction. *J Geophys Res* 95:19763–19772
  45. Kossobokov VG, Keilis-Borok VI, Cheng B (2000) Similarities of multiple fracturing on a neutron star and on the Earth. *Phys Rev E* 61(4):3529–3533
  46. Kravtsov YA (ed) (1993) Limits of Predictability. Springer, Berlin-Heidelberg
  47. Kuznetsov IV, Keilis-Borok VI (1997) The interrelation of earthquakes of the Pacific seismic belt. *Trans (Doklady) Russ Acad Sci, Earth Sci Sect* 355A(6):869–873
  48. Levshina T, Vorobieva I (1992) Application of algorithm for prediction of a strong repeated earthquake to the Joshua Tree and the Landers earthquakes' aftershock sequence. *EOS Trans AGU* 73(43), Fall Meet Suppl:382
  49. Sir Lighthill J (ed) (1996) A Critical Review of VAN. World Scientific, Singapore-New Jersey-London-Hong Kong
  50. Lomnitz-Adler J (1991) Model for steady state friction. *J Geophys Res* 96:6121–6131
  51. Lorenz EN (1963) Deterministic nonperiodic flow. *J Atmos Sci* 20:130–141
  52. Ma Z, Fu Z, Zhang Y, Wang C, Zhang G, Liu D (1990) Earthquake Prediction: Nine Major Earthquakes in China. Springer, New York
  53. Ma S-K (1976) Modern Theory of Critical Phenomena. WA Benjamin, Inc, Reading
  54. Mason IB (2003) Binary events. In: Jolliffe IT, Stephenson DB (eds) Forecast Verification. A Practitioner's Guide in Atmospheric Science. Wiley, Chichester, pp 37–76
  55. McKenzie DP, Morgan WJ (1969) The evolution of triple junctions. *Nature* 224:125–133
  56. McKenzie DP, Parker RL (1967) The North Pacific: An example of tectonics on a sphere. *Nature* 216:1276–1280
  57. Minster JB, Jordan TH (1984) In: Crouch JK, Bachman SB (eds) Tectonics and Sedimentation Along the California Margin: Pacific Section, vol 38. Academic, San Diego, pp 1–16
  58. Mogi K (1968) Migration of seismic activity. *Bull Earth Res Inst Univ Tokyo* 46(1):53–74
  59. Molchan GM (1990) Strategies in strong earthquake prediction. *Phys Earth Planet Inter* 61:84–98
  60. Molchan GM (1991) Structure of optimal strategies of earthquake prediction. *Tectonophys* 193:267–276
  61. Molchan GM (1994) Models for optimization of earthquake prediction. In: Chowdhury DK (ed) Computational Seismology and Geodynamics, vol 1. Am Geophys Un, Washington, DC, pp 1–10
  62. Molchan GM (1997) Earthquake prediction as a decision-making problem. *Pure Appl Geophys* 149:233–237
  63. Molchan GM (2003) Earthquake Prediction Strategies: A Theoretical Analysis. In: Keilis-Borok VI, Soloviev AA (eds) Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer, Berlin-Heidelberg, pp 209–237

64. Molchan G, Keilis-Borok V (2008) Earthquake prediction: Probabilistic aspect. *Geophys J Int* 173(3):1012–1017
65. Molchan GM, Dmitrieva OE, Rotwain IM, Dewey J (1990) Statistical analysis of the results of earthquake prediction, based on burst of aftershocks. *Phys Earth Planet Inter* 61:128–139
66. Newman W, Gabrielov A, Turcotte DL (eds) (1994) Nonlinear Dynamics and Predictability of Geophysical Phenomena. Am Geophys Un, Int Un Geodesy Geophys, Washington
67. Newman WI, Turcotte DL, Gabrielov AM (1995) Log-periodic behavior of a hierarchical failure model with application to precursory seismic activation. *Phys Rev E* 52:4827–4835
68. Pepke GF, Carlson JR, Shaw BE (1994) Prediction of large events on a dynamical model of fault. *J Geophys Res* 99:6769–6788
69. Pollitz FF, Burgmann R, Romanowicz B (1998) Viscosity of oceanic asthenosphere inferred from remote triggering of earthquakes. *Science* 280:1245–1249
70. Press F (ed) (1965) Earthquake Prediction: A Proposal for a Ten Year Program of Research. Ad Hoc Panel on Earthquake Prediction. White House Office of Science and Technology, Washington, DC, p 134
71. Press F, Allen C (1995) Patterns of seismic release in the southern California region. *J Geophys Res* 100(B4):6421–6430
72. Press F, Briggs P (1975) Chandler wobble, earthquakes, rotation and geomagnetic changes. *Nature* (London) 256:270–273
73. Prozorov AG (1975) Changes of seismic activity connected to large earthquakes. In: Keilis-Borok VI (ed) Interpretation of Data in Seismology and Neotectonics. Comput Seismol vol 8. Nauka, Moscow, pp 71–82 (in Russian)
74. Prozorov AG, Schreider SY (1990) Real time test of the long-range aftershock algorithm as a tool for mid-term earthquake prediction in Southern California. *Pure Appl Geophys* 133:329–347
75. Richter C (1964) Comment on the paper “One Regularity in the Occurrence of Strong Earthquakes” by Keilis-Borok VI and Malinovskaya LN. *J Geophys Res* 69:3025
76. Romanowicz B (1993) Spatiotemporal patterns in the energy-release of great earthquakes. *Science* 260:1923–1926
77. Rundkvist DV, Rotwain IM (1996) Present-day geodynamics and seismicity of Asia minor. In: Chowdhury DK (ed) Computational Seismology and Geodynamics, vol 1. Am Geophys Un, Washington, DC, pp 130–149
78. Rundle JB, Turcotte DL, Klein W (eds) (2000) Geocomplexity and the Physics of Earthquakes. Am Geophys Un, Washington, DC
79. Sammis CG, Sornette D, Saleur H (1996) Complexity and earthquake forecasting. In: Rundle JB, Turcotte DL, Klein W (eds) SFI Studies in the Science of Complexity, vol XXV. Addison-Wesley, Reading
80. Shebalin PN, Keilis-Borok VI (1999) Phenomenon of local “seismic reversal” before strong earthquakes. *Phys Earth Planet Inter* 111:215–227
81. Shebalin P, Keilis-Borok V, Gabrielov A, Zaliapin I, Turcotte D (2006) Short-term earthquake prediction by reverse analysis of lithosphere dynamics. *Tectonophysics* 413:63–75
82. Soloviev A, Ismail-Zadeh A (2003) Models of Dynamics of Block-and-Fault Systems. In: Keilis-Borok VI, Soloviev AA (eds) Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer, Berlin-Heidelberg, pp 71–139
83. Sornette D (2004) Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization, and Disorder. Concept and Tools, 2nd edn. Springer, Berlin-Heidelberg
84. Sornette D, Sammis CG (1995) Complex critical exponents from renormalization group theory of earthquakes: Implications for earthquake predictions. *J Phys I France* 5:607–619
85. Traskin VY, Skvortsova ZN (2006) Rehbinder effect in geodynamical processes. In: Kissin IG, Rusinov VL (eds) Fluids and Geodynamics. Nauka, Moscow, pp 147–164 (in Russian)
86. Tukey JW (1977) Exploratory Data Analysis. Addison-Wesley Series in Behavioral Science: Quantitative Methods. Addison-Wesley, Reading
87. Turcotte DL (1997) Fractals and Chaos in Geology and Geophysics, 2nd edn. Cambridge University Press, Cambridge
88. Turcotte DL (1999) Seismicity and self-organized criticality. *Phys Earth Planet Inter* 111:275–294
89. Turcotte DL, Newman WI, Gabrielov A (2000) A statistical physics approach to earthquakes. In: Geocomplexity and the Physics of Earthquakes. Am Geophys Un, Washington, DC
90. Vil'kovich EV, Shnirman MG (1983) Epicenter migration waves: Examples and models. In: Keilis-Borok VI, Levshin AL (eds) Mathematical models of the structure of the Earth and the earthquake prediction. Comput Sismol, vol 14. Allerton Press, New York, pp 27–36
91. Vorobieva IA, Levshina TA (1994) Prediction of a second large earthquake based on aftershock sequence. In: Chowdhury DK (ed) Computational Seismology and Geodynamics, vol 1. Am Geophys Un, Washington, DC, pp 27–36
92. Yamashita T, Knopoff L (1992) Model for intermediate-term precursory clustering of earthquakes. *J Geophys Res* 97:19873–19879
93. Zaliapin I, Keilis-Borok V, Ghil M (2003) A Boolean delay model of colliding cascades. II: Prediction of critical transitions. *J Stat Phys* 111(3–4):839–861

## Books and Reviews

- Agnew DC, Ellsworth WL (1991) Earthquake prediction and long-term hazard assessment. *Rev Geophys Suppl* 29:877–889
- Aki K (1981) A probabilistic synthesis of precursory phenomena. In: Simpson DV, Richards PG (eds) Earthquake Prediction. An International Review. Manrice Ewing Ser 4. Am Geophys Un, Washington, DC, pp 566–574
- Bolt BA (1993) Earthquakes – Newly Revised and Expanded. Freeman, New York
- Bongard MM (1967) The Problem of Recognition. Nauka, Moscow (in Russian)
- Conference and School on Predictability of Natural Disasters for our Planet in Danger. A System View: Theory, Models, Data Analysis, 25 June–6 July 2007, the Abdus Salam International Centre for Theoretical Physics, Trieste, [http://cdsagenda5.ictp.it/full\\_display.php?ida=a06204](http://cdsagenda5.ictp.it/full_display.php?ida=a06204)
- Jaumé SC, Sykes LR (1999) Evolving towards a critical point: A review of accelerating seismic moment/energy release prior to large and great earthquakes. *Pure Appl Geophys* 155:279–306
- Kanamori H (1977) The energy release in great earthquakes. *J Geophys Res* 82(B20):2981–2988
- Mandelbrot B (1983) The Fractal Geometry of Nature. Freeman, New York
- Ranzman Ela (1979) Places of Earthquakes and Morphostructures of Mountain Countries. Nauka, Moscow (in Russian)
- Varnes DJ (1989) Predicting earthquakes by analyzing accelerating precursory seismic activity. *Pure Appl Geophys* 130:661–686

- Wyss M (ed) (1991) Evaluation of Proposed Earthquake Precursors. Am Geophys Un, Washington DC
- Wyss M (1997) Second round of evaluation of proposed earthquake precursors. Pure Appl Geophys 149:3–16
- Wyss M, Habermann R (1988) Precursory seismic quiescence. Pure Appl Geophys 126:319–332
- Zoller G, Hainzl S, Kurths J (2001) Observation of growing correlation length as an indicator for critical point behaviour prior to large earthquakes. J Geophys Res 106:2167–2176

## Geometric Phase and Related Phenomena in Quantum Nanosystems

SHI-JIE XIONG

National Laboratory of Solid State Microstructures and Department of Physics, Nanjing University, Nanjing, China

### Article Outline

Glossary

Definition of the Subject

Introduction

Geometric Phase Induced in Different Types of Nanostructures

Berry Phase: Adiabatic Approximation

Theoretical Treatment Beyond Adiabatic Approximation

Effect of Phase Uncertainty

Future Directions

Acknowledgments

Bibliography

### Glossary

**Target system** The system under investigation. Usually the target system is a part of degrees of freedom in a nanostructure whose state is described quantum mechanically.

**Environment** The system which interacts with the target system and influences its state. The environment may be the parts of degrees of freedom in the nanostructure other than the target, the systems which surround the nanostructure and interact with the target system, or an external electric or magnetic field which is exerted on the target system.

**Hamiltonian** The quantum operator that describes the energy of the system and acts on the Hilbert space for the quantum states. It is usually denoted by  $\hat{H}$  with subscript indicating the described system.

**Eigen energy and eigen wave function** For time-independent (stationary) Hamiltonian the system can only take specific values of energy, called the eigen energies. Corresponding every eigen energy there is a state of the system described by the eigen wave function. All eigen wave functions are normalized and orthogonal to each other, forming a complete linear space.

**Hilbert space** A complete linear space spanned by all eigen wave functions of the time-independent Hamiltonian. Any states of the related system can be expressed as a vector in this linear space.

**Parameter space** The space spanned by parameters which specify the interactions of the environment on the target system. In a dynamical process in which the geometric phase is investigated, the parameters are assumed to be periodically varying in time. The period is denoted as  $T$ .

**Evolution of wave functions** During a period of variation of parameters, the Hamiltonian transverses a circle and returns to its initial situation at the end of the period. In this process the wave function of the target system also undergoes an evolution with the time.

**Dynamical phase** Even though the Hamiltonian is time independent, the wave function of the target system still has a time dependent phase factor denoted as  $e^{i\gamma_d(t)}$ . In a period  $T$  of a periodic dynamical process the phase  $\gamma_d$  acquired by the target wave function, which is calculated under the assumption that the Hamiltonian is the averaged one and stationary in this period, is called the dynamical phase.

**Geometric phase** If the Hamiltonian is periodically time-dependent, the phase  $\gamma$  acquired by the target wave function in a period of the evolution is different from the dynamical phase  $\gamma_d$ . The difference  $\gamma_g = \gamma - \gamma_d$  is called the geometric phase.

**Adiabatic approximation** If the evolution of the Hamiltonian is slow enough so that there is no transition between different eigen wave functions of the target system during the evolution, this evolution is regarded as adiabatic. The approximation based on the adiabatic assumption of the process is called adiabatic approximation.

**Berry phase** The geometric phase calculated in the adiabatic approximation is called the Berry phase.

**Quantum interference** In the case where the states of a nanostructure are quantum mechanically described by wave functions, the amplitude of a resultant wave function can be constructively enhanced when its components have the same phase, or can be destructively weakened when its components have opposite phases. Generically the resultant wave function de-

pends not only on the amplitudes of its components, but also depends on the phases of the components. This phenomenon is called quantum interference.

**Dephasing** The processes or mechanisms which cause spatial or temporal uncertainty of phase of the target wave functions.

### Definition of the Subject

Owing to the rapid growing of the nanotechnology, now it is possible to design and fabricate different types of nanostructures which have nanometer length scales in one, two, or three dimensions. Examples of nanostructures are: quantum wells [12], quantum wires [33], and quantum dots [15] in which the motion of charge carriers is restricted within nanometer length scale in one, two, and three dimensions, respectively. Usually quantum wells, wires, and dots are made of semiconductor materials and based on the related designing and fabricating technology. Other examples are those of carbon-based nano-materials, such as graphene [26], carbon nanotubes [17], and fullerenes  $C_{60}$  or  $C_{70}$  [10], in which the motion of electrons is also restricted within nanometer scale in one, two, and three dimensions, respectively. The electric, optical, and magnetic properties of these nanostructures are mainly depending on the states of electrons. So here we define the electron system in nanostructures as the target system under our investigation. Other degrees of freedom in the nanostructure, such as vibrations of nuclei and other collective excitations, and other materials attaching or near the nanostructure, can be regarded as the environment. In the nanometer scale of some dimensions of the nanostructure the quantum coherence is usually maintained and the effect of quantum interference becomes important. In a dynamical process of the environment, the quantum interference of the target system depends not only on the dynamical phase, but also on the geometric phase. Thus, the investigation of the geometric phase and its effect on nanostructures is of particular interest. Such investigation can reveal the physical origins of peculiar properties of nanostructures observed in experiments and suggest possible methods in designing devices with specific properties.

## Introduction

### Nanostructures

Recently there has been extensive interest in designing, fabricating, and investigating various types of nanostructures [29]. Owing to the advances of technology, people are now able to arrange atoms into structures that are

only a few nanometers in size. A nanometer is about four atoms in diameter. The goal in doing this is in specific self-assembly of nanostructures, which can produce large amounts of artificial materials with new properties. In order to connect nanostructures to the more familiar world of microstructures and microelectronics, one must build up larger assemblies by guided self-assembly.

A common and striking feature of nanostructures is in the length scale of nanometers in one or more dimensions. If an electron is confined within this length scale in one direction, its momentum in this direction will be well quantized due to the boundary condition and, correspondingly, the energy levels characterizing the kinetic motion in this direction are separated so that at low temperatures only one level plays the central role in determining the transport or other properties of the system. This is so-called quantum size effect in nanostructures. An example is the nanowire or carbon nanotube where the motions of electrons in the transverse directions are quantized but the motion in the longitudinal direction is confined and quantized. As a consequence there is only one or few effective energy levels in the transverse directions which correspond one or few traveling channels in the longitudinal direction. If there is only one channel, the system can be viewed as a true one-dimensional one in the longitudinal direction. The quantum size effect is an important manifestation of the quantum effects in nanostructures, implying the central role of the quantum principles.

Besides the quantum size effect, abundant and charming properties of various nanostructures arise from their specific crystal structures, adjusted shapes, and designed arrangement of self-assembly. Such particular structures can provide complicated networks of interfering paths and produce specific band structures which may be useful for applications. A simplest example is the Aharonov–Bohm (AB) rings made of quantum wires which can directly exhibit the quantum interference effect from the AB oscillations of the conductance through the rings as a function of the magnetic flux through the ring [34]. This device makes use of the magnetic flux as a method for controlling the quantum phase of the electron wave functions. Another example is the graphene, a single atomic layer of the graphite where carbon atoms are arranged on a honeycomb lattice. Owing to the specific lattice structure of the graphene, the electrons show massless Dirac relativistic dispersion relation [26]. As we will see below, such a specific band structure gives rise to the Berry or geometric phase and causes exotic transport properties.

The various shapes and structures provide possible methods of adjusting geometric phase  $j$  and, in turn, controlling the transport properties in nanostructures. The

physical ingredients taking part in these processes include: (i) the magnetic field or flux through some part of nanostructures [34]; (ii) the spin-orbit coupling which governs the spin precessions during the transport of electrons [28]; (iii) the Jahn-Teller electron-lattice interaction [16]; (iv) the exchange interaction between electrons and magnetic impurities; and (v) the superconducting pairing. Any of these ingredients can take action only in association with special shapes of the nanostructures. For instance, the AB oscillations can occur by varying the magnetic flux only in a ring-shaped device.

### Geometric or Berry Phase

We consider that a point (representing a car or a ship) carrying two vectors is traversing along a loop on the surface of a sphere. Both vectors stay tangential to the surface at all times and they remain as parallel as possible to the direction they were pointing before each infinitesimal displacement. In this special example the long vector points “South” and the short one points “East” all the time. After completing the loop, the vectors go back to the original point, but they find themselves rotated with respect to the directions they were pointing at the beginning. Note that the vectors “rotate” despite the fact that we have been careful to keep them parallel during the transport. If the loop is smaller, the rotation angle would have been smaller. For a loop encircling a 1/8, the rotation angle is  $\pi/2$ . A larger path surrounding for example one quarter of the sphere, would rotate the parallel-transported vectors by  $\pi$ . An even larger loop, a diameter, surrounding (on both sides) half the sphere gives a rotation angle of  $2\pi$ , i. e. no rotation at all. This rotation originates from the intrinsic curvature of the sphere surface. This phenomenon can not appear for the transport along a flat manifold, such as a plane or a cylinder. So the rotation angle is related to the integral of the curvature over the surface bounded by the loop.

In the quantum version such rotation angles correspond to the Berry phases. In quantum mechanics the state of the target system, i. e., an electron in a nanostructure, is characterized by the wave function  $\psi$  which has not only the magnitude  $|\psi|$  but also a phase factor  $e^{i\gamma}$ . Phase  $\gamma$  is equivalent to the angle in the complex plane. Several specific cases had been recognized for many years, in particular the well-studied Aharonov-Bohm effect in the ring of quantum wire threaded with a magnetic flux. In 1984, Berry published an influential formal systematization of the closed-path geometric phase in quantum mechanical problems [8]. In the following we summarize some major points about the Berry phase or geometric phase in nanostructures.

Wave functions of the target system are parameterized vector fields in the Hilbert space. Each wave function has a phase connected to its time evolution. The dynamical phase is not geometrical: it depends on the energy at which the parameterized path is followed, and on the arbitrary choice of a “zero phase” (technically called a gauge choice) at each point along the path. A recipe is needed to get rid of any dynamical phases, and single out the intrinsic, geometrical invariant phases. Let us consider the set of eigen wave functions of a Hamiltonian  $\hat{H}$ , depending on several classical “external” parameters collectively denoted by  $q$ . Changing  $q$  from one point to another in the multi-dimensional parameter space, the set of eigen wave functions and corresponding eigen energies of  $\hat{H}$  change to totally new and unrelated eigen system. In the adiabatic approximation smooth evolution of  $\hat{H}$  is assumed so that one expects that for infinitesimal change  $\Delta q$  in the parameters, the inner product between a given eigen wave function at some point  $q$ , and the evolved eigen wave function at a following neighboring point  $q + \Delta q$ , should be as close to 1 as possible with infinitesimal error of the order  $\Delta q^2$ .

We note that every stationary wave function “rotates” as  $e^{-iEt/\hbar}$ , where  $E$  is the energy of the state,  $t$  is time and  $\hbar$  is Planck’s constant, because of the Schrödinger time evolution. On top of this, since at each point the eigen system is unrelated to that at a different point, nothing forbids to multiply each eigen wave function by an additional arbitrary phase factor  $e^{-i\phi(q)}$  (a “gauge choice”). The correct rule for the “parallel transport”, as sketched above, permits to calculate the geometrical phase, such that it ignores completely any dynamical and gauge phase. The result depends only on the geometry of the loop. In this sense, Berry phase is said to be “gauge-invariant”.

In 1987, Aharonov and Anandan [1] extended this calculation to the non-adiabatic case. By exactly solving the time-dependent Hamiltonian one can obtain the total phase  $\gamma$  acquired by the wave function in the evolution from  $t = 0$  to  $t = T$ . At the same time the acquired dynamical phase in this period can be calculated as

$$\gamma_d = -\frac{1}{\hbar} \int_0^T \langle \psi(t) | \dot{\hat{H}}(t) | \psi(t) \rangle dt, \quad (1)$$

where  $\dot{\hat{H}}(t)$  is the time dependent Hamiltonian and  $|\psi(t)\rangle$  is the investigated wave function. Then the geometric phase acquired in this period is

$$\gamma_g = \gamma - \gamma_d. \quad (2)$$

In the following sections we will describe several important phenomena related to the geometric phase in nanostructures.

## Geometric Phase Induced in Different Types of Nanostructures

In this section, we outline different groups of phenomena related to the geometric phase in nanostructures. In following sections, detailed mathematical and physical illustrations will be given for particular examples.

As mentioned above, there are three ingredients for the occurrence of the geometric phase in nanostructures: (i) the target system representing the central degrees of freedom under investigation; (ii) the environment which makes the Hamiltonian of the target system evolve periodically in a parameter space; (iii) the coupling or interaction between the target system and the environment. According to these, there are three major groups of the related phenomena: (i) the target is a spin system; (ii) the target is the orbital motion of electrons interacting with a field; (iii) the target is the orbital motion of electrons interacting with nuclei vibrations. We will discuss these phenomena in the following subsections.

### Spins as the Target System

A nanostructure consists of electrons and atomic nuclei. For both of them there are two types of degrees of freedom: the orbital (spatial) motions and the spins. Usually the spatial motion of electrons in outer levels of atoms is important in determining the properties of the system. In some cases, however, this spatial motion of electrons is restricted by some factors, such as the potential barriers in quantum dots, so that the spin degrees of freedom become more important. In such cases the spins of the whole or a part of the nanostructure may be regarded as the target system. Under special environment the spin system can exhibit phenomena related to the geometric phase.

### Spinning Quantum System in an External Magnetic Field

If a spin, say,  $s = 1/2$  is aligned to a magnetic field, and the field is made to rotate adiabatically, so that the spinor remains aligned to it, at the end of the loop the spinor acquires a complex phase factor  $e^{i\gamma_g}$ . For the general component of spin-1/2 spinor along the magnetic field, the Berry phase factor is precisely  $e^{-i/2\phi_s}$ , where  $\phi_s$  is the solid angle enclosed in the circuit followed by the magnetic field. Thus, the field rotating within a plane (thus enclosing one half of the complete solid angle  $4\pi$ ), the value of  $\gamma_g$  is  $\pi$ . This amounts to a “180 degrees” rotation in the Hilbert space of wave function. Note, however, that in the real three-dimensional space, the spin has actually rotated by 360 degrees. Technically, the factor of two distinguishing integer and half integer s-spinors is related to the 1:1 and 2:1 representations of the three-dimensional rota-

tion group, and equivalently to the 2:1 relation of SU(2) to SO(3) [31].

### Spinning Quantum System in the Presence of Spin Orbit Coupling

Besides the external magnetic field, the spin degree of freedom as a target system can also be controlled by the spin-orbit coupling of electrons moving in a quantum wire ring. In the presence of spin-orbit interaction there is a momentum-dependent effective magnetic field coupled to the electron spin. If the momentum traverses a circuit in the momentum space which can be viewed as the parameter space for the spin Hamiltonian, and if the spin orientation keeps in the direction of the effective magnetic field, the spin state acquires a Berry phase in its cyclic evolution. This Berry phase is just the spin-orbit Berry phase. During a cyclic motion of the electron along the ring, the spin of electron is precessing due to the spin-orbit coupling. It has been shown that in this case the Berry phase is  $-1/2$  of the solid angle subtended by the circuit traced on a sphere by the spin orientation precessing along the ring [4,28]. This investigation has been recently extended to the case including both the external magnetic field and the spin-orbit coupling in a ring [11].

### Electron Motion in Magnetic Field

In this subsection we discuss the situation where the target is the spatial motion of electrons and the environment is an external magnetic field. In this case the orbital motion should be important in determining the properties of the system. So there should be some extended dimensions in the nanostructure. The typical nanostructures possessing extended dimensions are the quantum wells having two extended dimensions and quantum wires having one extended dimension. In the following we separately discuss the geometric phase in these two types of nanostructures.

**Aharonov–Bohm Effect** The Aharonov–Bohm effect occurs in a ring made of quantum wire threaded by a magnetic flux. In this case a geometric phase is acquired by electrons during its cyclic motion around the ring due to differences in local values of the magnetic vector potential. As one must have guessed, however, in most cases the Berry phase vanishes. As one might expect, usually coming back after a loop to the starting point, makes nothing important, and the vector quantities go back to exactly what they used to be before the loop, indicating that the space where those vectors are moved around is flat. So, nonzero phases are clues pointing to nontrivial topological properties underlying the relation between the vector quantities and the adiabatic parameters. In the Aharonov–

Bohm effect the flatness of the space is violated at the point where the flux penetrates. This leads to the geometric phase which depends on the flux. So the topological phase of Aharonov and Bohm can be viewed as a special case of the Berry geometric phase [8].

**Quantum Hall Effect** If the Aharonov–Bohm ring is widened, the system is changed to a plane. The quantum Hall effect, a striking quantum phenomenon, occurs in two-dimensional electrons in strong magnetic field. Of course the quantum Hall effect is a very complicated phenomenon under extensive investigations. However, owing to the breaking of the flatness of space similar to the Aharonov–Bohm effect, some methods or concept developed in the investigation of the geometric phase can be adopted in the investigation of the quantum Hall effect. Especially, the quantization of the Hall conductance observed in experiments may be explained as a topological invariant represented by the Chern number which is related to the Berry curvature of the space.

### Orbital Motion of Electrons Interacting with Nuclei Vibrations

With the Born–Oppenheimer separation of electronic and vibrational motion, the Berry phase is zero on a physical path when the electronic ground state is well separated from the first electronic excitation. An exception is the so-called Jahn–Teller (JT) effect where the ground state is degenerate at some specially symmetric  $q$ -point in the parameter space. The distortion of the crystal reduces the symmetry and corresponds to a displacement in the parameter  $q$ -space from the symmetric point. To illustrate this idea we consider here the instructive example of the triangular cluster ( $E \times e$  JT model). The equilibrium geometry of a triangular cluster is not the equilateral triangle, but a distorted isosceles geometry. In the equilateral configuration, the electronic ground state is twofold degenerate ( $E$ ). These two electronic wave functions split for any displacement that destroys the equilateral geometry. Due to the symmetry, it can be seen that equivalent isosceles distortions where any two atoms get closer are energetically equivalent. Actually, it may be verified that for linear electron-distortion coupling the equilibrium configurations constitute a flat circular valley when the Born–Oppenheimer potential energy is calculated against the  $q$  coordinates. The two Born–Oppenheimer potential surfaces are degenerate at the central  $q = 0$  high-symmetry point. So, this “JT valley” encircles a degeneracy point for the electronic Hamiltonian. As a consequence a Berry phase should be present and physically relevant, as this

loop is frequently accessed by the low-energy dynamics of the  $q$ -distortions. The electronic wave functions undergoes a Berry phase change of  $\pi$  when the distortions loop around the circular valley. In the adiabatic approximation we choose carefully the phase at each point in order to simulate the “parallel transport”, i. e., to have positive overlap to the immediately preceding configuration. When a loop along the JT valley is completed, the electronic state changes sign. Another loop is needed to restore the original sign of the electron wave function.

### Berry Phase: Adiabatic Approximation

In this section, we present the detailed calculation of the Berry phase of electronic wave functions due to the electron-lattice interaction in graphene as a concrete illustration of the adiabatic approximation.

Due to the advances of material science, the graphene, as a real two-dimensional (2D) system with only one layer of carbon atoms, has been fabricated recently. In spite of the simplest lattice structure, the graphene exhibits striking properties which attracted much attention of both experimentalists and theorists [2,3,6,7,19,21,22,23, 24,25,26,27,35,36]. Particularly, the Dirac dispersion relation of electrons in graphene and the Fermi level near the Dirac point lead to specific features different from those in usual metals and semiconductors. Its transport properties are very special. There exists a universal maximal resistivity for all samples with the Fermi level near the Dirac point, independent of their shapes and mobility. This behavior is difficult to be explained from quantum transport theory, as such a minimal conductivity can not be supported in 2D with limited number of tunneling channels. Moreover, it is found that the weak localization is strongly suppressed [22]. This suppression could be attributed to a dephasing effect similar to that of a random magnetic field causing phase uncertainty. Thus, the transport in graphene is more likely to be diffusive or classical-like rather than quantum. There are several theoretical studies using different methods addressing the unusual transport properties in graphene [2,3,19,21,23,24,27,36].

In [37] we discuss one of the possible mechanisms which may cause the phase uncertainty in fermion states of graphene: the geometric phase produced by periodic relative vibrations between two sublattices. By deriving the time-dependent fermion Hamiltonian and exactly solving it with the Floquet scheme, one shows that in an evolution period of the vibration the geometric phases acquired by the fermions can be expressed by the adiabatic theory only in the momentum region where the level splitting of fermions in the evolution path is much larger than  $\hbar\omega$  with

$\omega$  being the vibration frequency. Outside this region the adiabatic condition is not satisfied. We show that the geometric phase, the fermion average energy, and the energy uncertainty in the non-adiabatic region are essentially different from those in the adiabatic region. Since geometric phases can be easily produced by the fermion-lattice interaction during lattice vibrations and they exhibit random or chaotic behavior in the non-adiabatic region, the resultant phase uncertainty may be a possible dephasing mechanism in such a 2D material.

The band structure in graphene can be described by a tight-binding Hamiltonian with one  $\pi$  orbital per site on the 2D honeycomb lattice [32]:

$$H = \sum_{\langle nn' \rangle} t_{nn'} (a_n^\dagger a_{n'} + a_{n'}^\dagger a_n), \quad (3)$$

where  $a_n^\dagger (a_n)$  creates (annihilates) an electron at site  $n$ ,  $\langle \dots \rangle$  denotes the nearest-neighbor (NN) sites, and  $t_{nn'}$  is the NN hopping. Here the spin indices are not explicitly included. On a honeycomb lattice there are two sublattices, labeled as A and B, and the cells, each of which consists of two atoms, compose a triangle lattice. On a perfect lattice, all the NN hopping integrals are the same,  $t_{nn'} = t_0$ . When there is a displacement  $\mathbf{d}$  between two sublattices, the lengths of three NN bonds connected to a site of sublattice A become

$$l^{(j)} = \sqrt{[d \sin(\theta - \alpha^{(j)})]^2 + [l_0 - d \cos(\theta - \alpha^{(j)})]^2}, \quad (4)$$

where  $j = 1, 2, 3$  labels the three NN bonds,  $\theta$  is the angle of  $\mathbf{d}$  related to the  $x$ -axis,  $l_0$  is the original bond length, and  $\alpha^{(j)} = 0, 2\pi/3, -2\pi/3$  for  $j = 1, 2, 3$ , respectively, is the original azimuth angle of the  $j$ th bond. For a small displacement, we can keep only the first-order terms of  $d$  and obtain  $l^{(j)} \sim l_0 - d \cos(\theta - \alpha^{(j)})$ . Then the hopping integrals as functions of  $\mathbf{d}$  can be written as

$$t^{(j)}(\mathbf{d}) = t_0 + \lambda d \cos(\theta - \alpha^{(j)}), \quad (5)$$

where  $\lambda$  is a coefficient describing the linear dependence of  $t^{(j)}(\mathbf{d})$  on  $l^{(j)}$ .

If one uses the following Bloch transformation for electron operator:

$$a_n = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}_n} a(b)_{\mathbf{k}}, \quad \text{for } n \in \text{sublattice A(B)}, \quad (6)$$

where  $a(b)_{\mathbf{k}}$  is annihilation operator of electron on sublattice A(B) with Bloch wavevector  $\mathbf{k}$ ,  $\mathbf{r}_n$  is the position

of site  $n$ , and  $N$  is the total number of cells, the original Hamiltonian becomes

$$H_0 = \sum_{\mathbf{k}} t_0 \left[ e^{ik_x l_0} + 2e^{-ik_x l_0/2} \cos\left(\frac{\sqrt{3}}{2} k_y l_0\right) \right] a_{\mathbf{k}}^\dagger b_{\mathbf{k}} \\ + \text{H.c.} \quad (7)$$

Expanding it to the first order of  $\mathbf{k}$  around two irreducible K points  $(0, \mp 4\pi/(3\sqrt{3}l_0))$  in the Brillouin zone, one has

$$H_0 = \frac{3}{2} \sum_{\mathbf{k}} [t_0 l_0 k_y \hat{\tau}_z \otimes \hat{\sigma}_x - t_0 l_0 k_x \hat{\mathbf{1}} \otimes \hat{\sigma}_y], \quad (8)$$

where  $\hat{\mathbf{1}}$  and  $\hat{\tau}_z$  are unit and Pauli matrices acting on two valleys, and  $\hat{\sigma}_{x,y}$  are Pauli matrices on two sublattices.

At the same time, the interaction between electrons and lattice displacement  $\mathbf{d}$  can be described by

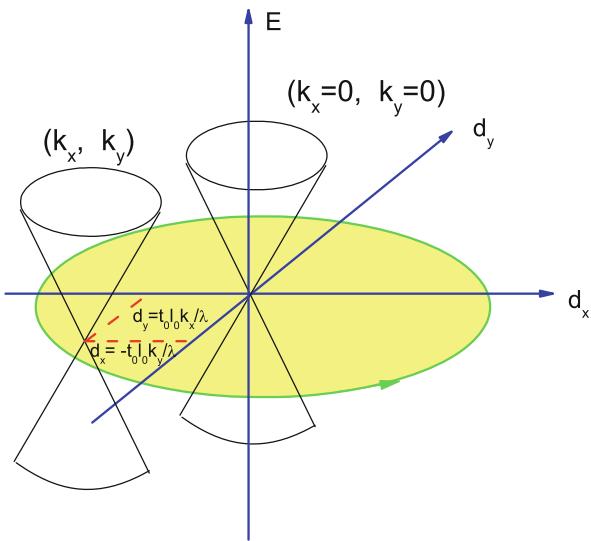
$$H_1 = \sum_{\langle nn' \rangle} (t_{nn'}(\mathbf{d}) - t_0) (a_n^\dagger a_{n'} + a_{n'}^\dagger a_n), \quad (9)$$

where  $t_{nn'}(\mathbf{d})$  is the hopping integral between sites  $n$  and  $n'$  under displacement  $\mathbf{d}$ . From Eqs. (5) and (6) and keeping only the terms of the first order of  $\mathbf{d}$ , the interaction Hamiltonian becomes

$$H_1 = \frac{3\lambda}{2} \sum_{\mathbf{k}} (d_x \hat{\mathbf{1}} \otimes \hat{\sigma}_x + d_y \hat{\tau}_z \otimes \hat{\sigma}_y). \quad (10)$$

In this paper we only consider vibrations at the long wavelength limit, i. e., the relative displacement  $\mathbf{d}$  between two sublattices is uniform over the whole system. In this case the vibrations only mix fermion states of particle and hole branches with the same momentum  $\mathbf{k}$  and in the same valley. As we will see below, the evolutions of fermion states with different  $\mathbf{k}$  during the vibrations are drastically different if the processes are not adiabatic, although the interaction terms in  $H_1$  are independent of  $\mathbf{k}$ . The vibration modes with finite wavelengths can mix fermion states of two branches with different momenta. This may produce a more complicated level structure of interacting fermions, but, as we will see below, the main features in the non-adiabatic region are chaotic and may not be very sensitive to the details of levels.

We consider the in-plane vibrations of the relative coordinates as  $d_x = (\eta_x/\lambda) \cos(\omega t + \alpha_x)$  and  $d_y = (\eta_y/\lambda) \cos(\omega t + \alpha_y)$ , where  $\eta_{x(y)}/\lambda$  and  $\alpha_{x(y)}$  are amplitude and initial phase of the vibration in the  $x(y)$  direction, respectively. To be specified in this paper, we do not consider other modes of vibrations which may produce similar effects.



**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 1**  
 $E_{1,2}^{(+)}$  as functions of  $d$  for different values of  $k$ . Green circle shows the evolution path in the vibrations

The Hamiltonian for electrons becomes time-dependent:

$$H(t) = \frac{3}{2} \sum_{\mathbf{k}} [(t_0 l_0 k_y \hat{\sigma}_z + \eta_x \cos(\omega t + \alpha_x) \hat{\sigma}_x) \otimes \hat{\sigma}_x - (t_0 l_0 k_x \hat{\sigma}_y - \eta_y \cos(\omega t + \alpha_y) \hat{\sigma}_z) \otimes \hat{\sigma}_y]. \quad (11)$$

There is no mixing between two valleys. For a given valley labeled by + or -, the Hamiltonian is

$$H_{\pm}(t) = \frac{3}{2} \sum_{\mathbf{k}} \{ t_0 l_0 (\pm k_y \hat{\sigma}_x - k_x \hat{\sigma}_y) + \eta_x \cos(\omega t + \alpha_x) \hat{\sigma}_x \pm \eta_y \cos(\omega t + \alpha_y) \hat{\sigma}_y \}. \quad (12)$$

At the adiabatic limit, the instantaneous eigen energies at time  $t$  can be obtained by diagonalizing  $H_{\pm}(t)$ :

$$E_m^{(\pm)}(t) = (-1)^m \frac{3}{2} |t_0 l_0 (\pm k_y + ik_x) + \eta_x \cos(\omega t + \alpha_x) + \mp i \eta_y \cos(\omega t + \alpha_y)|, \quad (13)$$

where  $m = 1$  and  $m = 2$  correspond to hole and particle branches, respectively. For given  $(k_x, k_y)$ , these two instantaneous eigen energies also exhibit conic dependence on parameters  $d_x$  and  $d_y$ , as shown in Fig. 1. It is interesting to note that the diabolical point, where the poles of two tapers coincide, is determined by ( $d_x = \mp t_0 k_y l_0 / \lambda$ ,  $d_y = \pm t_0 k_x l_0 / \lambda$ ), where plus and minus signs refer to two valleys. The vibrations of  $d_x$  and  $d_y$  result in circular motion with an elliptic track in the

$d_x - d_y$  plane. From the Berry theorem, a cyclic motion along a close track in a 2D parameter space can cause a Berry phase of  $\pm\pi$  in a wavefunction whose instantaneous eigen energy has a diabolical point enclosed in this circle [8]. Since the position of the diabolic point in the  $d_x - d_y$  plane is determined by  $(k_x, k_y)$ , the Berry phase acquired by the fermions with different momenta are different: It is  $\pm\pi$  for the fermions whose diabolic point ( $d_x = \mp t_0 k_y l_0 / \lambda$ ,  $d_y = \pm t_0 k_x l_0 / \lambda$ ) is enclosed in the track, while it is zero for the states with the diabolic point in the  $d_x - d_y$  parameter space outside the track.

### Theoretical Treatment Beyond Adiabatic Approximation

The above conclusion about the Berry phase, however, is valid only in the adiabatic condition, i.e., the energy difference between two branches in the track is always much larger than  $\hbar\omega$ . Owing to the existence of energy degeneracy, for any track and any finite frequency there always exist fermion states for which the adiabatic condition can not be satisfied. So we have to investigate the geometric phase from solutions of time-dependent Hamiltonian. Under this Hamiltonian, the wavefunctions of electrons can be expressed in the Floquet form [30]

$$\psi^{\pm}(l, \mathbf{k}, t) = e^{-i\frac{\epsilon^{\pm}(l, \mathbf{k})}{\hbar}t} u^{\pm}(l, \mathbf{k}, t), \quad (14)$$

where  $l$  is an index of electron states for given  $\mathbf{k}$ , and  $u^{\pm}(l, \mathbf{k}, t)$  is a periodic function of  $t$  with period  $T = 2\pi/\omega$ . Performing Fourier transformation for  $u^{\pm}(l, \mathbf{k}, t)$  with respect to  $t$ , one has

$$u^{\pm}(l, \mathbf{k}, t) = \sum_n \begin{pmatrix} A_n^{\pm}(l, \mathbf{k}) \\ B_n^{\pm}(l, \mathbf{k}) \end{pmatrix} e^{-in\omega t}, \quad (15)$$

where  $A_n^{\pm}$  and  $B_n^{\pm}$  are components on two sublattices. Substituting the Floquet wave function into the time-dependent Schrödinger equation [ $H_{\pm}(t) - i\hbar\partial_t$ ]  $\psi^{\pm}(l, \mathbf{k}, t) = 0$ , we obtain a set of linear homogeneous equations for the components [37]

$$\begin{aligned} & \frac{3t_0 l_0}{2} (\pm k_y + ik_x) B_n^{\pm} + \frac{3}{4} (e^{i\alpha_x} \eta_x \mp ie^{i\alpha_y} \eta_y) B_{n-1}^{\pm} \\ & + \frac{3}{4} (e^{-i\alpha_x} \eta_x \mp ie^{-i\alpha_y} \eta_y) B_{n+1}^{\pm} = (\epsilon^{\pm} + n\hbar\omega) A_n^{\pm}, \end{aligned} \quad (16)$$

$$\begin{aligned} & \frac{3t_0 l_0}{2} (\pm k_y - ik_x) A_n^{\pm} + \frac{3}{4} (e^{i\alpha_x} \eta_x \pm ie^{i\alpha_y} \eta_y) A_{n-1}^{\pm} \\ & + \frac{3}{4} (e^{-i\alpha_x} \eta_x \pm ie^{-i\alpha_y} \eta_y) A_{n+1}^{\pm} = (\epsilon^{\pm} + n\hbar\omega) B_n^{\pm}. \end{aligned} \quad (17)$$

From the requirement of the existence of nonzero solutions, for a given  $\mathbf{k}$ , one can solve the discrete quasienergies  $\epsilon^\pm(l, \mathbf{k})$  and the corresponding Floquet states  $\psi^\pm(l, \mathbf{k}, t)$ . It is noteworthy that two Floquet states whose quasienergies differ by  $n\hbar\omega$  with  $n$  being an integer are physically equivalent states [30]. So the quasienergies of physically different Floquet states can be reduced into region  $-\hbar\omega/2 \leq \epsilon^\pm \leq \hbar\omega/2$ .

In a period of evolution of Hamiltonian  $H^\pm(t)$ , the wavefunction  $\psi^\pm(l, \mathbf{k}, t = T) = e^{i\phi^\pm(l, \mathbf{k})} \psi^\pm(l, \mathbf{k}, t = 0)$  acquires a phase  $\phi^\pm(l, \mathbf{k})$ . It consists of two parts:

$$\phi^\pm(l, \mathbf{k}) = \alpha^\pm(l, \mathbf{k}) + \beta^\pm(l, \mathbf{k}), \quad (18)$$

where

$$\alpha^\pm(l, \mathbf{k}) = -\frac{\tilde{E}^\pm(l, \mathbf{k})T}{\hbar} \quad (19)$$

is the dynamical phase with

$$\tilde{E}^\pm(l, \mathbf{k}) = \frac{1}{T} \int_0^T dt \langle \psi^\pm(l, \mathbf{k}, t) | H_\pm(t) | \psi^\pm(l, \mathbf{k}, t) \rangle, \quad (20)$$

and

$$\beta^\pm(l, \mathbf{k}) = \phi^\pm(l, \mathbf{k}) - \alpha^\pm(l, \mathbf{k}) \quad (21)$$

is the geometrical phase [1]. Under the periodic vibrations the Floquet state  $\psi^\pm(l, \mathbf{k}, t)$  is stationary one which returns to the initial state other than phase  $\phi^\pm(l, \mathbf{k}) = -\epsilon^\pm(l, \mathbf{k})$  after a period of evolution. At the same time, the energy is no longer a good quantum number and the average energy of stationary state  $\psi^\pm(l, \mathbf{k}, t)$  can be calculated as

$$\begin{aligned} \tilde{E}^\pm(l, \mathbf{k}) &= \epsilon^\pm(l, \mathbf{k}) \\ &+ \hbar \sum_n n\omega \left( |A_n^\pm(l, \mathbf{k})|^2 + |B_n^\pm(l, \mathbf{k})|^2 \right). \end{aligned} \quad (22)$$

From this one obtains the geometric phase

$$\beta^\pm(l, \mathbf{k}) = 2\pi \sum_n n \left( |A_n^\pm(l, \mathbf{k})|^2 + |B_n^\pm(l, \mathbf{k})|^2 \right). \quad (23)$$

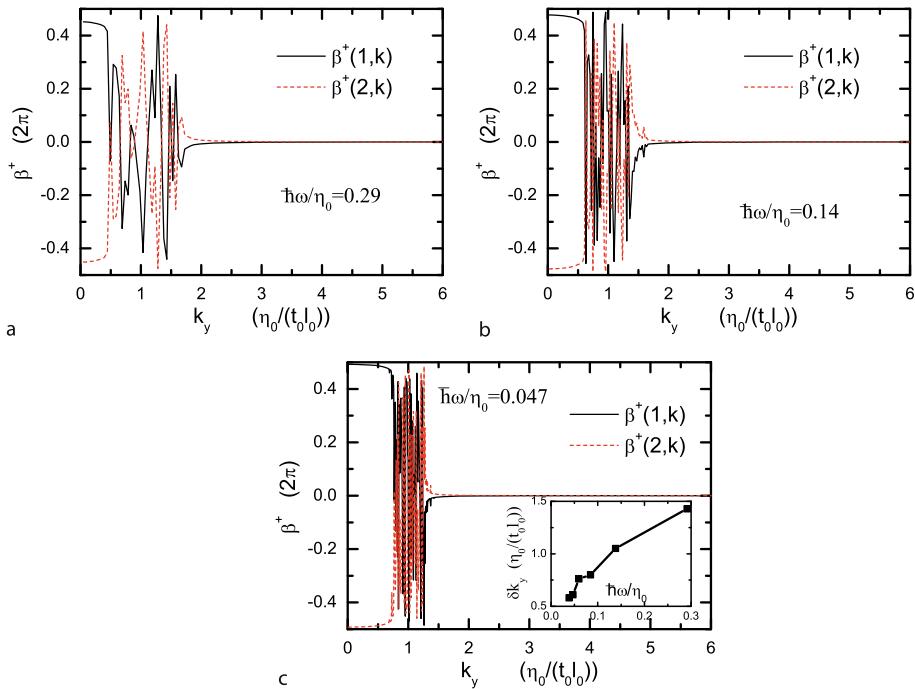
The energy uncertainty can be specified by the standard variance caused by the vibration

$$\begin{aligned} \Delta^\pm(l, \mathbf{k}) &\equiv \langle E^2 \rangle - \langle E \rangle^2 \\ &= \hbar^2 \omega^2 \left\{ \sum_n n^2 \left( |A_n^\pm(l, \mathbf{k})|^2 + |B_n^\pm(l, \mathbf{k})|^2 \right) \right. \\ &\quad \left. - \left[ \sum_n n \left( |A_n^\pm(l, \mathbf{k})|^2 + |B_n^\pm(l, \mathbf{k})|^2 \right) \right]^2 \right\}. \end{aligned} \quad (24)$$

From the superposition principle any time-dependent state can be expressed as a linear combination of the Floquet states. But this linear combination is usually not a stationary state, i.e., the state could not return to its initial one with only a phase difference after a period of evolution. For such states the geometric phase cannot be defined. So in this paper we only consider the geometric phase for the Floquet states. Since for given valley and given momentum there are only two unknowns having the same phonon number  $n$  in Eqs. (1) and (17), the number of quasienergies within range of  $[-\hbar\omega/2, \hbar\omega/2]$  is 2, corresponding to the particle and hole branches of the Dirac fermions with labels  $l = 1$  and  $l = 2$ , respectively.

Now we begin to investigate the properties of stationary states in the periodic vibrations. We are interested in: (i) the geometric phases acquired by fermion states with various momenta in a period of vibrations; (ii) the deviation of the average energy from the Dirac dispersion relation and the energy uncertainty caused by the vibration. From (i) we can shed some light on the phase variation of the Dirac fermions and the possible dephasing mechanism in graphene. From (ii) we can see the essential effect of vibrations on the basic dispersion relation.

In Fig. 2 we plot the geometric phase acquired by fermion states in a period of vibrations versus the fermion momentum. There is a  $\pi/2$  phase difference between the vibrations in the  $x$  and  $y$  directions and their amplitudes are the same, i.e.,  $\eta_x = \eta_y = \eta_0$ . So this is a circular vibration of the relative coordinates of two sublattices as shown in Fig. 1. According to the Berry theorem, the adiabatic Berry phase is  $\pm\pi$  if  $|\mathbf{k}| < (\eta_0)/(t_0 l_0)$  and is zero if  $|\mathbf{k}| > (\eta_0)/(t_0 l_0)$ . We note that from the calculations beyond the adiabatic approximation the obtained geometric phase is still roughly equal to the Berry phase: for  $|\mathbf{k}| \ll (\eta_0)/(t_0 l_0)$  it is nearly  $\pm\pi$  and for  $|\mathbf{k}| \gg (\eta_0)/(t_0 l_0)$  it is almost zero. However, now there appears a “chaotic” region around the border  $|\mathbf{k}| = (\eta_0)/(t_0 l_0)$  where the geometric phase randomly oscillates in changing the momentum. This is a direct consequence of the non-adiabaticity, as in this region the adiabatic condition, that the energy difference between two branches is much larger than  $\hbar\omega$  in the evolution path, is not satisfied. As can be seen from the inset of Fig. 2c, the width of this chaotic region is increased by increasing the frequency. The chaotic nature of the geometric phase reflects the phase uncertainty of the single-fermion states in this momentum region during vibrations. Such vibrations may exist in the graphene due to thermal excitations at a finite temperature, or due to the zero-point fluctuations even at very low temperatures. So such a fermion-lattice interaction can play roles of a dephasing mechanism for the single-fermion states. We also



**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 2**

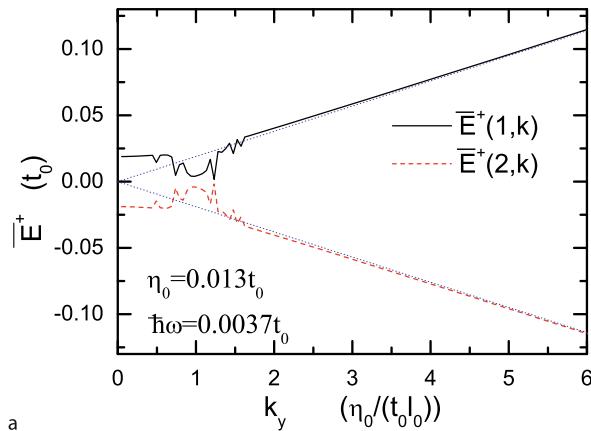
Geometric phase acquired in a period of vibrations by Dirac fermions in particle and hole branches of valley “+” as a function of the fermion momentum. The parameters for vibrations are:  $\eta_x = \eta_y = \eta_0$ ,  $\alpha_x = 0$ , and  $\alpha_y = -\pi/2$ .  $k_x = 0$  and units of  $k_y$  are set to be  $(\eta_0)/(t_0 l_0)$  so that the border between the  $\pm\pi$  Berry phase and the zero Berry phase is at  $k_y = 1$ . The inset of c shows the width of the chaotic region  $\delta k_y$  as a function of the frequency

note that in the whole region, even including the chaotic region, the geometric phases acquired by fermions of the same momentum in two branches are opposite to each other. This phase compensation effect of two branches implies that this dephasing mechanism has no effect if the fermions in two branches are paired.

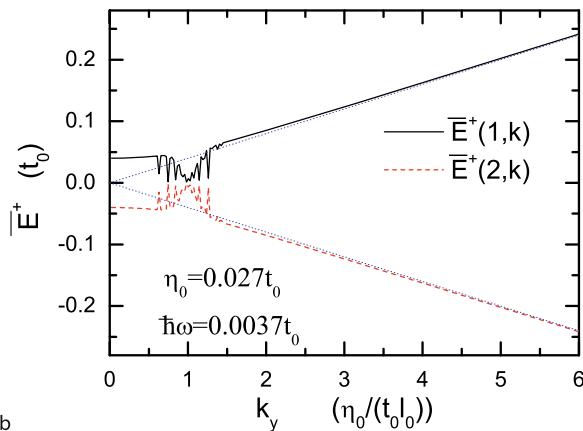
The above result indicates that the exchange of two branches leads to the sign reverse of the geometric phase. From the symmetry shown in Fig. 1 we can see that an equivalent operator which can also cause the sign reverse of the geometric phase is the reverse of the direction of the circular vibration from clockwise to counterclockwise or vice versa.

Another significant effect of vibrations is the uncertainty of fermion energies. As the single-particle energy is not a good quantum number, in Fig. 3 we plot the average energy of single-fermion states as a function of the momentum. From the comparison with the Dirac dispersion relation and with Fig. 2, the energy difference between two branches is enlarged near the Dirac point and for  $|k| \gg (\eta_0)/(t_0 l_0)$ , where the geometric phase is near its adiabatic value,  $\pm\pi$  or zero, while this energy difference shrinks in the chaotic region where the adiabatic approxi-

mation can not be used. Such opposite behaviors in these two regions imply that the failure of the adiabatic theory in the chaotic region has much more profound meaning than reflected from the values of geometric phase. Physically, in the chaotic region the oscillations of electrons between two branches are not able to follow the vibrations, this causes the loss of distinguishability of the two branches and the resultant states trend to take average energies in between, leading to a smaller energy spacing. On the contrary, in a quantum description the coupling of two states always enlarges their energy spacing. So one may expect that the chaotic region corresponds to a classical-like behavior and has a maximum uncertainty of single-fermion energies. To verify this we calculate the standard variance of Eq. (24) and show the results in Fig. 4. As expected, the standard variance exhibits a peak in the chaotic region, and the energy uncertainty increases by increasing the amplitudes of vibrations. Except for the peak in the chaotic region, the standard variance globally increases by increasing the momentum. Especially, the energy uncertainty becomes zero at the Dirac point for all the investigated vibration amplitudes. This originates from the fact that the energy uncertainty is due to the oscil-



a



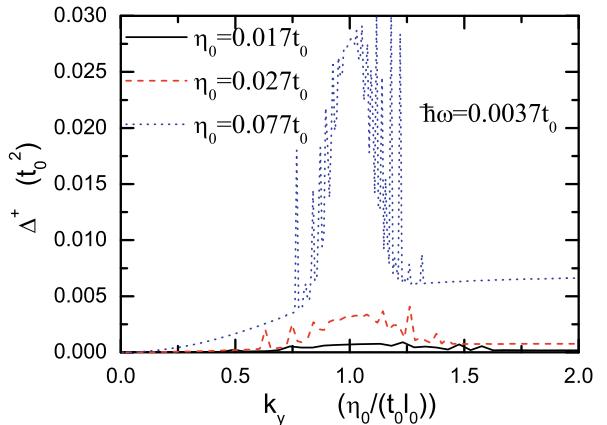
b

**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 3**

Average energy of fermions in particle and hole branches of valley “+” as a function of the fermion momentum. The parameters for vibrations are  $\eta_x = \eta_y = \eta_0$ ,  $\alpha_x = 0$ , and  $\alpha_y = -\pi/2$ .  $k_x = 0$ . The standard variance is the same for particle and hole branches

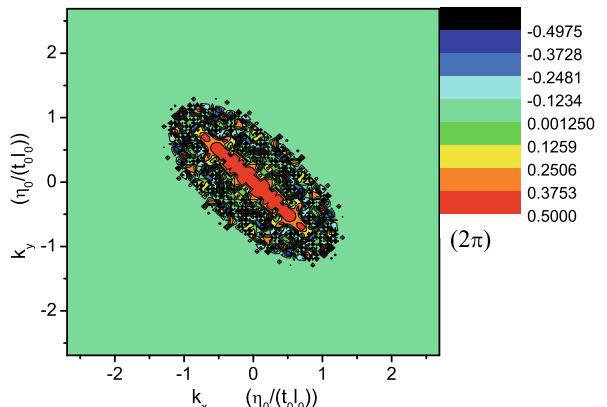
lution of the fermion states between two branches, but at the Dirac point the two branches coincide, leading to a zero amplitude of the oscillation. This structure, including the peak and the globally increasing background with the momentum, may catch major features of linewidth distribution obtained from the angle-resolved photoemission spectroscopy in graphene [9].

The evolution path plays a crucial role on the distribution of adiabatic and chaotic regions in the momentum space. To show this in Fig. 5 we plot the geometric phase as a function of  $k_x$  and  $k_y$  for an elliptic path. The chaotic region forms an orbicular area with a nearly fixed width along a loop defined by  $k_x = (\lambda d_y(t))/(t_0 l_0)$  and  $k_y = -(\lambda d_x(t))/(t_0 l_0)$ . The width depends on  $\omega$  as shown in the inset of Fig. 2c. As a result, the adiabatic region with



**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 4**

Standard variance of single-fermion energy of valley “+” as a function of the fermion momentum. The parameters for vibrations are  $\eta_x = \eta_y = \eta_0$ ,  $\alpha_x = 0$ , and  $\alpha_y = -\pi/2$ .  $k_x = 0$ . The standard variance is the same for particle and hole branches



**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 5**

Contour plot of geometric phase acquired by fermions in valley “+” and in the particle branch as a function of  $k_x$  and  $k_y$ . The parameters for vibrations are  $\eta_x = \eta_y = \eta_0 = 0.078t_0$ ,  $\hbar\omega = 0.0037t_0$ ,  $\alpha_x = 0$ , and  $\alpha_y = 0.6$

a  $\pm\pi$  Berry phase is compressed by reducing the short axis of the ellipse, but the chaotic region with random geometric phases can exist even for a linear vibration.

### Effect of Phase Uncertainty

The main features of the band structure in the graphene can be well described by a tight-binding Hamiltonian with one  $\pi$  orbital per site on the honeycomb lattice [32]:

$$H = \sum_n \epsilon_n a_n^\dagger a_n \sum_{\langle n, n' \rangle} + t_{nn'} (a_n^\dagger a_{n'} + a_{n'}^\dagger a_n), \quad (25)$$

where  $a_n^\dagger(a_n)$  creates (annihilates) an electron at site  $n$ ,  $\epsilon_n$  is the energy level at site  $n$ ,  $\langle \dots \rangle$  denotes the nearest-neighbor (NN) sites, and  $t_{nn'}$  is hopping integral between NN sites  $n$  and  $n'$ . Here, the spin indices are not explicitly included. In the absence of disorder,  $\epsilon_n = 0$  for all  $n$  and  $t_{nn'} = t_0$  for all  $\langle nn' \rangle$ , this Hamiltonian leads to the four-component Dirac fermion dispersion relation near the Dirac point  $E = 0$ . The diagonal and off-diagonal disorder can be introduced by adopting random variables  $\epsilon_n$  and  $t_{nn'}$  satisfying distribution probabilities  $P_d(\epsilon_n)$  and  $P_o(t_{nn'})$ , respectively. These probability functions define the strengths and the types of disorder in the system. The diagonal disorder represents the potential fluctuations due to impurities or due to the randomness on the substrate surface. A most commonly used distribution probability for the diagonal disorder is the square function

$$P_d(\epsilon_n) = \begin{cases} \frac{1}{W}, & \text{for } -\frac{W}{2} \leq \epsilon_n \leq \frac{W}{2}, \\ 0, & \text{otherwise,} \end{cases} \quad (26)$$

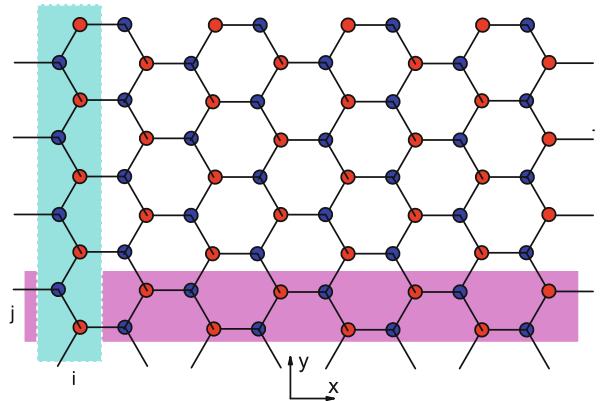
where  $W$  is the distribution width describing the strength of the disorder. For the OD disorder, there are different distribution functions describing different types of randomness. If we consider slight fluctuations of bond lengths around their average value due to lattice distortions, the following distribution is suitable

$$P_{o1}(t_{nn'}) = \begin{cases} \frac{1}{\lambda}, & \text{for } t_0 - \frac{\lambda}{2} \leq t_{nn'} \leq t_0 + \frac{\lambda}{2}, \\ 0, & \text{otherwise,} \end{cases} \quad (27)$$

where  $\lambda(<2|t_0|)$  is a measure of the bond-length fluctuations which make the NN hopping integrals randomly shift from  $t_0$ . In  $P_{o1}(t_{nn'})$  all the hopping integrals have the same sign as  $t_0$ , reflecting that only the changes of magnitudes of  $t_{nn'}$  are considered. In some cases, however,  $t_{nn'}$  not only fluctuate in magnitude, but also change in sign due to specific physical mechanisms. One of these possible mechanisms is the geometric phase which can easily produced for fermions in graphene since the Dirac point could play the roles of a degenerate point of two branches. We can assume that the opposite signs of  $t_{nn'}$  are also randomly distributed in the space. Then the distribution of  $t_{nn'}$  can be written as

$$P_o(t_{nn'}) = s P_{o1}(t_{nn'}) + (1-s) P_{o1}(-t_{nn'}), \quad (28)$$

where  $s(0 \leq s \leq 1)$  is a parameter describing how frequently the hopping integrals change sign in the space. When  $s = 0$  or  $s = 1$ , all the hopping integrals have the same sign and  $P_o(t_{nn'})$  reduces to  $P_{o1}(t_{nn'})$ . At  $s = 0.5$  the frequency of the sign changes is maximal. In  $P_o(t_{nn'})$  we



**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 6**

Structure of graphene lattice. Red and blue sites correspond to sublattices  $l = 1$  and  $l = -1$ ,  $i$  and  $j$  are indices of columns (light-cyan shaded) and rows (light-magenta shaded), respectively

include both the magnitude disorder and sign disorder of the off-diagonal elements in the Hamiltonian. Below we will set  $t_0$  as the energy units.

The wavefunction of a fermion can be written as

$$\psi = \sum_n c_n |n\rangle, \quad (29)$$

where  $c_n$  is the amplitude on site  $n$ . At a given energy  $E$  the amplitudes are given by the Schrödinger equation  $H\psi = E\psi$ . The propagation properties of a wavefunction on a finite lattice can be given by a relation between amplitudes on both sides. This is usually formulated by the transfer-matrix method or by the Green function. We note that the graphene lattice can be divided into two sub-lattices as distinguished by red and blue sites in Fig. 6. From the tight-binding nature of the Hamiltonian and the structure shown in Fig. 6, it can be seen that for given  $E$  all the amplitudes can be iteratively calculated along the  $x(y)$  direction if the amplitudes in the light-cyan shaded column (light-magenta shaded row) are given. From this we can establish a transfer matrix along the  $x$  or  $y$  direction for the graphene. For this purpose we identify a site  $n$  with three integers,  $n \equiv (i, j, l)$ , where  $i$  and  $j$  index the column and row, respectively, and  $l = \pm 1$  specifies the sublattice which the site  $n$  belongs to. In this notation the iterative relation for the amplitudes along the  $x$  direction can be derived from the Schrödinger equation as [38]

$$c_{i+1,j,-1} = \frac{(E - \epsilon_{i,j,1})c_{i,j,1} - t_{i,j,1;i,j,-1}c_{i,j,-1} - t_{i,j,1;i,j+(-1)^i,-1}c_{i,j+(-1)^i,-1}}{t_{i,j,1;i+1,j,-1}}, \quad (30)$$

$$c_{i+1,j,1} = \frac{(E - \epsilon_{i+1,j,-1})c_{i+1,j,-1} - t_{i+1,j,-1;i,j,1}c_{i,j,1} - t_{i+1,j,-1;i+1,j+(-1)^i,1}c_{i+1,j+(-1)^i,1}}{t_{i+1,j,-1;i+1,j,1}}. \quad (31)$$

For a system with finite width in the  $y$  direction, one can define vectors for the columns by  $\mathbf{u}_{i,l} \equiv (c_{i,1,l}, c_{i,2,l}, \dots, c_{i,M,l})^T$ , where  $M$  is the number of cells in a column. Then Eqs. (30) and (31) can be written in a matrix form:

$$\mathbf{u}_{i+1,-1} = \hat{T}_{1,i}\mathbf{u}_{i,1} + \hat{T}_{2,i}\mathbf{u}_{i,-1}, \quad (32)$$

$$\mathbf{u}_{i,1} = \hat{T}_{3,i}\mathbf{u}_{i+1,1} + \hat{T}_{4,i}\mathbf{u}_{i+1,-1}, \quad (33)$$

where the elements of the matrices are given by

$$\{\hat{T}_{1,i}\}_{j,j'} = \frac{E - \epsilon_{i,j,1}}{t_{i,j,1;i+1,j,-1}}\delta_{j,j'}, \quad (34)$$

$$\begin{aligned} \{\hat{T}_{2,i}\}_{j,j'} &= -\frac{t_{i,j,1;i,j,-1}}{t_{i,j,1;i+1,j,-1}}\delta_{j,j'} \\ &\quad - \frac{t_{i,j,1;i,j+(-1)^i,-1}}{t_{i,j,1;i+1,j,-1}}\delta_{j+(-1)^i,j'}, \end{aligned} \quad (35)$$

$$\begin{aligned} \{\hat{T}_{3,i}\}_{j,j'} &= -\frac{t_{i+1,j,-1;i+1,j,1}}{t_{i+1,j,-1;i,j,1}}\delta_{j,j'} \\ &\quad - \frac{t_{i+1,j,-1;i+1,j+(-1)^i,1}}{t_{i+1,j,-1;i,j,1}}\delta_{j+(-1)^i,j'}, \end{aligned} \quad (36)$$

$$\{\hat{T}_{4,i}\}_{j,j'} = \frac{E - \epsilon_{i+1,j,-1}}{t_{i+1,j,-1;i,j,1}}\delta_{j,j'}. \quad (37)$$

Here, one may use the periodic or open boundary condition at the ends  $j = 1$  and  $j = M$ . We have the iterative relation for the whole column

$$\begin{pmatrix} \mathbf{u}_{i+1,-1} \\ \mathbf{u}_{i+1,1} \end{pmatrix} = \hat{T}_i \begin{pmatrix} \mathbf{u}_{i,-1} \\ \mathbf{u}_{i,1} \end{pmatrix}, \quad (38)$$

where the transfer matrix is

$$\hat{T}_i = \begin{pmatrix} \hat{T}_{2,i}, & \hat{T}_{1,i} \\ -\hat{T}_{3,i}^{-1}\hat{T}_{4,i}\hat{T}_{2,i}, & \hat{T}_{3,i}^{-1} - \hat{T}_{3,i}^{-1}\hat{T}_{4,i}\hat{T}_{1,i} \end{pmatrix}. \quad (39)$$

If there are  $L$  columns in the  $x$ -direction, the total transfer matrix which gives the relation between amplitudes at two end columns is

$$\hat{T} = \prod_{i=1}^L \hat{T}_{L-i+1}. \quad (40)$$

If we want to investigate the propagation properties in the  $y$ -direction, we have to establish the transfer matrix for the rows. The relation between two adjacent rows is written as

$$\begin{pmatrix} \mathbf{v}_{j+1,-1} \\ \mathbf{v}_{j+1,1} \end{pmatrix} = \hat{V}_j \begin{pmatrix} \mathbf{v}_{j,-1} \\ \mathbf{v}_{j,1} \end{pmatrix}, \quad (41)$$

where  $\mathbf{v}_{j,l}$ 's are vectors denoting rows with length  $M$ ,  $\mathbf{v}_{j,l} \equiv (c_{1,j,l}, c_{2,j,l}, \dots, c_{M,j,l})^T$ , and the elements of the transfer matrix are

$$\begin{aligned} \{\hat{V}_j\}_{i,l;i',l'} &= \delta_{l,-(-1)^i} \left\{ C_{i,j}^{(1)}\delta_{-l,l'}\delta_{i,i'} - C_{i,j}^{(2)}\delta_{l,l'}\delta_{i,i'} \right. \\ &\quad - C_{i,j}^{(3)}\delta_{l,l'}\delta_{i+(-1)^i,i'} + C_{i,j}^{(4)} \left( C_{i,j}^{(1)}\delta_{l,l'}\delta_{i,i'} - C_{i,j}^{(2)}\delta_{l,-l'}\delta_{i,i'} \right. \\ &\quad - C_{i,j}^{(3)}\delta_{l,-l'}\delta_{i+(-1)^i,i'} \left. \right) - C^{(5)}\delta_{l,l'}\delta_{l,-(-1)^i}\delta_{i,i'} \\ &\quad - C^{(6)} \left( C_{i,-(-1)^i,j}^{(1)}\delta_{l,-l'}\delta_{i-(-1)^i,i'} \right. \\ &\quad - C_{i,-(-1)^i,j}^{(2)}\delta_{l,l'}\delta_{i-(-1)^i,i'} \\ &\quad \left. \left. - C_{i,-(-1)^i,j}^{(3)}\delta_{l,l'}\delta_{i-(-1)^i,2,i'} \right) \right\}, \end{aligned} \quad (42)$$

where

$$\begin{aligned} C_{i,j}^{(1)} &= \frac{E - \epsilon_{i,j,(-1)^i}}{t_{i,j,(-1)^i;i,j+1,-(-1)^i}}, \\ C_{i,j}^{(2)} &= \frac{t_{i,j,(-1)^i;i,j,-(-1)^i}}{t_{i,j,(-1)^i;i,j+1,-(-1)^i}}, \\ C_{i,j}^{(3)} &= \frac{t_{i,j,(-1)^i;i,j+(-1)^i,j,-(-1)^i}}{t_{i,j,(-1)^i;i,j+1,-(-1)^i}}, \\ C_{i,j}^{(4)} &= \frac{E - \epsilon_{i,j+1,-(-1)^i}}{t_{i,j+1,-(-1)^i;i,j+1,-(-1)^i}}, \\ C_{i,j}^{(5)} &= \frac{t_{i,j+1,-(-1)^i;i,j,(-1)^i}}{t_{i,j+1,-(-1)^i;i,j+1,-(-1)^i}}, \\ C_{i,j}^{(6)} &= \frac{t_{i,j+1,-(-1)^i;i,-(-1)^i,j+1,(-1)^i}}{t_{i,j+1,-(-1)^i;i,j+1,(-1)^i}}. \end{aligned}$$

The total transfer matrix for  $L$  rows is

$$\hat{V} = \prod_{i=1}^L \hat{V}_{L-i+1}. \quad (43)$$

For a finite  $M \times L$  system, there are  $M$  propagating channels along the length  $L$ . The Hermitian matrix,  $(\hat{T}^\dagger \hat{T})^{1/2L}$  or  $(\hat{V}^\dagger \hat{V})^{1/2L}$ , has  $2M$  eigen values. These  $2M$  eigen values should be positive and come in inverse pairs due to the unitarity of the matrix [5]. The  $M$  positive logarithms of the eigen values, denoted as  $\gamma_i$  in an ascending order with

index  $i$ , are Lyapunov exponents (LE) reflecting the exponential decay of the corresponding channels. Obviously, the logarithms of other  $M$  eigen values are the negatives of  $\gamma_i$ . The LEs will approach 0 when the disorder vanishes. The localization length is defined as the reciprocal of the smallest LE  $\xi \equiv 1/\gamma_1$ . The localization length plays a crucial role in the localization theory.

The zero-temperature conductance  $g$  of a rectangular  $M \times L$  system in units of  $2e^2/h$ , with two spins being taken into account, can be evaluated using the  $M$ -channel Landauer formula [13,14]

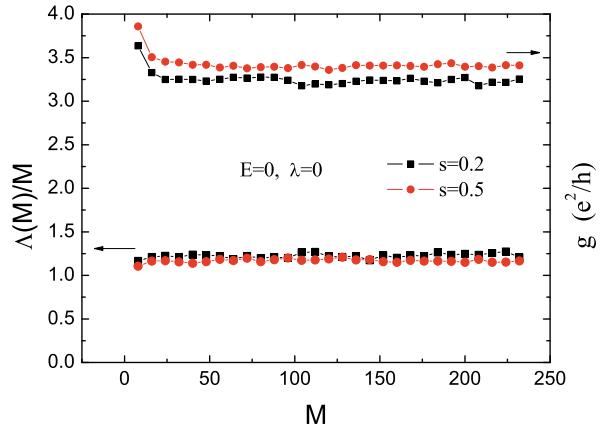
$$g = \text{Tr}(\hat{t}^\dagger \hat{t}), \quad (44)$$

where the  $M \times M$  transmission matrix  $\hat{t}$  describes the transmission of electrons from one lead to the other in the longitudinal direction. Here we assume that the leads are connected to all the  $M$  channels and they are perfect metal with band width much larger than that of the graphene. Formula (44) can be expressed with the LEs as [5]

$$g = \sum_{i=1}^M \frac{1}{\cosh^2(\gamma_i L)}. \quad (45)$$

Numerically, LEs can be calculated by using the standard method of Gram–Schmidt re-orthonormalization after each a few, say, ten steps of multiplication of the transfer matrices [20]. This is equivalent to the diagonalization of  $\hat{T}^\dagger \hat{T}$  or  $\hat{V}^\dagger \hat{V}$ , but avoids terrible overflow and loss of precision on the computer. On the other hand, although the calculations of the transfer matrices along the  $x$  and  $y$  directions are certainly different, there is no essential difference in the calculated results between these two directions. This is owing to the fact that the dispersion relation is conic around the Dirac point which is isotropic in the  $x$  and  $y$  directions. Below we will only present the results obtained for the transmission along the  $y$  direction.

The quantum transport processes could be drastically changed if the phases of wavefunctions become uncertain. This may happen due to the Berry phase which can be acquired by electrons during a cyclic evolution around a conic point. In the graphene the Berry phase can be easily created owing to the Dirac fermion dispersion relation. The attached Berry phase of  $\pi$  corresponds to sign changes of the wavefunction during the motion of electrons. If the sign changes randomly happen during the spatial motion, they can be modeled by a random distribution of positive and negative signs of hopping integrals. In probability  $P_o(t_{nn'})$  of Eq. (28), we include both the sign randomness and the magnitude randomness of the hopping integrals which are specified by parameters  $s$  and  $\lambda$ , respectively.

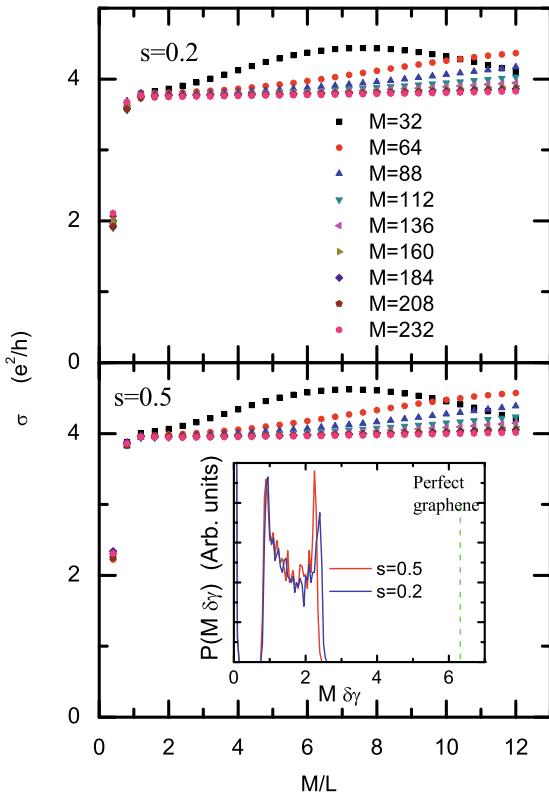


**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 7**

Scaling behavior of the rescaled localization length and the conductance of an  $M \times M$  system in the presence of sign randomness of hopping integrals obeying probability  $P_o(t_{nn'})$  with different values of  $s$

At first we set  $\lambda = 0$  and focus on the effect of sign randomness. In Fig. 7 we plot the scaling behavior of the rescaled localization length and the conductance of an  $M \times M$  system for different values of  $s$ . Similarly to the case of off-diagonal disorder, the states at the Dirac point in the presence of the sign randomness are also delocalized in the sense of finite-size scaling analysis. Effects of the sign randomness on the rescaled localization length and on the conductance are opposite:  $\Lambda(M)/M$  in  $s = 0.5$  is smaller than that in  $s = 0.2$ , but  $g$  in  $s = 0.5$  is larger than that in  $s = 0.2$ . This is strange because  $\Lambda(M)/M$  corresponds to the channel with the largest localization length which usually provides the leading contribution to the conductance. The only explanation for this is that the other channels also give significant contributions in the case of sign randomness. An expectable consequence of this is a new shape dependence of the conductivity, since the number of the tunneling channels is no longer restricted.

The shape dependence of the conductivity can be investigated by calculating the conductance of  $M \times L$  rectangular samples. The obtained conductivity as a function of the ratio  $M/L$  for various sizes is shown in Fig. 8. Except for small values of  $M/L$ , the conductivity has values near  $(4e^2)/h$ , independent of  $M/L$  and  $M$ , in consistence with the experimental findings. Surprisingly, this value is even larger than  $(4e^2)/(\pi h)$  obtained in the ballistic graphene [18] in spite of the disorder introduced by the sign randomness of hopping integrals. To get insight into this anomalous feature, we denote the average spacing between successive pairs of Lyapunov exponents as  $\delta\gamma$ ,



**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 8**

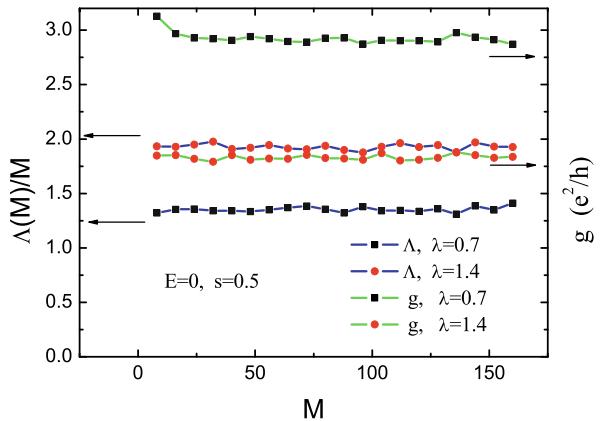
Calculated conductivity as a function of ratio  $M/L$  for rectangular sheets with different widths  $M$ . Inset in the lower panel: Distribution function  $P(M\delta\gamma)$  of  $M\delta\gamma$ . The green dashed line indicates the  $\delta$ -function distribution at  $2\pi$  in the case of perfect sheet

i.e.,  $\gamma_{2i+1} - \gamma_{2i-1} \sim \gamma_{2i+2} - \gamma_{2i} \sim \delta\gamma$ . Notice that in this case the Lyapunov exponents are grouped into pairs, corresponding to two equivalent valleys. By keeping only the leading exponential term of  $\cosh^2(\gamma_i L)$ , Eq. (45) can be approximately rewritten as

$$g \approx \sum_{i=1}^{M/2} 8e^{-2\gamma_{2i-1}L} \approx \frac{8e^{-2\gamma_1 L}(1 - e^{-(M+2)L\delta\gamma})}{1 - e^{-2L\delta\gamma}}, \quad (46)$$

in units of  $(2e^2)/h$ . In a perfect system of width  $M$  and using the periodic boundary condition, the spacing  $\delta\gamma$  for states at  $E = 0$  is  $\delta\gamma = 2\pi/M$  due to the Dirac fermion dispersion relation [18]. One may reasonably suppose that  $\delta\gamma \propto 1/M$  is still held in the case of disorder. Then, for  $M \gg L$ , the conductivity of an  $M \times L$  rectangular sample can be calculated as

$$\sigma = \frac{gL}{M} \approx \frac{4e^{-2\gamma_1 L}}{\delta\gamma M}. \quad (47)$$

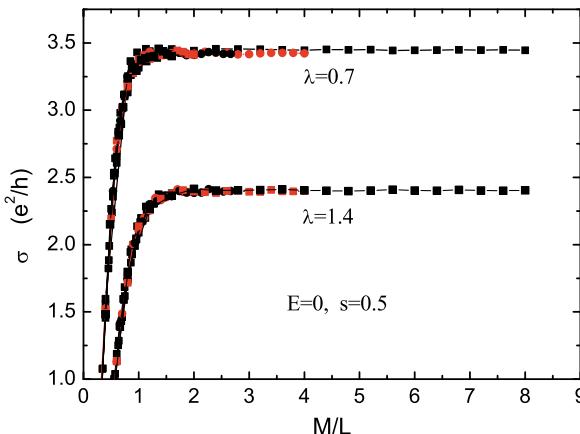


**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 9**

Scaling behavior of the rescaled localization length and the conductance of an  $M \times M$  system in the presence of both sign randomness and magnitude randomness of hopping integrals obeying probability  $P_0(t_{nn'})$

So for the perfect lattice  $\sigma = 2/\pi$  in units of  $(2e^2)/h$  [18]. In the case of sign randomness, the values of  $M\delta\gamma$  are in the range  $[0.9, 2.5]$  as shown in the inset of Fig. 8. Together with the factor  $e^{-2\gamma_1 L}$ , the obtained conductivity may be  $\pi$  times of that of the perfect lattice. Thus, the  $\pi$  factor difference in the conductivity is not trivial and reflects an essential change of the fermion properties due to the phase uncertainty.

It is interesting to investigate how the localization behavior changes in the presence of the magnitude randomness of hopping integrals. The scaling behaviors of the rescaled localization length and the conductance of an  $M \times M$  system with both nonzero  $s$  and  $\lambda$  are shown in Fig. 9. It can be seen that the states at  $E = 0$  are still delocalized and the opposite effects on the rescaled localization length and on the conductance are more evident than those shown in Fig. 7. In Fig. 10 we display the shape dependence of the conductivity of the  $M \times L$  rectangular samples for  $s = 0.5$  and different nonzero values of  $\lambda$ . Similarly to the case of  $\lambda = 0$ , the conductivity is still shape independent except for small values of  $M/L$ . This means that the sign-randomness induced shape independence of the conductivity is robust against the warping or corrugation disorder which may cause the randomness of magnitudes of hopping integrals. It can also be seen that the value of the conductivity is slightly reduced from  $(4e^2)/h$  by increasing the warping or corrugation disorder. For reasonable strength of the warping disorder ( $\lambda < 1$ ), however, the conductivity is above  $(3e^2)/h$ , in the range of measured values in experiments.



**Geometric Phase and Related Phenomena in Quantum Nanosystems, Figure 10**

Calculated conductivity as a function of ratio  $W/L$  for rectangular sheets with different widths  $M$  in the presence of both sign randomness and magnitude randomness of hopping integrals

### Future Directions

In physical applications, one often encounters Hamiltonian operators depending on several external parameters, thus potentially affected by Berry phases [38]. The Berry phase is observed by following a nontrivial loop in the parameters space, and performing some kind of interference between the state prior and after the loop. Standard examples are the parallel transport of a vector on a sphere, the Aharonov–Bohm effect, and analogous neutron interference experiments;  $q$  indicates some internal degrees of freedom of a system are approximately treated as “slow”, classical variables, affecting adiabatically the “fast” quantum dynamics of the other degrees of freedom. Of course, eventually the true quantum nature of the slow variables must be taken into account, as a further crucial step. The standard example for this is the Born–Oppenheimer separation of the “fast” quantum motion of electrons from the “slow” motion of nuclei in molecules and nanocrystals: the nuclear coordinates appear as external parameters defining the potential in which the electrons move. In turn, the electronic ground state energy enters in the potential energy for the adiabatic motion of the nuclei. When the latter is quantized, it gives rise to the vibrational/phononic states observed by spectroscopy. Here we shall be mostly concerned with the latter case. Berry phase for real Hamiltonians A physically relevant and rather simple case is that of the eigensystem of a real Hamiltonian operator  $H$  depending on a set of external parameters  $q$ . The eigenstates of a real operator may always be chosen to be real, at any  $q$ -point. It is straightforward to check that continuous real eigenstates realize parallel transport.

So there are following future directions of investigation on the geometric phase in nanostructures:

- (i) The transition from quantum to classical. It is well known that the phases of wave functions are essential for the manifestation of the quantum natures in nanostructures. If the fast degrees of freedom acquire geometric phases from random evolutions of slow degrees of freedom, i.e., the thermally excited vibrations of nuclei, the phase of the fast degrees of freedom becomes uncertain, leading to the transition from the quantum to classical.
- (ii) Investigation on the dephasing mechanism in nanostructures. Nanostructures are regarded as potential candidates for the quantum computing and quantum information processing. A prerequisite for such applications is the quantum coherence time as long as possible. Various dephasing mechanisms may prevent these applications as they destroy the quantum coherence of the target system. As one of the dephasing mechanisms the investigation of geometric phase in nanostructures is of particular importance.
- (iii) Designing of novel quantum interference devices employing the geometric phase in nanostructures. The geometric phase crucially depends on the geometry of the nanostructure. This provides a new way in designing interference devices by adjusting the shapes of nanostructures.
- (iv) Investigation of geometric phase in quantum many-body low-dimensional and nano systems. A many-body system is a complicated mixture of various degrees of freedom. Among them there may be some faster degrees of freedom which are crucially affected by geometric phases created from the slower degrees of freedom. Compared with the one-body nature of the fast degrees of freedom previously investigated, the geometric phases in a many-body fast system are much more complicated and need future investigations.

### Acknowledgments

This work was supported by the State Key Programs for Basic Research of China (2005CB623605 and 2006CB921803), and by National Foundation of Natural Science in China Grant Nos. 10474033 and 60676056.

### Bibliography

#### Primary Literature

1. Aharonov Y, Anandan A (1987) Phase change during a cyclic quantum evolution. *Phys Rev Lett* 58:1593–1596

2. Aleiner IL, Efetov KB (2006) Effect of disorder on transport in graphene. *Phys Rev Lett* 97:236801
3. Altland A (2006) Low-energy theory of disordered graphene. *Phys Rev Lett* 97:236802
4. Aronov AG, Lyanda-Geller YB (1993) Spin-orbit Berry phase in conducting rings. *Phys Rev Lett* 70:343–346
5. Beenakker CWJ (1997) Random-matrix theory of quantum transport. *Rev Mod Phys* 69:731–808
6. Berger C, Song Z, Li T, Li X, Ogbazghi AY, Feng R, Dai Z, Marchenkov AN, Conrad EH, First PN, de Heer WA (2004) Ultrathin epitaxial graphite: 2D electron gas properties and a route toward graphene-based nanoelectronics. *J Phys Chem B* 108:19912–19916
7. Berger C, Song Z, Li X, Wu X, Brown N, Naud C, Mayou D, Li T, Hass J, Marchenkov AN, Conrad EH, First PN, de Heer WA (2006) Electronic confinement and coherence in patterned epitaxial graphene. *Science* 312:1191–1196
8. Berry MV (1984) Quantal phase factors accompanying adiabatic changes. *Proc R Soc Lond A* 392:45–57
9. Bostwick A, Ohta T, Seyller T, Horn K, Rotenberg E (2007) Quasiparticle dynamics in graphene. *Nat Phys* 3:36–40
10. Calvert P (1992) Strength in disunity. *Nature* 357:365–366
11. Capozza R, Giuliano D, Lucignano P, Tagliacozzo A (2005) Quantum interference of electrons in a ring: tuning of the geometrical phase. *Phys Rev Lett* 95:226803
12. Chang LL, Ploog K (1985) Molecular beam epitaxy and heterostructures. Martinus Nijhoff, Dordrecht
13. Economou EN, Soukoulis CM (1981) Static conductance and scaling theory of localization in one dimension. *Phys Rev Lett* 46:618–621
14. Fisher DS, Lee PA (1981) *Phys Rev B* 23:R6851–6854
15. Goldhaber-Gordon D, Shtrikman H, Mahalu D, Abusch-Magder D, Kastner MA (1998) Kondo effect in a single-electron transistor. *Nature (London)* 391:156–159, Cronenwett SM, Oosterkamp TH, Kouwenhoven LP (1998) A tunable Kondo effect in quantum dots, *Science* 281:540–544
16. Ham FS (1987) Berry's geometrical phase and the sequence of states in the Jahn-Teller effect. *Phys Rev Lett* 58:725–728
17. Iijima S (1991) Helical microtubules of graphitic carbon. *Nature* 354:56–58
18. Katsnelson MI (2006) *Eur Phys J B* 51:157–160
19. Khveshchenko DV (2006) Electron localization properties in graphene. *Phys Rev Lett* 97:036802
20. MacKinnon A, Kramer B (1981) One-parameter scaling of localization length and conductance in disordered systems. *Phys Rev Lett* 47:1546–1549
21. McCann E, Kechedzhi K, Fal'ko VI, Suzuura H, Ando T, Altshuler BL (2006) Weak-localization magnetoresistance and valley symmetry in graphene. *Phys Rev Lett* 97:146805
22. Morozov SV, Novoselov KS, Katsnelson MI, Schedin F, Ponomarenko LA, Jiang D, Geim AK (2006) Strong suppression of weak localization in graphene. *Phys Rev Lett* 97:016801
23. Morpurgo AF, Guinea F (2006) Intervalley scattering, long-range disorder, and effective time-reversal symmetry breaking in graphene. *Phys Rev Lett* 97:196804
24. Nomura K, MacDonald AH (2007) Quantum transport of massless Dirac fermions. *Phys Rev Lett* 98:076602
25. Novoselov KS et al (2005) Two-dimensional gas of massless Dirac fermions in graphene. *Nature (London)* 438:197–200
26. Novoselov KS, Geim AK, Morozov SV, Jiang D, Zhang Y, Dubonos SV, Grigorieva IV, Firsov AA (2004) Electric field effect in atomically thin carbon films. *Science* 306:666–669
27. Pereira VM, Guinea F, dos Santos JMBL, Peres NMR, Neto AHC (2006) Disorder induced localized states in graphene. *Phys Rev Lett* 96:036801
28. Qian TZ, Su ZB (1994) Spin-orbit interaction and Aharonov-Anandan phase in mesoscopic rings. *Phys Rev Lett* 72:2311–2314
29. Reed MA, Randall JN, Aggarwal RJ, Matyi RJ, Moore TM, Wetzel AE (1988) Observation of discrete electronic states in a zero-dimensional semiconductor nanostructure. *Phys Rev Lett* 60:535–537
30. Sambe H (1973) Steady states and quasienergies of a quantum-mechanical system in an oscillating field. *Phys Rev A* 7:2203–2213
31. Simon B (1983) Holonomy, the quantum adiabatic theorem, and Berry's phase. *Phys Rev Lett* 51:2167–2170
32. Slonczewski JC, Weiss PR (1958) Band structure of graphite. *Phys Rev* 109:272–279
33. Teroffe J, Tromp RM (1993) Shape transition in growth of strained islands: spontaneous formation of quantum wires. *Phys Rev Lett* 70:2782–2785
34. Yacoby A, Heiblum M, Mahalu D, Shtrikman H (1995) Coherence and phase sensitive measurements in a quantum dot. *Phys Rev Lett* 74:4047–4050
35. Zhang Y, Tan YW, Stormer HL, Kim P (2005) Experimental observation of the quantum Hall effect and Berry's phase in graphene. *Nature (London)* 438:201–204
36. Ziegler K (2006) Robust transport properties in graphene. *Phys Rev Lett* 97:266802
37. Xiong SJ, Xiong Y (2007) Vibration-induced non-adiabatic geometric phase and energy uncertainty of fermions in graphene. *Euro Phys Letters* 80(6),60008438
38. Xiong SJ, Xiong Y (2007) Anderson localization of electron states in graphene in different types of disorder. *Phys Rev B* 76, 214204

## Books and Reviews

- Bohm A et al (2003) The geometric phase in quantum systems: foundations, mathematical concepts, and applications in molecular and condensed matter physics. Springer, Berlin
- Mead CA (1992) The geometric phase in molecular systems. *Rev Mod Phys* 64:51–85
- Yarkony DR (1996) Diabolical conical intersections. *Rev Mod Phys* 68:985–1013

## Glasses and Aging, A Statistical Mechanics Perspective on

LUDOVIC BERTHIER<sup>1</sup>, GIULIO BIROLI<sup>2</sup>

<sup>1</sup> Laboratoire des Colloïdes, Verres et Nanomatériaux, Université Montpellier II and CNRS, Montpellier, France

<sup>2</sup> CEA, DSM, Institut de Physique Théorique, IPhT, CNRS, MPPU, URA2306, Saclay, Gif-sur-Yvette, France

## Article Outline

### Glossary

### Definition of the Subject

### Phenomenology

### Taxonomy of ‘Glasses’ in Science

### Numerical Simulations

### Dynamic Heterogeneity

### Some Theory and Models

### Aging and Off-equilibrium

### Future Directions

### Acknowledgments

### Bibliography

## Glossary

In this preliminary section, a few concise definitions of the most important concepts discussed in this article are given.

**Glass transition** For molecular liquids, the glass transition denotes a crossover from a viscous liquid to an amorphous solid. Experimentally, the crossover takes place at the glass temperature,  $T_g$ , conventionally defined as the temperature where the liquid’s viscosity reaches the arbitrary value of  $10^{12}$  Pas. The glass transition more generally applies to many different condensed matter systems where a crossover or, less frequently, a true phase transition, takes place between an ergodic phase and a frozen, amorphous glassy phase.

**Aging** In the glass phase, disordered materials are characterized by relaxation times that exceed common observation timescales, so that a material quenched in its glass phase never reaches equilibrium (neither a metastable equilibrium). It exhibits instead an aging behaviour during which its physical properties keep evolving with time.

**Dynamic heterogeneity** Relaxation spectra of dynamical observables, e.g. the dynamical structure factor, are very broad in supercooled liquids. This is associated to a spatial distribution of timescales: at any given time, different regions in the liquid relax at different rates. Since the supercooled liquid is ergodic, slow regions eventually become fast, and vice versa. Dynamic heterogeneity refers to the existence of these non-trivial spatio-temporal fluctuations in the local dynamical behaviour, a phenomenon observed in virtually all disordered systems with slow dynamics.

**Effective temperature** An aging material relaxes very slowly, trying (in vain) to reach its equilibrium state. During this process, the system probes states that do not correspond to thermodynamic equilibrium, so that its thermodynamic properties can not be rigorously defined. Any practical measurement of its temperature

becomes a frequency-dependent operation. A ‘slow’ thermometer tuned to the relaxation timescale of the aging system measures an effective temperature corresponding to the ratio between spontaneous fluctuations (correlation) and linear response (susceptibility). This corresponds to a generalized form of the fluctuation-dissipation theorem for off-equilibrium materials.

**Frustration** Impossibility of simultaneously minimizing all the interaction terms in the energy function of the system. Frustration might arise from quenched disorder (as in spin glass models), from competing interactions (as in geometrically frustrated magnets), or from competition between a ‘locally preferred order’, and global, e.g. geometric, constraints (as in hard spheres packing problems).

## Definition of the Subject

Glasses belong to a well-known state of matter: we easily design glasses with desired mechanical or optical properties on an industrial scale, they are widely present in our daily life. Yet, a deep microscopic understanding of the glassy state of matter remains a challenge for condensed matter physicists [6,67]. Glasses share similarities with crystalline solids (they are both mechanically rigid), but also with liquids (they both have similar disordered structures at the molecular level). It is mainly this mixed character that makes them fascinating even to non-scientists.

A glass can be obtained by cooling the temperature of a liquid below its glass temperature,  $T_g$ . The quench must be fast enough that the more standard first order phase transition towards the crystalline phase is avoided. The glass ‘transition’ is not a thermodynamic transition at all, since  $T_g$  is only empirically defined as the temperature below which the material has become too viscous to flow on a ‘reasonable’ timescale (and it is hard to define the word ‘reasonable’ in any reasonable manner). Therefore,  $T_g$  does not play a fundamental role, as a phase transition temperature would. It is simply the temperature below which the material looks solid. When quenched in the glass phase below  $T_g$ , liquids slowly evolve towards an equilibrium state they cannot reach on experimental timescales. Physical properties are then found to evolve slowly with time in far from equilibrium states, a process known as ‘aging’ [152].

Describing theoretically and quantifying experimentally the physical mechanisms responsible for the viscosity increase of liquids approaching the glass transition and for aging phenomena below the glass transition certainly stand as central open challenges in condensed mat-

ter physics. Since statistical mechanics aims at understanding the collective behaviour of large assemblies of interacting objects, it comes as no surprise that it is a central tool in that field. We shall therefore summarize the understanding gained from statistical mechanics perspectives into the problem of glasses and aging.

The subject has quite broad implications. A material is said to be ‘glassy’ when its typical relaxation timescale becomes of the order of, and often much larger than, the typical duration of an experiment or a numerical simulation. With this generic definition, a large number of systems can be considered as glassy materials [173]. One can be interested in the physics of liquids (window glasses are then the archetype), in ‘hard’ condensed matter (for instance type II superconductors in the presence of disorder such as high- $T_c$  superconducting materials), charge density waves or spin glasses, ‘soft’ condensed matter with numerous complex fluids such as colloidal assemblies, emulsions, foams, but also granular materials, proteins, etc. All these materials exhibit, in some part of their phase diagrams, some sort of glassy dynamics characterized by a very rich phenomenology with effects such as aging, hysteresis, creep, memory, effective temperatures, rejuvenation, dynamic heterogeneity, non-linear response, etc.

This long list explains why this area of research has received increasing attention from physicists in the last two decades. ‘Glassy’ topics now go much beyond the physics of simple liquids (glass transition physics) and models and concepts developed for one system often find applications elsewhere in physics, from algorithmics to biophysics [55]. Motivations to study glassy materials are numerous. Glassy materials are everywhere around us and therefore obviously attract interest beyond academic research. At the same time, the glass conundrum provides theoretical physicists with deep fundamental questions since classical tools are sometimes not sufficient to properly account for the glass state. Moreover, simulating in the computer the dynamics of microscopically realistic material on timescales that are experimentally relevant is not an easy task, even with modern computers.

Studies on glassy materials constitute an exciting research area where experiments, simulations and theoretical calculations can meet, where both applied and fundamental problems are considered. How can one observe, understand, and theoretically describe the rich phenomenology of glassy materials? What are the fundamental quantities and concepts that emerge from these descriptions?

The outline of the article is as follows. In Sect. “[Phenomenology](#)” the phenomenology of glass-forming liquids

is discussed. In Sect. “[Taxonomy of ‘Glasses’ in Science](#)” other type of glasses are described, in particular colloids and granular materials. It is then described how computer simulations have provided deep insights into the glass problem in Sect. “[Numerical Simulations](#)”. The issue of dynamic heterogeneity is tackled in Sect. “[Dynamic Heterogeneity](#)”. The main theoretical perspectives currently available in the field are then summarized in Sect. “[Some Theory and Models](#)”. Aging and off-equilibrium phenomena occupy Sect. “[Aging and Off-equilibrium](#)”. Finally, issues that seem important for future research are discussed in Sect. “[Future Directions](#)”.

## Phenomenology

### Basic Facts

A vast majority of liquids (molecular liquids, polymeric liquids, etc.) form a glass if cooled fast enough in order to avoid the crystallisation transition [6]. Typical values of cooling rate in laboratory experiments are 0.1–100 K/min. The metastable phase reached in this way is called ‘super-cooled phase’. In this regime the typical timescales increase in a dramatic way and they end up to be many orders of magnitudes larger than microscopic timescales at  $T_g$ , the glass transition temperature.

For example, around the melting temperature  $T_m$ , the typical timescale  $\tau_\alpha$  on which density fluctuations relax, is of the order of  $\sqrt{ma^2/K_B T}$ , which corresponds to few picoseconds ( $m$  is the molecular mass,  $T$  the temperature,  $K_B$  the Boltzmann constant and  $a$  a typical distance between molecules). At  $T_g$ , which as a rule of thumb is about  $2/3T_m$ , the typical timescale has become of the order of 100 s, i. e. 14 orders of magnitude larger! This phenomenon is accompanied by a concomitant increase of the shear viscosity  $\eta$ . This can be understood by a simple Maxwell model in which  $\eta$  and  $\tau_\alpha$  are related by  $\eta = G_\infty \tau_\alpha$ , where  $G_\infty$  is the instantaneous (elastic) shear modulus which does not vary considerably in the super-cooled regime. In fact, viscosities at the glass transition temperature are of the order of  $10^{12}$  Pas. In order to grasp how viscous this is, recall that the typical viscosity of water at ambient temperature is of the order of  $10^{-2}$  Pas. How long would one have to wait to drink a glass of water with a viscosity  $10^{14}$  times larger?

As a matter of fact, the temperature at which the liquid does not flow anymore and becomes an amorphous solid, called a ‘glass’, is protocol dependent. It depends on the cooling rate and on the patience of the people carrying out the experiment: solidity is a timescale dependent notion. Pragmatically,  $T_g$  is defined as the temperature at which the shear viscosity is equal to  $10^{13}$  Poise (also  $10^{12}$  Pas).

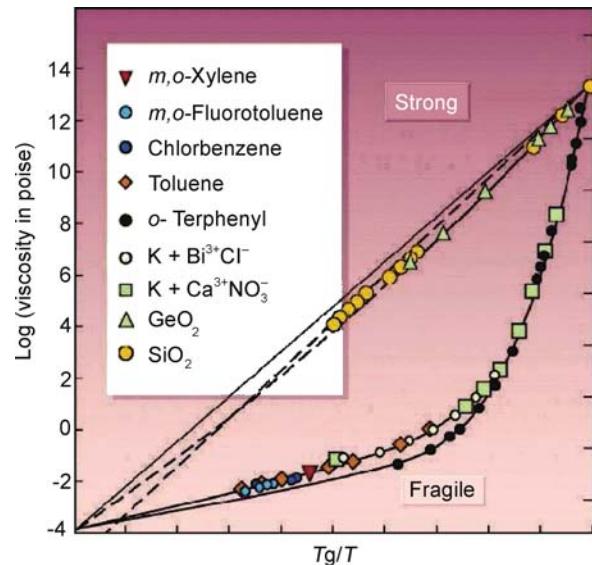
The increase of the relaxation timescale of supercooled liquids is remarkable not only because of the large number of decades involved but also because of its temperature dependence. This is vividly demonstrated by plotting the logarithm of the viscosity (or the relaxation time) as a function of  $T_g/T$ , as in Fig. 1. This is called the ‘Angell’ plot [6] and is very helpful in classifying supercooled liquids. A liquid is called strong or fragile depending on its position in the Angell plot. Straight lines correspond to ‘strong’ glass-formers and to an Arrhenius behaviour. In this case, one can extract from the plot an effective activation energy, suggesting quite a simple mechanism for relaxation by ‘breaking’ locally a chemical bond. The typical relaxation time is then dominated by the energy barrier to activate this process and, hence, has an Arrhenius behaviour. Window glasses fall in this category<sup>1</sup>. If one tries to define an effective activation energy for fragile glass-formers using the slope of the curve in Fig. 1, then one finds that this energy scale increases when the temperature decreases, a ‘super-Arrhenius’ behaviour. This increase of energy barriers immediately suggests that the glass formation is a collective phenomenon for fragile supercooled liquids. Support for this interpretation is provided by the fact that a good fit of the relaxation time or the viscosity is given by the Vogel–Fulcher–Tamman law (VFT):

$$\tau_\alpha = \tau_0 \exp \left[ \frac{DT_0}{(T - T_0)} \right], \quad (1)$$

which suggests a divergence of the relaxation time, and therefore a phase transition of some kind, at a finite temperature  $T_0$ . A smaller  $D$  in the VFT law corresponds to a more fragile glass. Note that there are other comparably good fits of these curves, such as the Bässler law [10],

$$\tau_\alpha = \tau_0 \exp \left( K \left( \frac{T_*}{T} \right)^2 \right),$$

that only lead to a divergence at zero temperature. Actually, although the relaxation time increases by 14 orders of magnitude, the increase of its logarithm, and therefore of the effective activation energy is very modest, and experimental data do not allow one to unambiguously determine the true underlying functional law without any reasonable doubt. For this and other reasons, physical interpretations



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 1

Arrhenius plot of the viscosity of several glass-forming liquids approaching the glass temperature  $T_g$  [67]. For ‘strong’ glasses, the viscosity increases in an Arrhenius manner as temperature is decreased,  $\log \eta \sim E/(K_B T)$ , where  $E$  is an activation energy and the plot is a straight line, as for silica. For ‘fragile’ liquids, the plot is bent and the effective activation energy increases when  $T$  is decreased towards  $T_g$ , as for ortho-terphenyl

in terms of a finite temperature phase transition must always be taken with a grain of salt.

However, there are other experimental facts that shed some light and reinforce this interpretation. Among them, is an empirical connection found between kinetic and thermodynamic behaviours. Consider the part of the entropy of the liquids,  $S_{\text{exc}}$ , which is in excess compared to the entropy of the corresponding crystal. Once this quantity, normalized by its value at the melting temperature, is plotted as a function of  $T$ , a remarkable connection with the dynamics emerges. As for the relaxation time one cannot follow this curve below  $T_g$  in thermal equilibrium. However, extrapolating the curve below  $T_g$  apparently indicates that the excess entropy vanishes at some finite temperature, called  $T_K$ , which is very close to zero for strong glasses and, generically, very close to  $T_0$ , the temperature at which a VFT fit diverges. This coincidence is quite remarkable: for materials with glass transition temperatures that vary from 50 K to 1000 K the ratio  $T_K/T_0$  remains close to 1, up to a few percents. Examples reported in [145] are provided in Table 1. The chosen subscript for  $T_K$  stands for Kauzmann [108] who recognized  $T_K$  as a very important temperature in the glass phase di-

<sup>1</sup>The terminology ‘strong’ and ‘fragile’ is not related to the mechanical properties of the glass but to the evolution of the short-range order close to  $T_g$ . Strong liquids, such as  $\text{SiO}_2$ , have a locally tetrahedral structure which persists both below and above the glass transition contrary to fragile liquids whose short-range amorphous structure disappears rapidly upon heating above  $T_g$ .

### Glasses and Aging, A Statistical Mechanics Perspective on, Table 1

Values of glass transition temperature, VFT singularity and Kauzmann temperatures for four supercooled liquids [145]

Substance	o-terphenyl	2-methyltetrahydrofuran	n-propanol	3-bromopentane
$T_g$	246	91	97	108
$T_0$	202.4	69.6	70.2	82.9
$T_K$	204.2	69.3	72.2	82.5
$T_K/T_0$	1.009	0.996	1.028	0.995

agram. Kauzmann further claimed that some change of behaviour (phase transition, crystal nucleation, etc.) must take place above  $T_K$ , because below  $T_K$  the entropy of the liquid, a disordered state of matter, becomes less than the entropy of the crystal, an ordered state of matter. This situation that seemed paradoxical at that time is not a serious problem. There is no general principle that would constraint the entropy of the liquid to be larger than that of the crystal. As a matter of fact, the crystallisation transition for hard spheres takes place precisely because the crystal becomes the state with the largest entropy at sufficiently high density [97].

On the other hand, the importance of  $T_K$  stands, partially because it is experimentally very close to  $T_0$ . Additionally, the quantity  $S_{\text{exc}}$  which vanishes at  $T_K$ , is thought to be a proxy for the so-called configurational entropy,  $S_c$ , which quantifies the number of metastable states. A popular physical picture due to Goldstein [91] is that close to  $T_g$  the system explores a part of the energy landscape (or configuration space) which is full of minima separated by barriers that increase when temperature decreases. The dynamic evolution in the energy landscape would then consist in a rather short equilibration inside the minima followed by ‘jumps’ between different minima. At  $T_g$  the barriers have become so large that the system remains trapped in one minimum, identified as one of the possible microscopic amorphous configurations of a glass. Following this interpretation, one can split the entropy into two parts. A first contribution is due to the fast relaxation inside one minimum, a second counts the number of metastable states,  $S_c = \log N_{\text{metastable}}$ , which is called the ‘configurational’ entropy. Assuming that the contribution to the entropy due to the ‘vibrations’ around an amorphous glass configuration is not very different from the entropy of the crystal, one finds that  $S_{\text{exc}} \approx S_c$ . In that case,  $T_K$  would correspond to a temperature at which the configurational entropy vanishes. This in turn would lead to a discontinuity (a downward jump) of the specific heat and would truly correspond to a thermodynamic phase transition.

### Static and Dynamic Correlation Functions

At this point the reader might have reached the conclusion that the glass transition may not be such a difficult problem: there are experimental indications of a diverging timescale and a concomitantly singularity in the thermodynamics. It simply remains to find static correlation functions displaying a diverging correlation length related to the emergence of ‘amorphous order’, which would indeed classify the glass transition as a standard second order phase transition. Remarkably, this remains an open and debated question despite several decades of research. Simple static correlation function are quite featureless in the supercooled regime, notwithstanding the dramatic changes in the dynamics. A simple static quantity is the structure factor defined by

$$S(q) = \left\langle \frac{1}{N} \delta\rho_q \delta\rho_{-q} \right\rangle ,$$

where the Fourier component of the density reads

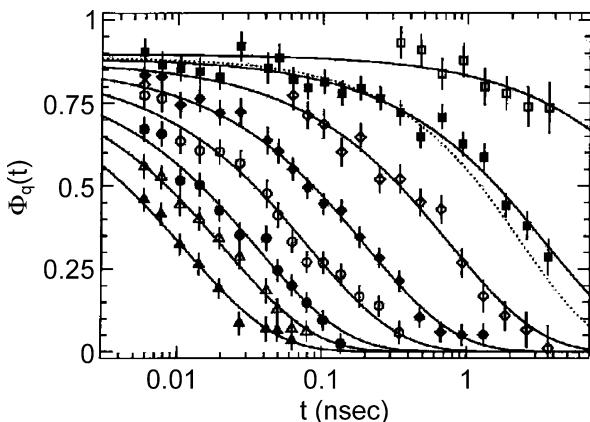
$$\delta\rho_q = \sum_{i=1}^N e^{i\mathbf{q}\cdot\mathbf{r}_i} - \frac{N}{V} \delta_{q,0} ,$$

with  $N$  is the number of particles,  $V$  the volume, and  $\mathbf{r}_i$  is the position of particle  $i$ . The structure factor measures the spatial correlations of particle positions, but it does not show any diverging peak in contrast to what happens, for example, at the liquid-gas tri-critical point where there is a divergence at small  $\mathbf{q}$ . More complicated static correlation functions have been studied [66], especially in numerical work, but until now there are no strong indications of a diverging, or at least substantially growing, static lengthscale [133]. A snapshot of a supercooled liquid configuration in fact just looks like a glass configuration, despite their widely different dynamic properties. What happens then at the glass transition? Is it a transition or simply a dynamic crossover? A more refined understanding can be gained studying dynamic correlations or response functions.

A dynamic observable studied in light and neutron scattering experiments is the intermediate scattering function,

$$F(\mathbf{q}, t) = \left\langle \frac{1}{N} \delta\rho_q(t) \delta\rho_{-q}(0) \right\rangle . \quad (2)$$

Different  $F(\mathbf{q}, t)$  measured by neutron scattering in supercooled glycerol [170] are shown for different temperatures in Fig. 2. These curves show a first, rather fast, relaxation to a plateau followed by a second, much slower, relaxation. The plateau is due to the fraction of density

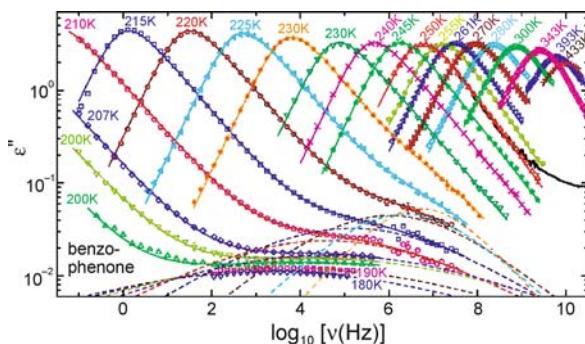


Glasses and Aging, A Statistical Mechanics Perspective on, Figure 2

Temperature evolution of the intermediate scattering function normalized by its value at time equal to zero for supercooled glycerol [170]. Temperatures decrease from 413 K to 270 K from left to right. The solid lines are fit with a stretched exponential with exponent  $\beta = 0.7$ . The dotted line represents another fit with  $\beta = 0.82$

fluctuations that are frozen on intermediate timescales, but eventually relax during the second relaxation. The latter is called ‘alpha-relaxation’, and corresponds to the structural relaxation of the liquid. This plateau is akin to the Edwards–Anderson order parameter,  $q_{EA}$ , defined for spin glasses which measures the fraction of frozen spin fluctuations [33]. Note that  $q_{EA}$  continuously increases from zero below the spin glass transition. Instead, for structural glasses, a finite plateau appears above any transition.

The intermediate scattering function can be probed only on a relatively small regime of temperatures. In order to track the dynamic slowing down from microscopic to macroscopic timescales, other correlators have been studied. A popular one is obtained from the dielectric susceptibility, which is related by the fluctuation-dissipation theorem to the time correlation of polarization fluctuations. It is generally admitted that different dynamic probes reveal similar temperature dependences for the relaxation time. The temperature evolution of the imaginary part of the dielectric susceptibility,  $\epsilon''(\omega)$ , is shown in Fig. 3 which covers a very wide temperature window [142]. At high temperature, a good representation of the data is given by a Debye law,  $\epsilon(\omega) = \epsilon(\infty) + \Delta\epsilon/(1 + i\omega\tau_\alpha)$ , which corresponds to an exponential relaxation in the time domain. When temperature is decreased, however, the relaxation spectra become very broad and strongly non-Debye. One particularly well-known feature of the spectra is that they are well fitted, in the time domain, for times corresponding to the alpha-



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 3

Temperature evolution of the dielectric susceptibility of the glass-former benzophenone measured over more than 10 decades of relaxation times [142]. Dynamics slows down dramatically as temperature is decreased and relaxation spectra become very broad at low temperature

relaxation with a stretched exponential,  $\exp(-(t/\tau_\alpha)^\beta)$ . In the Fourier domain, forms such as the Havriliak–Negami law are used,  $\epsilon(\omega) = \epsilon(\infty) + \Delta\epsilon/(1 + (i\omega\tau_\alpha)^\alpha)^\gamma$ , which generalizes the Debye law. The exponents  $\beta$ ,  $\alpha$  and  $\gamma$  depend in general on temperature and on the particular dynamic probe chosen, but they capture the fact that relaxation is increasingly non-exponential when  $T$  decreases towards  $T_g$ . A connection was empirically established between fragility and degree of non-exponentiality, more fragile liquids being characterized by broader relaxation spectra [67].

To sum up, there are many remarkable phenomena that take place when a supercooled liquid approaches the glass transition. Striking ones have been presented, but many others have been left out for lack of space [6,33,66,67]. We have discussed physical behaviours, relationships or empirical correlations observed in a broad class of materials. This is quite remarkable and suggests that there is some physics (and not only chemistry) to the problem of the glass transition, which we see as a collective (critical?) phenomenon which is relatively independent of microscopic details. This justifies our statistical mechanics perspective on this problem.

### Taxonomy of ‘Glasses’ in Science

We now introduce some other systems whose phenomenological behaviour is close or, at least, related, to the one of glass-forming liquids, showing that glassiness is truly ubiquitous. It does not only appear in many different physical situations but also in more abstract contexts, such as computer science.

## The Jamming Transition of Colloids and Grains

Colloidal suspensions consist of big particles suspended in a solvent [121]. The typical radii of the particles are in the range  $R = 1\text{--}500\text{ nm}$ . The solvent, which is at equilibrium at temperature  $T$ , renders the short-time dynamics of the particles Brownian. The microscopic timescale for this diffusion is given by  $\tau = R^2/D$  where  $D$  is the short-time self-diffusion coefficient. Typical values are of the order of  $\tau \sim 1\text{ ms}$ , and thus are much larger than the ones for molecular liquids (in the picosecond regime). The interaction potential between particles depends on the systems, and this large tunability makes colloids very attractive objects for technical applications. A particularly relevant case, on which we will focus in the following, is a purely hard sphere potential, which is zero when particles do not overlap and infinite otherwise. In this case the temperature becomes irrelevant, apart from a trivial rescaling of the microscopic timescale. Colloidal hard spheres systems have been intensively studied [121] in experiments, simulations and theory varying their density  $\rho$ , or their volume fraction  $\phi = 4/3\pi R^3 \rho$ . Hard spheres display a fluid phase from 0 to intermediate volume fractions, a freezing-crystallisation transition at  $\phi \simeq 0.494$ , and a melting transition at  $\phi \simeq 0.545$ . Above this latter value the system can be compressed until the close packing point  $\phi \simeq 0.74$ , which corresponds to the FCC crystal. Interestingly for our purposes, a small amount of polydispersity (particles with slightly different sizes) suppresses crystallization. In this case, the system can be more easily ‘supercompressed’ above the freezing transition without nucleating the crystal, at least on experimental timescales. In this regime the relaxation timescale increases very fast [144]. At a packing fraction  $\phi_g \simeq 0.58$  it becomes so large compared to typical experimental timescales that the system does not relax anymore: it is jammed. This ‘jamming transition’ is obviously reminiscent of the glass transition of molecular systems. In particular, the location  $\phi_g$  of the colloidal glass transition is as ill-defined as the glass temperature  $T_g$ .

Actually, the phenomena that take place increasing the volume fraction are analogous to the ones seen in molecular supercooled liquid: the relaxation timescales increases very fast and can be fitted [52] by a VFT law in density as in Eq. (1), dynamical correlation functions display a broad spectrum of timescales and develop a plateau, no static growing correlation length has been found, etc. Also the phenomenon of dynamic heterogeneity that will be addressed in Sect. “Dynamic Heterogeneity” is present in both cases [109,165]. However, it is important to underline a major difference: because the microscopic timescale for colloids is so large, experiments can only track the

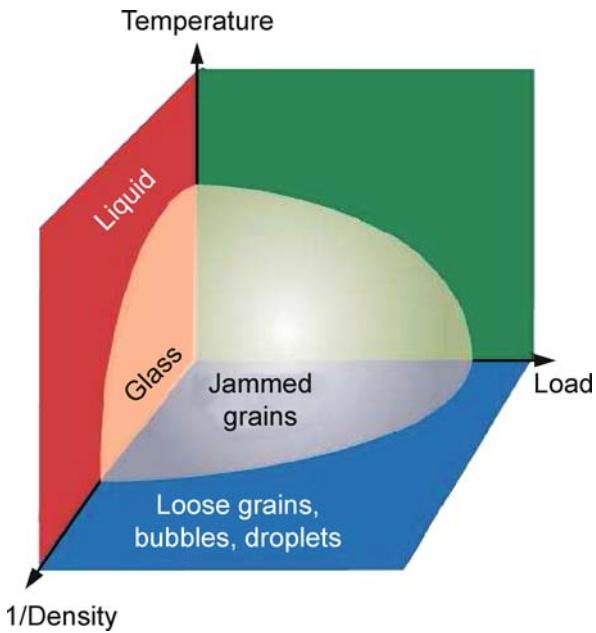
first 5 decades of slowing down. A major consequence is that the comparison between the glass and colloidal transitions must be performed by focusing in both cases on the first 5 decades of the slowing down, which corresponds to relatively high temperatures in molecular liquids. Understanding how much and to what extent the glassiness of colloidal suspensions is related to the one of molecular liquids remains an active domain of research.

Another class of systems that have recently been studied from the point of view of their glassiness is driven granular media. Grains are macroscopic objects and, as a consequence, do not have any thermal motion. A granular material is therefore frozen in a given configuration if no energy is injected into the system [104]. However, it can be forced in a steady state by an external drive, such as shearing or tapping. The dynamics in this steady state shows remarkable similarities (and differences) with simple fluids. The physics of granular materials is a very wide subject [104]. In the following we only address briefly what happens to a polydisperse granular fluid at very high packing fractions, close to its random close packed state. As for colloids, the timescales for relaxation or diffusion increase very fast when density is increased, without any noticeable change in structural properties. Again, it is now established [62,110,127] that many phenomenological properties of the glass and jamming transitions also occur in granular assemblies. As for colloids, going beyond the mere analogy and understanding how much these different physical systems are related is a very active domain of research.

This very question has been asked in a visual manner by Liu and Nagel [124] who rephrased it in a single picture, reproduced in Fig. 4. By building a common phase diagram for glasses, colloids and grains, they ask whether the glass and jamming transitions of molecular liquids, colloids and granular media are different facets of the same phase. In this unifying ‘phase diagram’, the ‘phase’ close to the origin is glassy and can be reached either by lowering the temperature as in molecular liquids, or increasing the packing fraction or decreasing the external drive in colloids and granular media. It remains to provide precise answers to this elegantly formulated, but rather broad, set of questions.

## Other ‘Glasses’ in Physics and Beyond

There are many other physical contexts in which glassiness plays an important role [173]. One of the most famous examples is the field of spin glasses. Real spin glasses are magnetic impurities interacting by quenched random couplings. At low temperatures, their dynamics be-



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 4

The ‘great unification’ phase diagram of jamming and glass transitions [124]. Glassy phases occur at low temperature, low external drive, or high density in different systems

come extremely slow and they freeze in amorphous spin configuration dubbed a ‘spin glass’ by P. W. Anderson. There are many other physical systems, often characterized by quenched disorder, that show glassy behaviour, like Coulomb glasses, Bose glasses, etc. In many cases, however, one does expect quite a different physics from structural glasses: the similarity between these systems is therefore only qualitative.

Finally, and quite remarkably, glassiness emerges even in other branches of science [55]. In particular, it has been discovered recently that concepts and techniques developed for glassy systems turn out to apply and be very useful tools in the field of computer science. Problems like combinatorial optimization display phenomena completely analogous to phase transitions, actually, to glassy phase transitions. *A posteriori*, this is quite natural, because a typical optimization problem consists in finding a solution in a presence of a large number of constraints. This can be defined, for instance, as a set of  $N$  Boolean variables that satisfies  $M$  constraints. For  $N$  and  $M$  very large at fixed  $\alpha = M/N$ , this problem very much resembles finding a ground state in a statistical mechanics problem with quenched disorder. Indeed one can define an energy function (a Hamiltonian) as the number of unsatisfied constraints, that has to be minimized, as in a  $T = 0$

statmech problem. The connection with glassy systems originates from the fact that in both cases the energy landscape is extremely complicated, full of minima and saddles. The fraction of constraints per degree of freedom,  $\alpha$ , plays a role similar to the density in a hard sphere system. A detailed presentation of the relationship between optimization problems and glassy systems is clearly out of the scope of the present review. We simply illustrate it pointing out that a central problem in optimization, random k-satisfiability, has been shown to undergo a glass transition when  $\alpha$  increases that is analogous to the one of structural glasses [117].

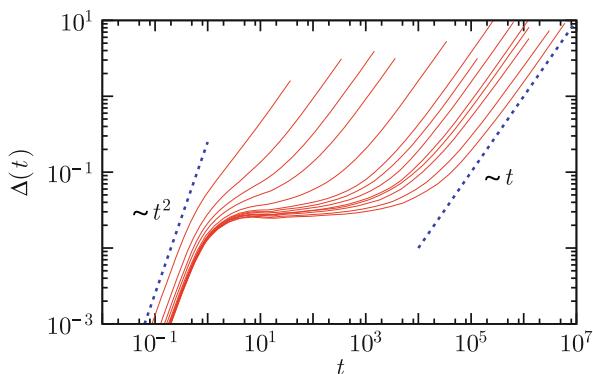
### Numerical Simulations

Studying the glass transition of molecular liquids at a microscopic level is in principle straightforward since one must answer a very simple question: how do particles move in a liquid close to  $T_g$ ? It is of course a daunting task to attempt answering this question experimentally because one should then resolve the dynamics of single molecules to be able to follow the trajectories of objects that are a few Angstroms large on timescales of tens or hundreds of seconds, which sounds like eternity when compared to typical molecular dynamics usually lying in the picosecond regime. In recent years, such direct experimental investigations have been started using time and space resolved techniques such as atomic force microscopy [161] or single molecule spectroscopy [3], but this remains a very difficult task.

In numerical simulations, by contrast, the trajectory of each particle in the system can, by construction, be followed at all times. This allows one to quantify easily single particle dynamics, as proved in Fig. 5 where the averaged mean-squared displacement  $\Delta(t)$  measured in a simple Lennard-Jones glass-former is shown. It is defined by

$$\Delta(t) = \left\langle \frac{1}{N} \sum_{i=1}^N |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \right\rangle,$$

where  $\mathbf{r}_i(t)$  represents the position of particle  $i$  at time  $t$  in a system composed of  $N$  particles; the brackets indicate an ensemble average. The particle displacements considerably slow down when  $T$  is decreased and the self-diffusion constant decreases by orders of magnitude, mirroring the behaviour of the viscosity shown in Fig. 1 for real systems. Moreover, a rich dynamics is observed, with a plateau regime at intermediate timescales, corresponding to an extended time window during which particles vibrate around their initial positions, exactly as in a crystalline solid. The difference with a crystal is of course that this localization is only transient, and all particles eventu-

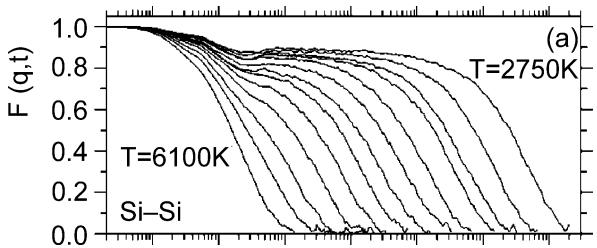


Glasses and Aging, A Statistical Mechanics Perspective on, Figure 5

Mean-squared displacements of individual particles in a simple model of a glass-forming liquid composed of Lennard-Jones particles observed on a wide time window. When temperature decreases (from left to right), the particle displacements become increasingly slow with several distinct time regimes corresponding, in this order, to ballistic, localized, and diffusive regimes

ally escape and diffuse at long times with a diffusion constant  $D_s$ , so that  $\Delta(t) \sim 6D_s t$  when  $t \rightarrow \infty$ .

In recent years, computer experiments have played an increasingly important role in glass transition studies. It could almost be said that particle trajectories in numerical work have been studied under so many different angles that probably very little remains to be learnt from such studies in the regime that is presently accessible using present day computers. Unfortunately, this does not imply complete knowledge of the physics of supercooled liquids. As shown in Fig. 5, it is presently possible to follow the dynamics of a simple glass-forming liquid over more than eight decades of time, and over a temperature window in which average relaxation timescales increase by more than five decades. This might sound impressive, but a quick look at Fig. 1 shows, however, that at the lowest temperatures studied in the computer, the relaxation timescales are still orders of magnitude faster than in experiments performed close to the glass transition temperature. They can be directly compared to experiments performed in this high temperature regime, but this also implies that simulations focus on a relaxation regime that is about eight to ten decades of times faster than in experiments performed close to  $T_g$ . Whether numerical works are useful to understand the glass transition itself at all is therefore an open, widely debated, question. We believe that it is now possible to numerically access temperatures which are low enough that many features associated to the glass transition physics can be observed: strong decoupling phenomena, clear deviations from fits to the mode-coupling the-



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 6

Intermediate scattering function at wavevector  $1.7 \text{ \AA}^{-1}$  for the Si particles at  $T = 2750 \text{ K}$  obtained from molecular dynamics simulations of a model for silica [98]

ory (which are experimentally known to hold only at high temperatures), and crossovers towards truly activated dynamics.

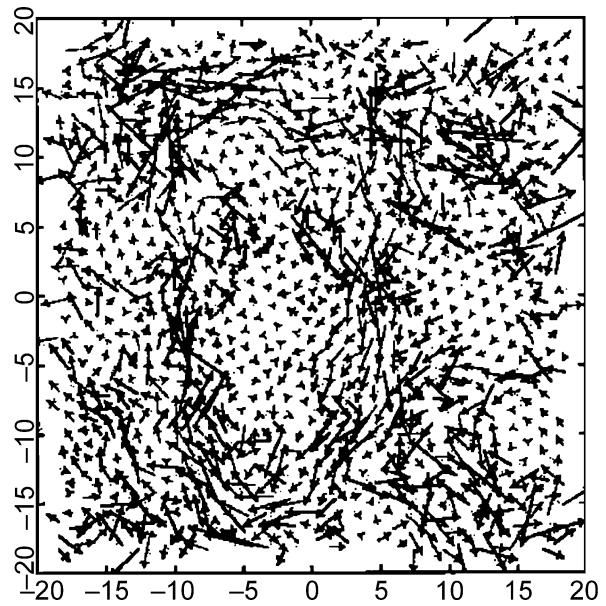
Classical computer simulations of supercooled liquids usually proceed by solving a cleverly discretized version of Newton's equations for a given potential interaction between particles [4]. If quantitative agreement with experimental data on an existing specific material is sought, the interaction must be carefully chosen in order to reproduce reality, for instance by combining classical to ab initio simulations. From a more fundamental perspective one rather seeks the simplest model that is still able to reproduce qualitatively the phenomenology of real glass-formers, while being considerably simpler to study. The implicit, but quite strong, hypothesis is that molecular details are not needed to explain the behaviour of supercooled liquids, so that the glass transition is indeed a topic for statistical mechanics, not for chemistry. A considerable amount of work has therefore been dedicated to studying models such as hard spheres, soft spheres, or Lennard-Jones particles. More realistic materials are also studied focusing for instance on the physics of network forming materials, multi-component ones, anisotropic particles, or molecules with internal degrees of freedom. Connections to experimental work can be made by computing quantities that are experimentally accessible such as the intermediate scattering function, static structure factors,  $S(\mathbf{q})$ , or thermodynamic quantities such specific heat or configurational entropy, which are directly obtained from particle trajectories and can be measured in experiments as well. As an example we show in Fig. 6 the intermediate scattering function  $F(\mathbf{q}, t)$  obtained from a molecular dynamics simulation of a classical model for  $\text{SiO}_2$  as a function of time for different temperatures [98].

An important role is played by simulations also because a large variety of dynamic and static quantities can

be simultaneously measured in a single model system. As we shall discuss below, there exist scores of different theoretical approaches to describe the physics of glass-formers, and they sometimes have their own set of predictions that can be readily tested by numerical work. Indeed, quite a large amount of numerical papers have been dedicated to testing in detail the predictions formulated by the mode-coupling theory of the glass transition, as reviewed recently in [94]. Here, computer simulations are particularly well-suited as the theory specifically addresses the relatively high temperature window that is studied in computer simulations.

While Newtonian dynamics is mainly used in numerical work on supercooled liquids, a most appropriate choice for these materials, it can be interesting to consider alternative dynamics that are not deterministic, or which do not conserve the energy. In colloidal glasses and physical gels, for instance, particles undergo Brownian motion arising from collisions with molecules in the solvent, and a stochastic dynamics is more appropriate. Theoretical considerations might also suggest the study of different sorts of dynamics for a given interaction between particles, for instance, to assess the role of conservation laws and structural information. Of course, if a given dynamics satisfies detailed balance with respect to the Boltzmann distribution, all structural quantities remain unchanged, but the resulting dynamical behaviour might be very different. Several papers [27,88,153] have studied in detail the influence of the chosen microscopic dynamics on the dynamical behaviour in glass-formers using either stochastic dynamics (where a friction term and a random noise are added to Newton's equations, the amplitude of both terms being related by a fluctuation-dissipation theorem), Brownian dynamics (in which there are no momenta, and positions evolve with Langevin dynamics), or Monte-Carlo dynamics (where the potential energy between two configurations is used to accept or reject a trial move). Quite surprisingly, the equivalence between these three types of stochastic dynamics and the originally studied Newtonian dynamics was established at the level of the averaged dynamical behaviour [27,88,153], except at very short times where obvious differences are indeed expected. This strongly suggests that an explanation for the appearance of slow dynamics in these materials originates from their amorphous structure. However, important differences were found when dynamic fluctuations were considered [21,22,27], even in the long-time regime comprising the structural relaxation.

Another crucial advantage of molecular simulations is illustrated in Fig. 7. This figure shows a spatial map of single particle displacements recorded during the simulation



**Glasses and Aging, A Statistical Mechanics Perspective on, Figure 7**

Spatial map of single particle displacements in the simulation of a binary mixture of soft spheres in two dimensions [99]. Arrows show the displacement of each particle in a trajectory of length about 10 times the structural relaxation time. The map reveals the existence of particles with different mobilities during relaxation, but also the existence of spatial correlations between these dynamic fluctuations

of a binary soft sphere system in two dimensions [99]. This type of measurement, out of reach of most experimental techniques that study the liquid state, reveals that dynamics might be very different from one particle to another. More importantly, Fig. 7 also unambiguously reveals the existence of spatial correlations between these dynamic fluctuations. The presence of non-trivial spatio-temporal fluctuations in supercooled liquids is now called 'dynamic heterogeneity' [72]. This is the phenomenon we discuss in more detail in the next section.

## Dynamic Heterogeneity

### Existence of Spatio-temporal Dynamic Fluctuations

A new facet of the relaxational behaviour of supercooled liquids has emerged in the last decade thanks to a considerable experimental and theoretical effort. It is called 'dynamic heterogeneity' (DH), and plays now a central role in modern descriptions of glassy liquids [72]. As anticipated in the previous section, the phenomenon of dynamic heterogeneity is related to the spatio-temporal fluctuations of the dynamics. Initial motivations stemmed from the

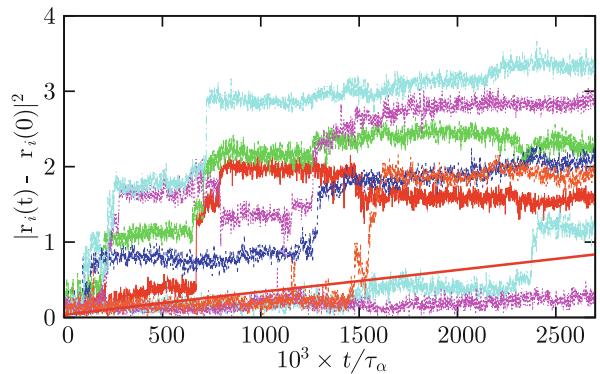
search for an explanation of the non-exponentiality of relaxation processes in supercooled liquids, related to the existence of a broad relaxation spectrum. Two natural, but fundamentally different, explanations can be put forward. (1) The relaxation is locally exponential, but the typical relaxation timescale varies spatially. Hence, global correlation or response functions become non-exponential upon spatial averaging over this spatial distribution of relaxation times. (2) The relaxation is complicated and inherently non-exponential, even locally. Experimental and theoretical works [72] suggest that both mechanisms are likely at play, but definitely conclude that relaxation is spatially heterogeneous, with regions that are faster and slower than the average. Since supercooled liquids are ergodic materials, a slow region will eventually become fast, and vice versa. A physical characterization of DH entails the determination of the typical lifetime of the heterogeneities, as well as their typical lengthscale.

A clear and more direct confirmation of the heterogeneous character of the dynamics also stems from simulation studies. For example, whereas the simulated average mean-squared displacements are smooth functions of time, time signals for individual particles clearly exhibit specific features that are not observed unless dynamics is resolved both in space and time. These features are displayed in Fig. 8. What do we see? We mainly observe that particle trajectories are not smooth but rather composed of a succession of long periods of time where particles simply vibrate around well-defined locations, separated by rapid ‘jumps’. Vibrations were previously inferred from the plateau observed at intermediate times in the mean-squared displacements of Fig. 5, but the existence of jumps that are clearly statistically widely distributed in time cannot be guessed from averaged quantities only. The fluctuations in Fig. 8 suggest, and direct measurements confirm, the importance played by fluctuations around the averaged dynamical behaviour.

A simple type of such fluctuations has been studied in much detail. When looking at Fig. 8, it is indeed natural to ask, for any given time, what is the distribution of particle displacements. This is quantified by the self-part of the van-Hove function defined as

$$G_s(\mathbf{r}, t) = \left\langle \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{r} - [\mathbf{r}_i(t) - \mathbf{r}_i(0)]) \right\rangle.$$

For an isotropic Gaussian diffusive process, one gets  $G_s(\mathbf{r}, t) = \exp(-|\mathbf{r}|^2/(4D_s t))/(4\pi D_s t)^{3/2}$ . Simulations reveal instead strong deviations from Gaussian behaviour on the timescales relevant for structural relaxation [116]. In particular they reveal ‘fat’ tails in the distributions that



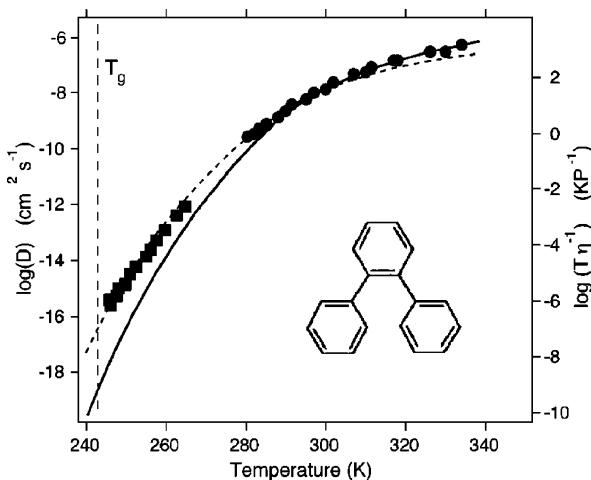
**Glasses and Aging, A Statistical Mechanics Perspective on, Figure 8**

Time resolved squared displacements of individual particles in a simple model of a glass-forming liquid composed of Lennard-Jones particles. The average is shown as a smooth full line. Trajectories are composed of long periods of time during which particles vibrate around well-defined positions, separated by rapid jumps that are widely distributed in time underlying the importance of dynamic fluctuations

are much wider than expected from the Gaussian approximation. These tails are in fact well described by an exponential, rather than Gaussian, decay in a wide time window comprising the structural relaxation, such that  $G_s(\mathbf{r}, t) \sim \exp(-|\mathbf{r}|/\lambda(t))$  [51]. Thus, they reflect the existence of a population of particles that moves distinctively further than the rest and appears therefore to be much more mobile. This observation implies that relaxation in a viscous liquid differs qualitatively from that of a normal liquid where diffusion is close to Gaussian, and that a non-trivial statistics of single particle displacements exists.

A long series of questions immediately follows this seemingly simple observation. Answering them has been the main occupation of many workers in this field over the last decade. What are the particles in the tails effectively doing? Why are they faster than the rest? Are they located randomly in space or do they cluster? What is the geometry, time and temperature evolution of the clusters? Are these spatial fluctuations correlated to geometric or thermodynamic properties of the liquids? Do similar correlations occur in all glassy materials? Can one predict these fluctuations theoretically? Can one understand glassy phenomenology using fluctuation-based arguments? Can these fluctuations be detected experimentally?

Another influential phenomenon that was related early on to the existence of DH is the decoupling of self-diffusion ( $D_s$ ) and viscosity ( $\eta$ ). In the high temperature liquid self-diffusion and viscosity are related by the Stokes–Einstein relation [95],  $D_s \eta / T = \text{const}$ . For a large

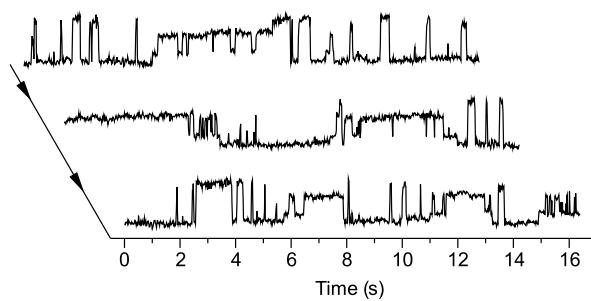


Glasses and Aging, A Statistical Mechanics Perspective on, Figure 9

Decoupling between viscosity (full line) and self-diffusion coefficient (symbols) in supercooled ortho-terphenyl [126]. The dashed line shows a fit with a ‘fractional’ Stokes–Einstein relation,  $D_s \sim (T/\eta)^\xi$  with  $\xi \sim 0.82$

particle moving in a fluid the constant is equal to  $1/(6\pi R)$  where  $R$  is the particle radius. Physically, the Stokes–Einstein relation means that two different measures of the relaxation time  $R^2/D_s$  and  $\eta R^3/T$  lead to the same timescale up to a constant factor. In supercooled liquids this phenomenological law breaks down, as shown in Fig. 9 for ortho-terphenyl [126]. It is commonly found that  $D_s^{-1}$  does not increase as fast as  $\eta$  so that, at  $T_g$ , the product  $D_s\eta$  has increased by 2–3 orders of magnitude as compared to its Stokes–Einstein value. This phenomenon, although less spectacular than the overall change of viscosity, is a significative indication that different ways to measure relaxation times lead to different answers and, thus, is a strong hint of the existence of a distribution of relaxation timescales.

Indeed, a natural explanation of this effect is that different observables probe differently way the underlying distribution of relaxation times [72]. For example, the self-diffusion coefficient of tracer particles is dominated by the more mobile particles whereas the viscosity or other measures of structural relaxation probe the timescale needed for every particle to move. An unrealistic but instructive example is a model where there is a small, non-percolative subset of particles that are blocked forever, coexisting with a majority of mobile particles. In this case, the structure never relaxes but the self-diffusion coefficient is non-zero because of the mobile particles. Of course, in reality all particles move, eventually, but this shows how different observables are likely to probe different moments of the dis-



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 10

Time series of polarization in the AFM experiment performed by Vidal Russell and Israeloff [161] on PVAc at  $T = 300$  K. The signal intermittently switches between periods with fast or slow dynamics, suggesting that extended regions of space indeed transiently behave as fast and slow regions

tribution of timescales, as explicitly shown within several theoretical frameworks [106,154].

The phenomena described above, although certainly an indication of spatio-temporal fluctuations, do not allow one to study how these fluctuations are correlated in space. This is, however, a fundamental issue both from the experimental and theoretical points of view. How large are the regions that are faster or slower than the average? How does their size depend on temperature? Are these regions compact or fractal? These important questions were first addressed in pioneering works using four-dimensional NMR [160], or by directly probing fluctuations at the nanoscopic scale using microscopy techniques. In particular, Vidal Russel and Israeloff using Atomic Force Microscopy techniques [161] measured the polarization fluctuations in a volume of size of few tens of nanometers in a supercooled polymeric liquid (PVAc) close to  $T_g$ . In this spatially resolved measurement, the hope is to probe a small enough number of dynamically correlated regions, and detect their dynamics. Indeed, the signal shown in Fig. 10 shows a dynamics which is very intermittent in time, the dynamics switching between moments with intense activity, and moments with no dynamics at all, suggesting that extended regions of space indeed transiently behave as fast and slow regions. A much smoother signal would have been measured if these such dynamically correlated ‘domains’ were not present. Spatially resolved and NMR experiments are quite difficult. They give undisputed information about the typical lifetime of the DH, but their determination of a dynamic correlation lengthscale is rather indirect and/or performed on a small number of liquids in a small temperature window. Nevertheless, the outcome is that

a non-trivial dynamic correlation length emerges at the glass transition, where it reaches a value of the order of 5–10 molecule diameters [72].

### Multi-point Correlation Functions

More recently, substantial progress in characterizing spatio-temporal dynamical fluctuations was obtained from theoretical [21,22,79,159] and numerical results [14, 28,76,99,172]. In particular, it is now understood that dynamical fluctuations can be measured and characterized through the use of four-point correlation functions. These multi-point functions can be seen as a generalization of the spin glass susceptibility measuring the extent of amorphous long-range order in spin glasses. In this subsection, we introduce these correlation functions and summarize the main results obtained using them.

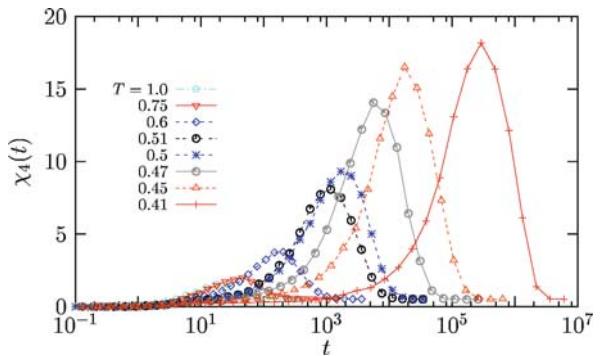
Standard experimental probes of the averaged dynamics of liquids give access to the time-dependent auto-correlation function of the spontaneous fluctuations of some observable  $O(t)$ ,  $F(t) = \langle \delta O(0)\delta O(t) \rangle$ , where  $\delta O(t) = O(t) - \langle O \rangle$  represents the instantaneous value of the deviation of  $O(t)$  from its ensemble average  $\langle O \rangle$  at time  $t$ . One can think of  $F(t)$  as being the average of a two-point quantity,  $C(0, t) = \delta O(0)\delta O(t)$ , characterizing the dynamics. A standard example corresponds to  $O$  being equal to the Fourier transform of the density field. In this case  $F(t)$  is the dynamical structure factor as in Eq. (2). More generally, the correlation functions  $F(t)$  measure the global relaxation in the system. Intuitively, in a system with important dynamic correlations, the fluctuations of  $C(0, t)$  will be stronger. Quantitative information on the amplitude of those fluctuations is provided by the variance

$$\chi_4(t) = N \langle \delta C(0, t)^2 \rangle , \quad (3)$$

where  $\delta C(0, t) = C(0, t) - F(t)$ , and  $N$  is the total number of particles in the system. The associated spatial correlations show up more clearly when considering a ‘local’ probe of the dynamics, like for instance an orientational correlation function measured by dielectric or light scattering experiments, which can be expressed as

$$C(0, t) = \frac{1}{V} \int d^3r c(\mathbf{r}; 0, t) , \quad (4)$$

where  $V$  is the volume of the sample and  $c(\mathbf{r}; 0, t)$  characterizes the dynamics between times 0 and  $t$  around point  $\mathbf{r}$ . For example, in the above mentioned case of orientational correlations,  $c(\mathbf{r}; 0, t) \propto \frac{V}{N} \sum_{i,j=1}^N \delta(\mathbf{r} - \mathbf{r}_i) Y(\Omega_i(0))Y(\Omega_j(t))$ , where  $\Omega_i$  denotes the angles describing the orientation of molecule  $i$ ,  $\mathbf{r}_i(0)$  is the position of that molecule at time 0, and  $Y(\Omega)$  is some appropriate rotation matrix element. Here, the ‘locality’ of the



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 11

Time dependence of  $\chi_4(t)$  quantifying the spontaneous fluctuations of the intermediate scattering function in a Lennard-Jones supercooled liquid. For each temperature,  $\chi_4(t)$  has a maximum, which shifts to larger times and has a larger value when  $T$  is decreased, revealing the increasing lengthscale of dynamic heterogeneity in supercooled liquids approaching the glass transition

probe comes from the fact that it is dominated by the self-term involving the same molecule at times 0 and  $t$ , or by the contribution coming from neighboring molecules. The dynamic susceptibility  $\chi_4(t)$  can thus be rewritten as

$$\chi_4(t) = \rho \int d^3r G_4(\mathbf{r}; 0, t) , \quad (5)$$

where

$$G_4(\mathbf{r}; 0, t) = \langle \delta c(\mathbf{0}; 0, t) \delta c(\mathbf{r}; 0, t) \rangle , \quad (6)$$

and translational invariance has been taken into account ( $\rho = N/V$  denotes the mean density). The above equations show that  $\chi_4(t)$  measures the extent of spatial correlation between dynamical events between times 0 and  $t$  at different points of the system, i.e., the spatial extent of dynamically heterogeneous regions over a time span  $t$ .

The function  $\chi_4(t)$  has been measured by molecular dynamics, Brownian and Monte Carlo simulations in different liquids [14,28,29,76,163]. An example is shown in Fig. 11 for a Lennard-Jones liquid. The qualitative behaviour is similar in all cases [21,79,159]: as a function of time  $\chi_4(t)$  first increases, it has a peak on a timescale that tracks the structural relaxation timescale and then it decreases<sup>2</sup>. The peak value measures thus the volume

<sup>2</sup>The decrease at long times constitutes a major difference with spin glasses. In a spin glass,  $\chi_4$  would be a monotonically increasing function of time whose long-time limit coincides with the static spin glass susceptibility. Physically, the difference is that spin glasses develop long-range static amorphous order while structural glasses do not, or, at least, in a different and more subtle way.

on which the structural relaxation processes are correlated. It is found to increase when the temperature decreases and the dynamics slows down. By measuring directly  $G_4(\mathbf{r}; 0, t)$  it has also been checked that the increase of the peak of  $\chi_4(t)$  corresponds, as expected, to a growing dynamic lengthscale  $\xi$  [14,21,28], although these measurements are much harder in computer simulations, because very large systems need to be simulated to determine  $\xi$  unambiguously. Note that if the dynamically correlated regions were compact, the peak of  $\chi_4$  would be proportional to  $\xi^3$  in three dimensions, directly relating  $\chi_4$  measurements to that of the relevant lengthscale of DH.

These results are also relevant because many theories of the glass transition assume or predict, in a way or another, that the dynamics slows down because there are increasingly large regions on which particles have to relax in a correlated or cooperative way. However, this lengthscale remained elusive for a long time. Measures of the spatial extent of dynamic heterogeneity, in particular  $\chi_4(t)$  and  $G_4(\mathbf{r}; 0, t)$ , seem to provide the long-sought evidence of this phenomenon. This in turn suggests that the glass transition is indeed a critical phenomenon characterized by growing timescales and lengthscales. A clear and conclusive understanding of the relationship between the lengthscale obtained from  $G_4(\mathbf{r}; 0, t)$  and the relaxation timescale is still the focus of an intense research activity.

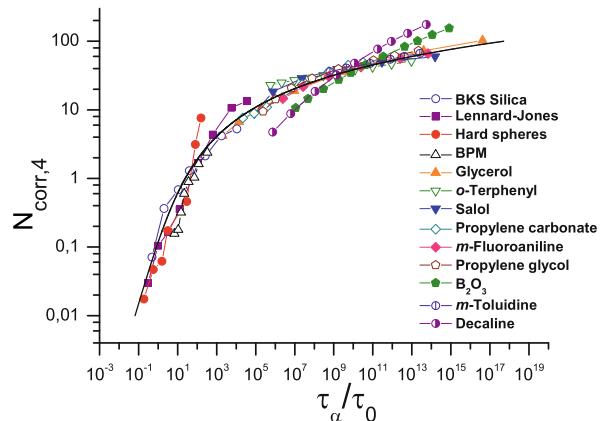
One major issue is that obtaining information on the behaviour of  $\chi_4(t)$  and  $G_4(\mathbf{r}; 0, t)$  from experiments is difficult. Such measurements are necessary because numerical simulations can only be performed rather far from  $T_g$ , see Sect. “Numerical Simulations”. Up to now, direct experimental measurements of  $\chi_4(t)$  have been restricted to colloidal [166] and granular materials [65,110] close to the jamming transition, because dynamics is more easily spatially resolved in those cases. Unfortunately, similar measurements are currently not available in molecular liquids.

Recently, an approach based on fluctuation-dissipation relations and rigorous inequalities has been developed in order to overcome this difficulty [20,21,22]. The main idea is to obtain a rigorous lower bound on  $\chi_4(t)$  using the Cauchy–Schwarz inequality  $\langle \delta H(0) \delta C(0, t) \rangle^2 \leq \langle \delta H(0)^2 \rangle \langle \delta C(0, t)^2 \rangle$ , where  $H(t)$  denotes the enthalpy at time  $t$ . By using fluctuation-dissipation relations the previous inequality can be rewritten as [20]

$$\chi_4(t) \geq \frac{K_B T^2}{c_P} [\chi_T(t)]^2 , \quad (7)$$

where the multi-point response function  $\chi_T(t)$  is defined by

$$\chi_T(t) = \left. \frac{\partial F(t)}{\partial T} \right|_{N,P} = \frac{N}{K_B T^2} \langle \delta H(0) \delta C(0, t) \rangle .$$



**Glasses and Aging, A Statistical Mechanics Perspective on, Figure 12**

Universal dynamic scaling relation between number of dynamically correlated particles,  $N_{\text{corr},4}$ , and relaxation timescale,  $\tau_\alpha$ , for a number of glass-formers [63], determined using Eq. (7)

In this way, the experimentally accessible response  $\chi_T(t)$  which quantifies the sensitivity of average correlation functions  $F(t)$  to an infinitesimal temperature change, can be used in Eq. (7) to yield a lower bound on  $\chi_4(t)$ . Moreover, detailed numerical simulations and theoretical arguments [21,22] strongly suggest that the right hand side of (7) actually provides a good estimation of  $\chi_4(t)$ , not just a lower bound.

Using this method, Dalle-Ferrier et al. [63] have been able to obtain the evolution of the peak value of  $\chi_4$  for many different glass-formers in the entire supercooled regime. In Fig. 12 we show some of these results as a function of the relaxation timescale. The value on the  $y$ -axis, the peak of  $\chi_4$ , is a proxy for the number of molecules,  $N_{\text{corr},4}$  that have to evolve in a correlated way in order to relax the structure of the liquid. Note that  $\chi_4$  is expected to be equal to  $N_{\text{corr},4}$  up to a proportionality constant which is not known from experiments, probably explaining why the high temperature values of  $N_{\text{corr},4}$  are smaller than one. Figure 12 also indicates that  $N_{\text{corr},4}$  grows faster when  $\tau_\alpha$  is not very large, close to the onset of slow dynamics, and a power law relationship between  $N_{\text{corr},4}$  and  $\tau_\alpha$  is good in this regime ( $\tau_\alpha/\tau_0 < 10^4$ ). The growth of  $N_{\text{corr},4}$  becomes much slower closer to  $T_g$ . A change of 6 decades in time corresponds to a mere increase of a factor about 4 of  $N_{\text{corr},4}$ , suggesting logarithmic rather than power law growth of dynamic correlations. This is in agreement with several theories of the glass transition which are based on activated dynamic scaling [85,155,171].

Understanding quantitatively this relation between timescales and lengthscales is one of the main recent topics addressed in theories of the glass transition, see

Sect. “[Some Theory and Models](#)”. Furthermore, numerical works are also devoted to characterizing better the geometry of the dynamically heterogeneous regions [7,69].

### [Some Theory and Models](#)

We now present some theoretical approaches to the glass transition. It is impossible to cover all of them in a brief review, simply because there are way too many of them, perhaps the clearest indication that the glass transition remains an open problem. We choose to present approaches that are keystones and have a solid statistical mechanics basis. Loosely speaking, they have an Hamiltonian, can be simulated numerically, or studied analytically with statistical tools. Of course, the choice of Hamiltonians is crucial and contains very important assumptions about the nature of the glass transition. All these approaches have given rise to unexpected results. One finds more in them than what was supposed at the beginning, which leads to new, testable predictions. Furthermore, with models that are precise enough, one can test (and hopefully falsify!) these approaches by working out all their predictions in great detail, and comparing the outcome to experimental data. This is not possible with ‘physical pictures’, or simpler approaches of the problem which we have therefore avoided.

Before going into the models, we would like to state the few important questions that face theoreticians.

- Why do the relaxation time and the viscosity increase when  $T_g$  is approached? Why is this growth super-Arrhenius?
- Can one understand and describe quantitatively the average dynamical behaviour of supercooled liquids, in particular broad relaxation spectra, non-exponential behaviour, and their evolution with fragility?
- Is there a relation between kinetics and thermodynamics (like  $T_0 \simeq T_K$ ), and why?
- Can one understand and describe quantitatively the spatio-temporal fluctuations of the dynamics? How and why are these fluctuations related to the dynamic slowing down?
- Is the glass transition a collective phenomenon? If yes, of which kind? Is there a finite temperature or zero temperature ideal glass transition? In this case, is the transition of static or purely dynamic origin?
- Is there a geometric, real space explanation for the dynamic slowing down that takes into account molecular degrees of freedom?

The glass transition appears as a kind of ‘intermediate coupling’ problem, since for instance typical growing length-

scales are found to be at most a few tens of particle large close to  $T_g$ . It would therefore be difficult to recognize the correct theory even if one bumped into it. To obtain quantitative, testable predictions, one must therefore be able to work out also preasymptotic effects. This is particularly difficult, especially in cases where the asymptotic theory itself has not satisfactorily been worked out. As a consequence, at this time, theories can only be judged by their overall predictive power and their theoretical consistency.

### [Cooperativity, Chaotic Energy Landscapes and Random First Order Theory](#)

In the last two decades, three *independent* lines of research approaches, Adam–Gibbs theory [2], mode-coupling theory [94] and spin glass theory [137], have merged to produce a theoretical ensemble that now goes under the name of Random First Order Theory (RFOT), a terminology introduced by Kirkpatrick, Thirumalai and Wolynes [111] who also played a major role in its development. Instead of following the rambling development of history, we summarize it in a more modern and unified way.

A key ingredient of RFOT is the existence of a chaotic or complex free energy landscape and its evolution with temperature and/or density. Analysing it in a controlled way for three dimensional interacting particles system is of course an impossible task. This can be achieved, however, in simplified models or using mean-field approximation, that have therefore played a crucial role in the development of RFOT.

A first, concrete example is given by ‘lattice glass models’ [37]. These are models containing hard particles sitting on the sites of a lattice. The Hamiltonian is infinite if there is more than one particle on a site or if the number of occupied neighbors of an occupied site is larger than a parameter,  $m$ , but the Hamiltonian is zero otherwise. Tuning the parameter  $m$ , or changing the type of lattice, in particular its connectivity, yields different models. Lattice glasses are constructed as simple statmech models to study the glassiness of hard sphere systems. The constraint on the number of occupied neighbors mimicks the geometric frustration [139] encountered when trying to pack hard spheres in three dimensions. Other models, which have a finite energy and, hence, are closer to molecular glass-formers, can be also constructed [131]. These models can be solved exactly on a Bethe lattice<sup>3</sup>, which reveals an astonishing physical behaviour [147]. In particular their free energy landscape can be analyzed in full details and turns

<sup>3</sup>In order to have a well-defined thermodynamics, Bethe lattices are generated as random graphs with fixed connectivity, also called random regular graphs.

out to have the properties that are also found in several ‘generalized spin glasses’.

Probably the most studied example of such spin glasses is the  $p$ -spin model, defined by the Hamiltonian [93]

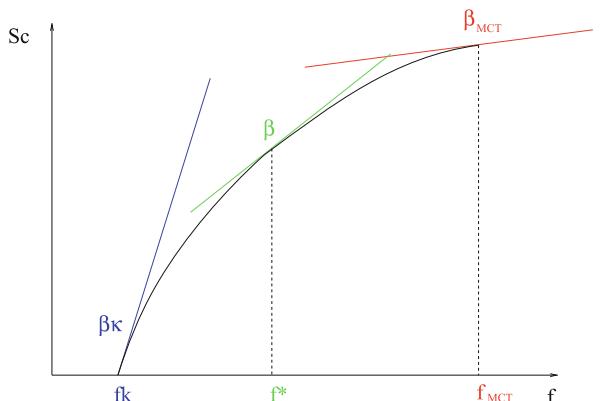
$$H = - \sum_{i_1, \dots, i_p} J_{i_1, \dots, i_p} S_{i_1} \dots S_{i_p}, \quad (8)$$

where the  $S_i$ s are Ising or spherical spins,  $p > 2$  and  $J_{i_1, \dots, i_p}$  quenched random couplings with zero mean variance  $p!/(2N^{p-1})$ .

All these models (lattice glasses, their finite energy generalizations and their quenched disorder counterparts) belong the class of one-step replica symmetry breaking systems [137]. This makes reference to the ansatz that is needed [93] when replica techniques are used to compute the thermodynamic behaviour of the model in Eq. (8). This corresponds to the universality class of chaotic (or random) free energy landscapes, as we now explain.

The free energy landscape of these systems is ‘rugged’, characterized by many minima and saddle points. Actually, the number of stationary points is so large that in order to count them one has to introduce an entropy, called configurational entropy or complexity,  $s_c = 1/N \log \mathcal{N}(f)$ , where  $\mathcal{N}(f)$  is the number of stationary points with a given free energy density  $f$ . The density profile corresponding to one given minimum is amorphous and lacks any type of periodic long-range order, and different mimima are very different. Defining a similarity measure between them, an ‘overlap’, one typically finds that two minima with the same free energy  $f$  have zero overlap. The typical shape of the configurational entropy as a function of  $f$  is plotted in Fig. 13.

At high temperature, there is typically a single minimum, the high temperature liquid state. There is a temperature below which an exponentially large (in the system size) number of minima appears. Within mean-field models, corresponding to Bethe or completely connected lattices, these minima correspond to macroscopic physical states analogous to the periodic minimum corresponding to the crystal<sup>4</sup>. Once the system is in one of these states it remains trapped in it forever, since barriers separating states diverge with the system size. However, when transposed to finite dimensional systems, these states become metastable and have a finite lifetime. As a consequence, in order to compute thermodynamic properties, one has to sum over all of them using the Boltzmann



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 13

Typical shape of the configurational entropy,  $s_c$ , as a function of free energy density,  $f$  in the range  $T_K < T < T_{MCT}$  for random first order landscapes. A graphic solution of Eq. (10) is obtained by finding the value of  $f$  at which the slope of the curve is  $\beta$ . Note that  $s_c$  is also a function of temperature, so this curve in fact changes with  $T$

weight  $\exp(-\beta N f_\alpha)$  for each state  $\alpha$  [135]:

$$Z = \sum_{\alpha} e^{-\beta N f_{\alpha}} = \int df \exp[N s_c(f; T)] e^{-\beta N f}, \quad (9)$$

where  $\beta = 1/(K_B T)$ . Evaluating this sum by saddle point method yields three regimes. At high temperature,  $T > T_{MCT}$ , the liquid corresponding to a flat density profile dominates the sum. The landscape is simple and has a single minimum. This is followed by an intermediate temperature regime,  $T_K < T < T_{MCT}$ , where the sum is dominated by all terms with free energy density satisfying

$$\left. \frac{\partial s_c(f, T)}{\partial f} \right|_{f=f^*} = \beta. \quad (10)$$

There are many of them, the logarithm of their number being given by  $N s_c(f^*, T)$ , see Fig. 13 for a graphical solution of Eq. (10). Upon decreasing the temperature,  $s_c(f^*, T)$  decreases until a temperature,  $T_K$ , below which the sum in Eq. (9) becomes dominated by only few terms corresponding to states with free energy density  $f_K$  given by  $s_c(f_K, T) = 0$ , see Fig. 13. The entropy in the intermediate temperature range above  $T_K$  has two contributions: the one counting the number of minima, given by  $s_c$ , and the intra-state entropy,  $s_{in}$ , counting the number configurations inside each state. At  $T_K$ , the configurational entropy vanishes,  $s_c(T_K) = 0$ . As a consequence the specific heat undergoes a jump towards a smaller value across  $T_K$ ,

<sup>4</sup>There is of course no crystal state in disordered systems such as in Eq. (8). In the case of lattice glass models, there is a crystal phase but it can disappear depending whether the Bethe lattice is a Cayley tree or a random regular graph.

an exact realization of the ‘entropy vanishing’ mechanism conjectured by Kauzmann [108].

Let us discuss the dynamical behaviour which results from the above analysis. We have already mentioned that relaxation processes do not occur below  $T_{\text{MCT}}$  because states have an infinite lifetime. The stability of these states can be analyzed by computing the free energy Hessian in the minima [48]. One finds that states become more fragile when  $T \rightarrow T_{\text{MCT}}^-$ , are marginally stable at  $T = T_{\text{MCT}}$ , unstable for  $T > T_{\text{MCT}}$ . The dynamics of these models can be analyzed exactly [58]. Coming from high temperature, the dynamics slows down and the relaxation time diverges at  $T_{\text{MCT}}$  in a power law manner,

$$\tau_\alpha \sim \frac{1}{(T - T_{\text{MCT}})^\gamma}, \quad (11)$$

where  $\gamma$  is a critical exponent. The physical reason is the incipient stable states that appear close to  $T_{\text{MCT}}$ . The closer the temperature is to  $T_{\text{MCT}}$ , the longer it takes to find an unstable direction to relax.

Amazingly, the dynamical transition that appears upon approaching  $T_{\text{MCT}}$  in random first order landscapes is completely analogous to the one predicted to occur in supercooled liquids by the Mode-Coupling Theory of the glass transition, and developed independently by Leuthegger, Bengzelius, Götze, Sjölander and coworkers [94]. Actually, MCT can be considered as an approximation which becomes controlled and exact for these mean-field models. Originally, MCT was developed using projector operator formalism [13,122] and field-theory methods [64] to yield closed integro-differential equations for the dynamical structure factor in supercooled liquids. These approaches were recently generalized [34,36] to deal with dynamic heterogeneity and make predictions for the multi-point susceptibilities and correlation functions discussed in Sect. “Dynamic Heterogeneity”. Within MCT, the relaxation timescale diverges in a power law fashion at  $T_{\text{MCT}}$ , as in Eq. (11). This divergence is accompanied by critical behaviour that appears both in space (long range spatial dynamic correlations), and in time (power laws in time).

Comparing Eqs. (1) and (11) makes it clear that MCT cannot be used to describe viscosity data close to  $T_g$  since it does not predict activated behaviour. It is now recognized that an MCT transition at  $T_{\text{MCT}}$  does not occur in real materials, so that  $T_{\text{MCT}}$  is, at best, a dynamical crossover. A central advantage of MCT, compared to many other theories (this includes the  $T \approx T_K$  regime of RFOT itself) is that it can yield quantitative predictions from microscopic input obtained for a particular material. As such it has been applied to yield predictions for scores of different sys-

tems that can be directly confronted to experimental or numerical measurements. A major drawback is the freedom offered by the ‘crossover’ nature of the MCT transition, so that ‘negative’ results can often be attributed to corrections to asymptotic predictions rather than deficiencies of the theory itself. Nevertheless, MCT has proven to be useful and continues to be developed, applied and generalized to study many different physical situations [94], including aging systems and non-linear rheology of glassy materials [18,83,134], see Sect. “Aging and Off-equilibrium”.

What happens below  $T_{\text{MCT}}$  in finite dimensional system if the relaxation time does not diverge as predicted in Eq. (11)? Why is the transition avoided? In fact, the plethora of states that one finds in mean-field are expected to become (at best) metastable in finite dimension, with a finite lifetime, even below  $T_{\text{MCT}}$ . What is their typical lifetime, and how these metastable states are related to the structural relaxation are issues that still await for a complete microscopic analysis.

There exist, however, phenomenological arguments [38,112,171], backed by microscopic computations [71, 80] that yield a possible solution dubbed ‘mosaic state’ by Kirkpatrick, Thirumalai and Wolynes [112]. Schematically, the mosaic picture states that, in the regime  $T_K < T < T_{\text{MCT}}$ , the liquid is composed of domains of linear size  $\xi$ . Inside each domain, the system is in one of the mean-field states. The length of the domains is fixed by a competition between energy and configurational entropy. A state in a finite but large region of linear size  $l$  can be selected by appropriate boundary conditions that decrease its free energy by an amount which scales as  $\Upsilon l^\theta$  with  $\theta \leq 2$ . On the other hand, the system can gain entropy, which scales as  $s_c l^3$ , if it visits the other numerous states. Entropy obviously gains on large lengthscales, the crossover length  $\xi$  being obtained by balancing the two terms,

$$\xi = \left( \frac{\Upsilon}{Ts_c(T)} \right)^{1/(3-\theta)}. \quad (12)$$

In this scenario, the configurational entropy on scales smaller than  $\xi$  is too small to stir the configurations efficiently and win over the dynamically generated pinning field due to the environment, while ergodicity is restored at larger scale. Hence, the relaxation time of the system is the relaxation time,  $\tau(\xi)$ , of a finite size regions of the system. Barriers are finite, unlike in the mean-field treatment. Smaller length scales are faster but unable to decorrelate, whereas larger scales are orders of magnitude slower. Assuming thermal activation over energy barriers which are supposed to grow with size as  $\xi^\psi$ , one predicts finally, us-

ing Eq. (12), that [38]

$$\log \left( \frac{\tau_\alpha}{\tau_0} \right) = c \frac{\Upsilon}{K_B T} \left( \frac{\Upsilon}{T_{S_c}(T)} \right)^{\psi/(3-\theta)}, \quad (13)$$

where  $c$  is a constant.

The above argument is rather generic and therefore not very predictive. Recent microscopic computations [71,80] aimed at putting these phenomenological arguments on a firmer basis and computing the exponents  $\theta$  and  $\psi$ . The results are unfortunately not yet conclusive because they involve replica calculations with uncontrolled assumptions, but they do confirm the phenomenological scenario presented above and suggest that  $\theta = 2$ . Some other phenomenological arguments suggest the value of  $\theta = 3/2$  [112]. There are no computation available for  $\psi$ , only the suggestion that  $\psi = \theta$  [112].

Note that using the value  $\theta = 3/2$  with  $\theta = \psi$  simplifies Eq. (13) into a form that is well-known experimentally and relates  $\log \tau_\alpha$  directly to  $1/S_c$ , which is the celebrated Adam–Gibbs relation [2] between relaxation time and configurational entropy that is in rather good quantitative agreement with many experimental results [5,96,105]. The Random First Order Theory can be considered, therefore, as a microscopic theory that reformulates and generalizes the Adam–Gibbs mechanism. Furthermore, using the fact that the configurational entropy vanishes linearly at  $T_K$  one predicts also a VFT divergence of the relaxation time as in Eq. (1), with the identification that

$$T_0 = T_K. \quad (14)$$

The equality (14) between two temperatures that are commonly used in the description of experimental data certainly constitutes a central achievement of RFOT since it accounts for the empirical relation found between the kinetics and the thermodynamics of supercooled liquids. Furthermore RFOT naturally contains MCT, which can be used to describe the first decades of the dynamical slowing down, while the spin glass side of RFOT qualitatively explains the dynamics in terms of the peculiar features of the free energy landscape that have been detailed above. Dynamics first slows down because there appear incipient metastable states, and once these metastable states are formed, the dynamics becomes dominated by the thermally activated barrier crossing from one metastable state to another, which is consistent with the relation between dynamical correlation length and timescale discussed in Sect. “Dynamic Heterogeneity”. Quite importantly, microscopic computations of  $T_{MCT}$  and  $T_0$  for realistic models of liquids are possible [136]. Remarkably, the jamming transition of hard spheres systems has been also studied

with these techniques and a clear connection with the glass transition has emerged within RFOT [143]. This quantitative “side” of RFOT is a most desirable feature, even if the results are not always quantitatively accurate [54,138,156].

Probably the most serious weakness of the RFOT construction is that the theory, although worked out in full details within mean-field models, has remained elusive for finite dimensional systems, for which it has a highly speculative flavour. Worrying is the fact that no simple three-dimensional glassy model, let alone interacting particles in the continuum, has been discovered, for which this theory has been shown to apply, and the entropy driven nucleation theory that leads to the VFT law is not understood completely. Although the ultimate consequences of the theory are sometimes in very good agreement with experiments, as Eq. (14), one should not conclude that RFOT is correct. In fact direct tests of the mosaic state picture are rare, and rather inconclusive [49]. One can hope that in the next few years, joint theoretical and experimental efforts will drive RFOT into a corner, to a point where it can be decided whether it is truly a valid theory for the glass transition.

### Free Volume, Defects, and Facilitated Models

In this subsection we motivate and briefly summarize studies of a different family of statistical mechanics models that turns out to yield a rich variety of physical behaviours. Their starting point are physical assumptions that might seem similar to the models described in Sect. “Cooperativity, Chaotic Energy Landscapes and Random First Order Theory”, but the outcome yields a different physical explanation of the glass transition. Although the two theoretical approaches cannot be simultaneously correct, they both have been influential and very instructive in order to develop a theoretical understanding of glassy phenomena. Furthermore, despite the ‘great unification’ phase diagram in Fig. 4, it could be that glass and jamming transitions in colloids, granular media and glass-formers have a different nature, so that different theories could apply to different phenomena.

As in Sect. “Cooperativity, Chaotic Energy Landscapes and Random First Order Theory”, we start from the packing considerations that are more appropriate for hard spheres systems. We follow first Kob and Andersen [115] and again use a lattice gas description of the physics and work on a three dimensional cubic lattice. As in a hard sphere system, we assume no interaction between particles beyond the hard-core constraint that the occupation number  $n_i$  at site  $i$  is at most equal to 1,

$$H[\{n_i\}] = 0, \quad n_i = 0, 1. \quad (15)$$

Contrary to the lattice glass model presented above, all configurations respecting the hard-core constraint are allowed and are equally probable. Geometric frustration is instead introduced at the level of the kinetic rules, that are defined as constrained local moves. Namely, a particle can jump to a nearest neighbor site only if that site is empty (to satisfy the hard-core constraint), but, additionnally, only if the sites occupied before and after the move have less than  $m$  neighbors,  $m$  being an adjustable parameter, which Kob and Andersen choose as  $m = 4$  for  $d = 3$  ( $m = 6$  corresponds to the unconstrained lattice gas). The model captures the idea that if the liquid is locally very dense, no movement is possible while regions with low density move more easily.

Of course, such kinetically constrained lattice gases have been studied in various spatial dimensions, for different values of  $m$ , for different constraints, or even different lattice geometries [146]. They can be thought of as models capturing the idea of a ‘cage’ effect in a strict sense, utilizing the notion that a particle with a dense neighbor shell cannot diffuse. Although the cage seems a purely local concept, it turns out that diffusion in constrained lattice gases arises from cooperative rearrangements, so that slow dynamics can be directly shown to be driven by the growth of dynamic lengthscales for these cooperative moves [77,141,157]. This strongly suggests that such cooperative moves are most probably at work also in real liquids.

In this lattice gas picture, the connection with liquid is not obvious because density (‘free volume’), rather than temperature controls the dynamics. Thermal models with similar features can in fact be defined along the following lines. In a liquid, low temperature implies a very small probability to find a location with enough free volume to move. The idea of a small concentration of ‘hot spots’ is in fact reminiscent of another picture of the glass transition based on the idea of ‘defects’ which is captured by the defect model proposed by Glarum [87] in the 60’s, where relaxation proceeds via the diffusion of a low concentration of independent defects. In the mid-80’s, using the conjugated ideas of kinetic constraints and rare defects, Fredrickson and Andersen defined a family of kinetic Ising models for the glass transition [81]. They study an assembly of non-interacting spins,

$$H[\{n_i\}] = \sum_{i=1}^N n_i, \quad n_i = 0, 1, \quad (16)$$

where  $n_i = 1$  represent the defects, whose concentration becomes exponentially small at low temperature,  $\langle n_i \rangle \approx \exp(-1/T)$ . As for the Kob–Andersen lattice gas, the non-

trivial ingredient lies in the chosen rates for the transition between states. The kinetic rules stipulate that a transition at site  $i$  can happen with a usual Glauber rate, but only if site  $i$  is surrounded by at least  $k$  defects ( $k = 0$  corresponds to the unconstrained limit). Again, one can easily imagine studying such models in different spatial dimensions, on different lattices, and with slightly different kinetic rules, yielding a large number of possible behaviours [125,146]. The similarity between those spin facilitated models and the kinetically constrained lattice gases is striking. Altogether, they form a large family of models generically called kinetically constrained models (KCMs) [146].

The connection between KCMs and the much older concept of free volume is obvious from our presentation. Free volume models are among the most widely used models to analyze experimental data, especially in polymeric systems. They have been thoroughly reviewed before [53,66], and the main prediction is that dynamic slowing down occurs because the free volume available to each particle,  $v_f$ , vanishes at some temperature  $T_0$  as  $v_f \approx \alpha(T - T_0)$ , a relation which connects volume to temperature. Statistical arguments then relate relaxation timescales to free volume assuming that movement is possible if locally there is ‘enough’ free volume available, more than a typical value  $v_0$ . This is clearly reminiscent of the above idea of a kinetic constraint for local moves in lattice gases. An appealing VFT divergence is then predicted:

$$\frac{\tau_\alpha}{\tau_0} \sim \exp\left(\gamma \frac{v_0}{v_f}\right) \sim \exp\left(\frac{\gamma v_0/\alpha}{[T - T_0]^\mu}\right), \quad (17)$$

where  $\gamma$  is a numerical factor and  $\mu = 1$ . Predictions such as Eq. (17) justify the wide use of free volume approaches, despite the many (justified) criticisms that have been raised.

Initially it was suggested that KCMs would similarly display finite temperature or finite density dynamic transitions similar to the one predicted by the mode-coupling theory of supercooled liquids [81], but it was soon realized [46,82] that most KCMs do not display such singularity, and timescales in fact only diverge in the limit of zero temperature ( $T = 0$ ) or maximal density ( $\rho = 1$ ). Recently, models displaying a  $T_0 > 0$  or  $\rho_0 < 1$  transition have been introduced and analyzed [158]. They provide a microscopic realization, based on well-defined statistical mechanics models, of the glass transition predicted by free volume arguments. Their relaxation timescale diverges with a VFT-like form but with an exponent  $\mu \simeq 0.64$ . Understanding their universality classes and how general is the mechanism leading to the transition is still an open problem.

Extensive studies have shown that KCMs have a macroscopic behaviour which resembles the phenomenology of supercooled liquids, displaying in particular Arrhenius or super-Arrhenius increase of relaxation timescales on decreasing the temperature and non-exponential relaxation functions at equilibrium [146]. Early studies also demonstrated that, when suddenly quenched to very low temperatures, the subsequent non-equilibrium aging dynamics of the models compares well with experimental observations on aging liquids [82]. Moreover, the many possibilities to define the models mean that they might exhibit a broad variety of possible behaviours. This is both a positive and a negative aspect: on the one hand one can explore various scenarios to describe glass transition phenomena, but on the other hand, one would like to be able to decide what particular model should be used if one wants to get a predictive quantitative description for a particular liquid. In fact, contrary to MCT, no microscopic calculations have been performed using the framework of KCMs. Rather than predicting the quantitative behaviour of a material in all its microscopic details, it is perhaps more appropriate to use KCMs as theoretical tools to define concepts and obtain new ideas.

It is precisely in this perspective that interest in KCMs continues to increase, in large part since it was realized that their dynamics is spatially heterogeneous [46,77,84], a central feature of supercooled liquids dynamics. It is only fair to say that in-depth studies of KCMs have greatly contributed to our theoretical understanding of the spatially heterogeneous dynamics in glass and jamming problems. Remarkably, virtually all the aspects related to dynamic heterogeneity mentioned in Sect. “[Dynamic Heterogeneity](#)” can be investigated and rationalized, at least qualitatively, in terms of KCMs. The dynamics of these systems can be understood in terms of defects motion [146]. Depending on the particular model, defects can diffuse or have a more complicated motion. Furthermore, they can simply be point-like, or ‘cooperative’ (formed by point-like defects moving in a cooperative way). A site can relax when it is visited by a defect. As a consequence, the heterogeneous character of the dynamics is entirely encoded in the defect configuration and defect motion [84]. For instance, a snapshot similar to Fig. 7 in a KCM shows clusters which have relaxed within the time interval  $t$  [26,167]. These are formed by all sites visited by a defect between 0 and  $t$ . The other sites are instead frozen in their initial state. In these models the dynamics slows down because the defect concentration decreases. As a consequence, in the regime of slow dynamics there are few defects and strong dynamic heterogeneity. Detailed numerical and analytical studies have indeed shown that in these

systems, non-exponential relaxations patterns do stem from a spatial, heterogeneous distribution of timescales, directly connected to a distribution of dynamic lengthscales [84,103,141,157,158,167]. Decoupling phenomena appear in KCMs and can be shown to be very direct, quantifiable, consequences of the dynamic heterogeneity [106], which also deeply affects the process of self-diffusion in a system close to its glass transition [24]. More fundamentally, multi-point susceptibilities, multi-point spatial correlation functions such as the ones defined in Eqs. (3) and (6) can be studied in much greater detail than in molecular systems, to the point that scaling relations between timescales, lengthscales, and dynamic susceptibilities can be established [22,50,141,159,168]. This type of scaling behaviour has been observed close to  $T = 0$  and  $\rho = 1$  in spin models and lattice gases without a transition<sup>5</sup>. These particular points of the phase diagram have been shown, by various theoretical means, to correspond to true critical points where timescales and dynamic lengthscales diverge with well-defined critical laws [103,168]. Such ‘dynamic criticality’ is a useful concept because it implies the possibility that some universal behaviour emerges in the physics of supercooled liquids, precisely of the type observed in Fig. 12.

A central criticism about the free volume approach, that is equally relevant for KCMs concerns the identification, at the molecular level, of the vacancies (in lattice gases), mobility defects (in spin facilitated models), or free volume itself. Attempts to provide reasonable coarse-graining from molecular models with continuous degrees of freedom to lattice models with kinetic rules are so far very limited, and not really convincing [70,163]. On the other hand the proof that kinetic rules can emerge effectively and induce a slow dynamics has been obtained for simple lattice spin models [86], whose dynamics directly maps onto constrained models. Several examples are available but here we only mention the simple case of the bidimensional plaquette model defined by a Hamiltonian of a  $p$ -spin type, but in two dimensions on a square lattice of linear size  $L$ ,

$$H = -J \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} S_{i,j} S_{i+1,j} S_{i,j+1} S_{i+1,j+1}, \quad (18)$$

where  $S_{i,j} = \pm 1$  is an Ising variable lying at node  $(i, j)$  of the lattice. Contrary to KCMs, the Hamiltonian in Eq. (18) contains genuine interactions, which are no less (or no more) physical than  $p$ -spin models discussed in

<sup>5</sup>A critical (different) behaviour is expected and predicted for models having a transition [158].

Sect. “[Cooperativity, Chaotic Energy Landscapes and Random First Order Theory](#)”. Interestingly the dynamics of this system is (trivially) mapped onto that of a KCM by analyzing its behaviour in terms of plaquette variables,  $p_{i,j} \equiv S_{i,j}S_{i+1,j}S_{i,j+1}S_{i+1,j+1}$ , such that the Hamiltonian becomes a non-interacting one,  $H = -J \sum_{i,j} p_{i,j}$ , as in Eq. (16). More interestingly, the analogy also applies to the dynamics [86]. The fundamental moves are spin-flips, but when a single spin is flipped the states of the four plaquettes surrounding that spin change. Considering the different types of moves, one quickly realizes that excited plaquettes,  $p_{i,j} = +1$ , act as sources of mobility, since the energetic barriers to spin flips are smaller in those regions. This observation allows to identify the excited plaquettes as defects, by analogy with KCMs. Spatially heterogeneous dynamics, diverging lengthscales accompanying diverging timescales and scaling behaviour sufficiently close to  $T = 0$  can be established by further analysis [101], providing a simple, but concrete example, of how an interacting many body system might effectively behave as a model with kinetic constraints<sup>6</sup>.

Another essential drawback of facilitated models is that among the microscopic ‘details’ thrown away to arrive at simple statmech models such as the ones in Eqs. (15) and (16), information on the thermodynamic behaviour of the liquids has totally disappeared. In particular, a possible coincidence between VFT and Kauzmann temperatures,  $T_0$  and  $T_K$  is not expected, nor can the dynamics be deeply connected to thermodynamics, as in Adam–Gibbs relations. The thermodynamic behaviour of KCMs appears different from the one of real glass-formers close to  $T_g$  [35]. This is probably the point where KCMs and RFOT approaches differ more evidently. Even though the dynamics of KCMs shares similarities with systems characterized with a complex energy landscape [25,169], thermodynamical behaviours are widely different in both cases, as has been recently highlighted in [102] by focusing on the concrete examples of plaquette models such as in Eq. (18).

Finally, when KCMs were first defined, they were argued to display a dynamic transition of a nature very similar to the one predicted by MCT [81]. Although the claim has been proven wrong<sup>7</sup>, it bears some truth: both approaches basically focus on the kinetic aspects of the glass transition and they both predict the existence of some dynamic criticality with diverging lengthscales and

<sup>6</sup>This type of plaquette models, and other spin models, were introduced originally [123,149] to show how ultra-slow glassy dynamics can emerge because of growing free energy barriers.

<sup>7</sup>Most KCMs do not have a finite temperature dynamical transition and the ones displaying a transition have critical properties different from MCT.

timescales. This similarity is even deeper, since a mode-coupling singularity is truly present when (some) KCMs are studied on the Bethe lattice [157], but is ‘avoided’ when more realistic lattice geometries are considered. This underlies the similarity of these two approaches while emphasizing further the mean-field character of the MCT approach.

### Geometric Frustration, Avoided Criticality, and Coulomb Frustrated Theories

In all of the above models, ‘real space’ was present in the sense that special attention was paid to different lengthscales characterizing the physics of the models that were discussed. However, apart from the ‘packing models’ with hard-core interactions, no or very little attention was paid to the geometric structure of local arrangements in molecular liquids close to a glass transition. This slight ‘oversight’ is generally justified using concepts such as ‘universality’ or ‘simplicity’, meaning that one studies complex phenomena using simple models, a typical statistical mechanics perspective. However, important questions remain: what is the liquid structure within mosaic states? How do different states differ? What is the geometric origin of the defects invoked in KCMs? Are they similar to defects (disclinations, dislocations, vacancies, etc.) found in crystalline materials?

There exists a line of research in this field which attempts to provide answers to these questions. It makes heavy use of the concept of geometric frustration. Broadly speaking, frustration refers to the impossibility of simultaneously minimizing all the interaction terms in the energy function of the system. Frustration might arise from quenched disorder (as in the spin glass models described above), but liquids have no quenched randomness. In that case, frustration has a purely geometric origin. It is attributed to a competition between a short-range tendency for the extension of a ‘locally preferred order’, and global constraints that prevent the periodic tiling of space with this local structure.

This can be illustrated by considering once more the packing problem of spheres in three dimensions. In that case, the locally preferred cluster of spheres is an icosahedron. However, the 5-fold rotational symmetry characteristic of icosahedral order is not compatible with translational symmetry, and formation of a periodic icosahedral crystal is impossible [75]. The geometric frustration that affects spheres in three dimensional Euclidean space can be relieved in curved space [139]. In Euclidian space, the system possesses topological defects (disclination lines), as the result of forcing the ideal icosahedral ordering into

a ‘flat’ space. Nelson and coworkers have developed a solid theoretical framework based on this picture to suggest that the slowing down of supercooled liquids is due to the slow wandering of these topological defects [139], but their treatment remains so complex that few quantitative, explicit results have been obtained.

This picture of sphere packing disrupted by frustration has been further developed in simple statistical models characterized by geometric frustration, in a pure statistical mechanics approach [155]. To build such a model, one must be able to identify, then capture, the physics of geometric frustration. Considering a locally ordered domain of linear size  $L$ , Kivelson et al. [114] suggest that the corresponding free energy scales as

$$F(L, T) = \sigma(T)L^2 - \phi(T)L^3 + s(T)L^5, \quad (19)$$

The first two terms expresses the tendency of growing local preferred order and they represent respectively the energy cost of having an interface between 2 phases and a bulk free energy gain inside the domain. Geometric frustration is encoded in the third term which represents the strain free energy resulting from the frustration. The remarkable feature of Eq. (19) is the super-extensive scaling of the energy cost due to frustration which opposes the growth of local order. The elements in Eq. (19) can then be directly incorporated into ferromagnetic models where ‘magnetization’ represents the local order, ferromagnetic interactions the tendency to local ordering, and Coulombic anti-ferromagnetic interactions the opposite effect of the frustration. The following Hamiltonian possesses these minimal ingredients:

$$H = -J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + K \sum_{i \neq j} \frac{\mathbf{S}_i \cdot \mathbf{S}_j}{|\mathbf{x}_i - \mathbf{x}_j|}, \quad (20)$$

where the spin  $\mathbf{S}_i$  occupies the site  $i$  at position  $\mathbf{x}_i$ . Such Coulomb frustrated models have been studied in great detail, using various approximations to study models for various space and spin dimensions [155].

The general picture is that the ferromagnetic transition occurring at  $T = T_c^0$  in the pure model with no frustration,  $K = 0$ , is either severely depressed to lower temperatures for  $K > 0$ , sometimes with a genuine discontinuity at  $K \rightarrow 0$ , yielding the concept of ‘avoided criticality’. For the simple case of Ising spins in  $d = 3$ , the situation is different since the second order transition becomes first order between a paramagnetic phase and a spatially modulated phase (stripes). For  $K > 0$  and  $T < T_c^0$  the system is described as a ‘mosaic’ of domains corresponding to some local order, whose size increases (but does not diverge!) when  $T$  decreases. Tarjus, Kivelson and co-workers clearly

demonstrate that such a structuration into mesoscopic domains allows one to understand most of the fundamental phenomena occurring in supercooled liquids [155]. Their picture as a whole is very appealing because it directly addresses the physics in terms of the ‘real space’, and the presence of domains of course connects to ideas such as cooperativity, dynamic heterogeneity and spatial fluctuations, that directly explains, at least qualitatively, non-exponential relaxation, decoupling phenomena or super-Arrhenius increase of the viscosity. However, as for the RFOT mosaic picture, direct confirmations of this scenario are rare [56]. Icosahedral order has not been clearly linked to the dynamics of hard spheres, while the very notion of local order in more realistic glass-forming liquids (such binary mixtures of spheres, or larger molecules with internal degrees of freedom) is problematic and not easily defined. This makes the basis of the scenario very fragile, and its practical applicability for a particular material difficult.

## Aging and Off-equilibrium

### Why Aging?

We have dedicated most of the above discussion to properties of materials approaching the glass transition at thermal equilibrium. We discussed a rich phenomenology and serious challenges for both our numerical and analytical capabilities to account for these phenomena. For most people, however, glasses are interesting below the glass transition, so deep in the glass phase that the material seems to be frozen forever in a seemingly arrested amorphous state, endowed with enough mechanical stability for a glass to retain, say, the liquid it contains (preferentially a nice red wine). Does this mean that there is no interesting physics in the glass state?

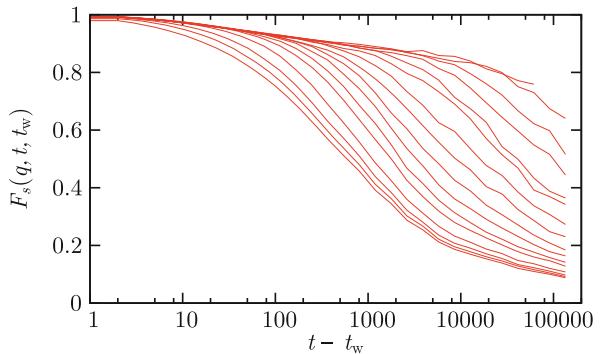
The answer is clearly ‘no’. There is still life (and physics) below the glass transition. We recall that for molecular glasses,  $T_g$  is defined as the temperature below which relaxation is too slow to occur within an experimental timescale. Much below  $T_g$ , therefore, the equilibrium relaxation timescale is so astronomically large that thermal equilibrium is out of reach. One enters therefore the realm of off-equilibrium dynamics. A full physical understanding of the non-equilibrium glassy state remains a central challenge [9,173].

A first consequence of studying materials in a time window smaller than equilibrium relaxation timescales is that the system can, in principle, remember its complete history, a most unwanted experimental situation since all details of the experimental protocol may then matter. The simplest protocol to study aging phenomena in the glass

phase is quite brutal [152]: take a system equilibrated above the glass transition and suddenly quench it at a low temperature at a ‘waiting time’  $t_w = 0$  which corresponds to the beginning of the experiment. For  $t_w > 0$  the system is left unperturbed at constant temperature where it tries to slowly reach thermal equilibrium, even though it has no hope to ever get there. Aging means that the system never forgets the time  $t_w$  spent in the glass phase, its ‘age’. The evolution of one time quantities, e. g. the energy, as a function of time are not a good evidence of aging. In order to show that the system never equilibrates two time quantities, such as density-density or spin-spin correlation functions, are much more useful. A typical example is presented in Fig. 14 where the self-part of the intermediate function in Eq. (2) is shown for a Lennard-Jones molecular liquid at low temperature. Immediately after the quench, the system exhibits a relatively fast relaxation: particles still move substantially. However, when the age of the system increases, dynamics slows down and relaxation becomes much slower. When  $t_w$  becomes very large, relaxation becomes too slow to be followed in the considered time window and the system seems frozen on that particular timescale: it has become a glass. A striking feature conveyed by these data is that an aging system not only remains out-of-equilibrium for all practical purposes, but its typical relaxation time is in fact set by its age  $t_w$ . In simple cases, the effective relaxation time after waiting a time  $t_w$  scales at  $t_w$  itself, which means that since equilibration timescales have diverged,  $t_w$  is the only remaining relevant timescale in the problem.

A popular interpretation of this phenomenon is given by considering trap models [40]. In this picture, reminiscent of the Goldstein view of the glass transition mentioned above [91], the system is described as a single particle evolving in a complex energy landscape with a broad distribution of trap depths—a paradigmatic mean-field approach. Aging in this perspective arises because the system visits traps that are increasingly deep when  $t_w$  increases, corresponding to more and more stable states. It takes therefore more and more time for the system to escape, and the dynamics slows down with time, as observed in Fig. 14. This implies that any physical property of the glass becomes an age-dependent quantity in aging protocols, and more generally dependent on how the glass was prepared. One can easily imagine using this property to tune mechanical or optical characteristics of a material by simply changing the way it is prepared, like how fast it is cooled to the glassy state.

A real space alternative picture was promoted in particular in the context of spin glass studies, based on the ideas of scaling and renormalization [42,74]. The physi-



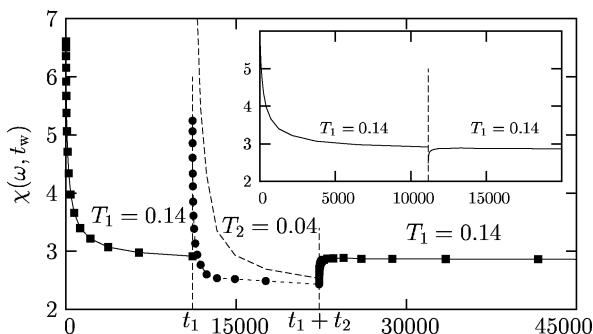
**Glasses and Aging, A Statistical Mechanics Perspective on, Figure 14**

Aging dynamics in a Lennard-Jones glass-forming liquid at low temperature. The system is quenched at time  $t_w = 0$  at low  $T$ , where the temperature is kept constant. Two-time self-intermediate scattering functions are then measured for 20 logarithmically spaced waiting times  $t_w$  from  $t_w = 1$  to  $t_w = 10^5$  (from left to right). The relaxation becomes slower when  $t_w$  increases: the system ages

cal picture is that of a coarsening process, where the system develops long-range order by growing extended domains of lengthscale  $\ell(t_w)$ . On lengthscales less than  $\ell(t_w)$  the system has ordered since the quench at  $t_w = 0$ . The domain walls evolve in a random environment. In order to move they have to overcome free energy barriers. It is then assumed an activated dynamic scaling which states that the typical barrier to extend the domain from linear size  $\ell(t_w)$  to, say,  $2\ell(t_w)$  scales as  $\ell^\psi$ , where  $\psi$  is some ‘barrier’ exponent. Using the Arrhenius law to relate dynamics to barriers, one gets that aging corresponds to the logarithmic growth with time of spatially correlated domains,  $\ell \sim (T \log t_w)^{1/\psi}$ . A domain growth picture of aging in spin glasses can be directly confirmed by numerical simulations [113], only indirectly by experiments.

### Memory and Rejuvenation Effects

Since the complete history of a sample in the glass phase matters, there is no reason to restrain experimental protocols to the simple aging experiment mentioned above. Indeed, experimentalists have investigated scores of more elaborated protocols that have revealed an incredibly rich, and sometimes quite unexpected, physics [173]. We restrain ourselves here to a short discussion of memory and rejuvenation effects observed during temperature cycling experiments [148] (one can imagine applying a magnetic field or a mechanical constraint, be they constant in time or sinusoidal, etc.). These two effects were first observed in spin glasses, but the protocol was then repeated in



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 15

Memory and rejuvenation effects obtained in the numerical simulation of a three-dimensional Heisenberg spin glass. There is a first aging step,  $0 < t_w < t_1$ , during which the system slowly tries to reach thermal equilibrium at temperature  $T_1$ . The system ‘rejuvenates’ in the second step at  $T_2$ ,  $t_1 < t < t_1 + t_2$ , and it restarts aging (rejuvenation). Finally in the third step, temperature is back to  $T_1$ , and memory of the first step is kept intact, as shown in the inset where relaxation during the second step of the experiment is taken away

many different materials, from polymers and organic liquids to disordered ferroelectrics. After several unsuccessful attempts, similar effects are now observed in numerical work as well. Results obtained from simulations of a three-dimensional Heisenberg spin glass [32] are presented in Fig. 15.

There are three steps in temperature cycling experiments [148]. The first one is a standard aging experiment, namely a sudden quench from high to low temperature at time  $t_w = 0$ . The system then ages for a duration  $t_1$  at constant temperature  $T_1$ . The system slowly relaxes towards equilibrium and its dynamics slows down, as observed in our spin glass example in Fig. 15 through the measurement of the magnetic susceptibility  $\chi(\omega, t_w)$ . Temperature is then suddenly shifted to  $T_2 < T_1$  at time  $t_1$ . There, the material restarts aging (almost) as if the first step had not taken place. This is called ‘rejuvenation effect’, because the system seems to forget it is already ‘old’. At total time  $t_1 + t_2$ , temperature is then shifted back to its initial value  $T_1$ . Then, aging is found to proceed as a quasi-perfect continuation of the first step, as if the second step had not taken place. The system has kept the ‘memory’ of the first part of the experiment, despite the rejuvenation observed in the intermediate part. The memory effect becomes more spectacular when relaxation during the second step is removed, as in the inset of Fig. 15. The third relaxation appears indeed as a perfect continuation of the first one.

On top of being elegant and quite intriguing, such protocols are relevant because they probe more deeply the dy-

namics of aging materials, allowing one to ask more precise questions beyond the simplistic observation that ‘this material displays aging’. Moreover, the observation of similar effects in many different glassy materials implies that these effects are intrinsic to systems with slow dynamics. Interesting also are the subtle differences observed from one material to the other.

Several experimental, numerical and theoretical papers have been devoted to this type of experiments, and these effects are not ‘mysterious’ anymore [31]. A clear link between memory effects and typical lengthscales over which the slow dynamics takes place has been established. Because lengthscales depend so sensitively on timescales and on the working temperature, experiments performed at two different temperatures typically probe very different lengthscales, allowing the system to store memory of its state at different temperatures at different lengthscales [23,39]. In return, this link has been elegantly exploited to obtain a rather precise experimental estimate of dynamic lengthscales involved in the aging dynamics of spin glass materials [15], which seems to confirm the slow logarithmic growth law mentioned before.

Discussion of the rejuvenation effect is slightly more subtle. It is indeed not yet obvious that the effect as it is observed in computer simulations and reported, e.g., in Fig. 15 is exactly similar to the one observed in experiments. The difficulty comes from the fact that some seemingly innocuous details of the experimental protocol, such as the necessary use in experiments of finite cooling rates, in fact play a crucial role and influence the physics so that direct comparison between experiments and simulations is difficult. In numerical work, rejuvenation can be attributed to a gradual change with temperature of the nature of spatial correlations between spins that develop with time [23,32]. More drastic changes are predicted to occur in disordered systems as a result of the chaotic evolution with temperature of the metastable states in a spin glass (so-called ‘chaos effect’ [43]), that could also be responsible for the observed rejuvenation effect [107]. This scenario can be directly discarded in simulations, where spatial correlations can be easily measured and chaos sought (in vain) in a very direct manner. Understanding the very origin of the rejuvenation effect observed in experiments remains, however, a challenge.

### Mean-Field Aging and Effective Temperatures

Theoretical studies of mean-field glassy models have provided important insights into the aging dynamics of both structural and spin glasses [60,61]. Although such models are defined in terms of spin degrees of freedom interact-

ing via infinite-ranged interactions, the deep connections between them and the mode-coupling theory of the glass transition make them serious candidates to investigate glassy states in general, not only thermodynamic properties at thermal equilibrium but also non-equilibrium aging dynamics. Despite their often reported ‘simplicity’, it took several years to derive a proper asymptotic solution of the long-time dynamics for a series of mean-field spin glasses [58]. These results have then triggered an enormous activity [57] encompassing theoretical, numerical and also experimental work trying to understand further these results, and to check in more realistic systems whether they have some reasonable range of applicability beyond mean-field. This large activity, by itself, easily demonstrates the broad interest of these results.

In these mean-field models, thermal equilibrium is never reached, and aging proceeds by downhill motion in an increasingly flat free energy landscape [119], with subtle differences between spin glass and structural glass models. In both cases, however, time translational invariance is broken, and two-time correlation and response functions depend on both their time arguments. In fact, the exact dynamic solution of the equations of motion for time correlators displays behaviours in strikingly good agreement with the numerical results reported in Fig. 14.

In these systems, the equations of motion in the aging regime involve not only time correlations, but also time-dependent response functions. At thermal equilibrium response and correlations are not independent, since the fluctuation-dissipation theorem (FDT) relates both quantities. In aging systems, there is no reason to expect the FDT to hold and both quantities carry, at least in principle, distinct physical information. Again, the asymptotic solution obtained for mean-field models quantitatively establishes that the FDT does not apply in the aging regime. Unexpectedly, the solution also shows that a generalized form of the FDT holds at large waiting times [60]. This is defined in terms of the two-time connected correlation function for some generic observable  $A(t)$ ,

$$C(t, t_w) = \langle A(t)A(t_w) \rangle - \langle A(t) \rangle \langle A(t_w) \rangle, \quad (21)$$

with  $t \geq t_w$ , and the corresponding two-time (impulse) response function

$$R(t, t_w) = T \frac{\delta \langle A(t) \rangle}{\delta h(t_w)} \Bigg|_{h=0}. \quad (22)$$

Here  $h$  denotes the thermodynamically conjugate field to the observable  $A$  so that the perturbation to the Hamiltonian (or energy function) is  $\delta E = -hA$ , and angled brackets indicate an average over initial conditions and any

stochasticity in the dynamics. Note that we have absorbed the temperature  $T$  in the definition of the response, for convenience. The associated generalized FDT reads then

$$R(t, t_w) = X(t, t_w) \frac{\partial}{\partial t_w} C(t, t_w), \quad (23)$$

with  $X(t, t_w)$  the so-called fluctuation-dissipation ratio (FDR). At equilibrium, correlation and response functions are time translation invariant, depending only on  $\tau = t - t_w$ , and equilibrium FDT imposes that  $X(t, t_w) = 1$  at all times. A parametric fluctuation-dissipation (FD) plot of the step response or susceptibility

$$\chi(t, t_w) = \int_{t_w}^t dt' R(t, t'),$$

against

$$\Delta C(t, t_w) = C(t, t) - C(t, t_w),$$

is then a straight line with unit slope. These simplifications do not occur in non-equilibrium systems. But the definition of an FDR through Eq. (23) becomes significant for aging systems [60,61]. In mean-field spin glass models the dependence of the FDR on both time arguments is only through the correlation function,

$$X(t, t_w) \sim X(C(t, t_w)), \quad (24)$$

valid at large wait times,  $t_w \rightarrow \infty$ . For mean-field structural glass models, the simplification (24) is even more spectacular since the FDR is shown to be characterized by only two numbers instead of a function, namely  $X \sim 1$  at short times (large value of the correlator) corresponding to a quasi-equilibrium regime, with a crossover to a non-trivial number,  $X \sim X^\infty$  for large times (small value of the correlator). This implies that parametric FD plots are simply made of two straight lines with slope 1 and  $X^\infty$ , instead of the single straight line of slope 1 obtained at equilibrium.

Since any kind of behaviour is in principle allowed in non-equilibrium situations, getting such a simple, equilibrium-like structure for the FD relations is a remarkable result. This immediately led to the idea that aging systems might be characterized by an effective thermodynamic behaviour and the idea of quasi-equilibration at different timescales [59]. In particular, generalized FD relations suggest to define an effective temperature, as

$$T_{\text{eff}} = \frac{T}{X(t, t_w)}, \quad (25)$$

such that mean-field glasses are characterized by a unique effective temperature,  $T_{\text{eff}} = T/X^\infty$ . It is thought of as the

temperature at which slow modes are quasi-equilibrated. One finds in general that  $0 < X^\infty < 1$ , such that  $T_{\text{eff}} > T$ , as if the system had kept some memory of its high temperature initial state.

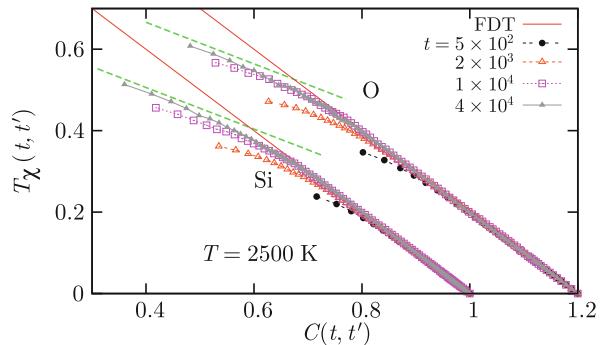
The name ‘temperature’ for the quantity defined in Eq. (25) is not simply the result of a dimensional analysis but has a deeper, physically appealing meaning that is revealed by asking the following questions. How does one measure temperatures in a many-body system whose relaxation involves well-separated timescales? What is a thermometer (and a temperature) in a far from equilibrium aging material? Answers are provided in Refs. [59, 120] both for mean-field models and for additional toy models with multiple relaxation timescales. The idea is to couple an additional degree of freedom, such as a harmonic oscillator,  $x(t)$ , which plays the role of the thermometer operating at frequency  $\omega$ , to an observable of interest  $A(t)$  via a linear coupling,  $-\lambda x(t)A(t)$ . Simple calculations show then that the thermometer ‘reads’ the following temperature,

$$\frac{1}{2}K_B T_{\text{meas}}^2 \equiv \frac{1}{2}\omega^2 \langle x^2 \rangle = \frac{\omega C'(\omega, t_w)}{2\chi''(\omega, t_w)}, \quad (26)$$

where  $C'(\omega, t_w)$  is the real part of the Fourier transform of Eq. (21), and  $\chi(\omega, t_w)$  the imaginary part of the Fourier transform of Eq. (22), with  $h = \lambda x$ . The relation (26) indicates that the bath temperature is measured,  $T_{\text{meas}} = T$ , if frequency is high and FDT is satisfied, while  $T_{\text{meas}} = T_{\text{eff}} > T$  if frequency is slow enough to be tuned to that of the slow relaxation in the aging material. The link between the FDR in Eq. (23) and the effective temperature measured in Eq. (26) was numerically confirmed in the computer simulation of a glassy molecular liquid in Ref. [19].

More generally, relaxation in glassy systems occurs in well-separated time sectors [61]; it is then easy to imagine that each sector could be associated with an effective temperature [120]. A thermodynamic interpretation of effective temperatures has also been put forward, relating them to the concept of replica symmetry breaking [78]. Interestingly, the full-step or one-step replica symmetry breaking schemes needed to solve the static problem in these models have a counterpart as the FDR being a function or a number, respectively, in the aging regime. Moreover, we note that these modern concepts are related to, but make much more precise, older ideas of quasi-equilibrium and fictive temperatures in aging glasses [152].

Taken together, these results make the mean-field description of aging very appealing, and they nicely complement the mode-coupling/RFOT description of the equilibrium glass transition described above. Moreover, they



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 16

Parametric correlation-response plots measured in the aging regime of a numerical model for a silica glass,  $\text{SiO}_2$  [30]. The plots for both species smoothly converge towards a two-straight line plot of slope 1 at short times (large  $C$  values), and of slope  $X^\infty \approx 0.51$  at large times (small values of  $C$ ), yielding an effective temperature of about  $T_{\text{eff}} = T/X^\infty \approx 4900$  K. Note that the glass transition temperature of  $\text{SiO}_2$  is 1446 K

have set the agenda for a large body of numerical and experimental work, as reviewed in [57]. In Fig. 16 we present recent numerical data obtained in an aging silica glass [30], presented in the form of a parametric response-correlation plot. The measured correlation functions are the self-part of the intermediate scattering functions defined in Eq. (2), while the conjugated response functions quantify the response of particle displacements to a spatially modulated field conjugated to the density. Plots for silicon and oxygen atoms at different ages of the system are presented. They seem to smoothly converge towards a two-straight line plot, as obtained in mean-field models (note, however, that this could be just a pre-asymptotic, finite “ $t_w$ ”, effect). Moreover, the second, non-trivial part of the plot is characterized by a slope that appears to be independent of the species, and of the wavevector chosen to quantify the dynamics, in agreement with the idea of a unique asymptotic value of the FDR, possibly related to a well-defined effective temperature.

### Beyond Mean-Field

Despite successes such as shown in Fig. 16, the broader applicability of the mean-field scenario of aging dynamics remains unclear, however. While some experiments and simulations indeed seem to support the existence of well-behaved effective temperatures [1, 92, 164], other studies also reveal the limits of the mean-field scenario. Experiments have for instance reported anomalously large FDT violations associated with intermittent

dynamics [11,12,44,45], while theoretical studies of model systems have also found non-monotonic or even negative response functions [68,118,140,162], and ill-defined or observable-dependent FDRs [73]. In principle, these discrepancies with mean-field predictions are to be expected, since there are many systems of physical interest in which the dynamics are not of mean-field type, displaying both activated processes and spatial heterogeneity.

It is thus an important task to understand from the theoretical point of view when the mean-field concept of an FDR-related effective temperature remains viable. However, studying theoretically the interplay between relevant dynamic lengthscales and thermally activated dynamics in the non-equilibrium regime of disordered materials is clearly a challenging task. Nevertheless, this problem has been approached in different ways, as we briefly summarize in this subsection.

A first class of system that displays aging and spatial heterogeneity is given by coarsening systems. The paradigmatic situation is that of an Ising ferromagnetic model (with a transition at  $T_c$ ) suddenly quenched in the ferromagnetic phase at time  $t_w = 0$ . For  $t_w > 0$  domains of positive and negative magnetizations appear and slowly coarsen with time. The appearance of domains that grow with time proves the presence of both aging and heterogeneity in this situation.

The case where the quench is performed down to  $T < T_c$  is well understood. The system becomes scale invariant [41], since the only relevant lengthscales is the growing domain size,  $\ell(t_w)$ . Correlation functions display aging, and scale invariance implies that  $C(t, t_w) \sim f(\ell(t)/\ell(t_w))$ . Response functions can be decomposed into two contributions [8,17]: one part stems from the bulk of the domains and behaves as the equilibrium response, and a second one from the domain walls and becomes vanishingly small in the long time limit where  $\ell(t_w) \rightarrow \infty$  and the density of domain walls vanishes. This implies that for coarsening systems in  $d \geq 2$ , one has  $X^\infty = 0$ , or equivalently an infinite effective temperature,  $T_{\text{eff}} = \infty$ . The case  $d = 1$  is special because  $T_c = 0$  and the response function remains dominated by the domain walls, which yields the non-trivial value  $X^\infty = 1/2$  [89].

Another special case has retained attention. When the quench is performed at  $T = T_c$ , there is no more distinction between walls and domains and the above argument yielding  $X^\infty = 0$  does not hold. Instead one studies the growth with time of critical fluctuations, with  $\xi(t_w) \sim t_w^{1/z}$  the correlation length at time  $t_w$ , where  $z$  is the dynamic exponent. Both correlation and response functions become non-trivial at the critical point [90]. It proves useful in that case to consider the dynam-

ics of the Fourier components of the magnetization fluctuations,  $C_q(t, t_w) = \langle m_q(t)m_{-q}(t_w) \rangle$ , and the conjugated response  $R_q(t, t_w) = (\delta \langle m_q(t) \rangle)(\delta h_{-q}(t_w))$ . From Eq. (23) a wavevector dependent FDR follows,  $X_q(t, t_w)$ , which has interesting properties [128] (see [47] for a review).

In dimension  $d = 1$ , it is possible to compute  $X_q(t, t_w)$  exactly in the aging regime at  $T = T_c = 0$ . An interesting scaling form is found, and numerical simulations performed for  $d > 1$  confirm its validity:

$$X_q(t, t_w) = X(q^2 t_w), \quad (27)$$

where the scaling function  $X(x)$  is  $X(x \rightarrow \infty) \rightarrow 1$  at small lengthscales,  $q\xi \gg 1$ , and  $X(x \rightarrow 0) \rightarrow 1/2$  (in  $d = 1$ ) at large distance,  $q\xi \ll 1$ ; recall that for  $z = 2$  in that case.

Contrary to mean-field systems where geometry played no role, here the presence of a growing correlation lengthscales plays a crucial role in the off-equilibrium regime since  $\xi(t_w)$  allows one to discriminate between fluctuations that satisfy the FDT at small lengthscales,  $X_q \sim 1$ , and those at large lengthscales which are still far from equilibrium,  $0 < X_q \sim X^\infty < 1$ . These studies suggest therefore that generalized fluctuation-dissipation relation in fact have a strong lengthscales dependence—a result which is not predicted by mean-field approaches.

Another interesting result is that the FDT violation for global observables (i.e. those at  $q = 0$ ) takes a particularly simple form, since the introduction of a single number is sufficient, the FDR at zero wavevector,  $X_{q=0}(t, t_w) \equiv X^\infty = 1/2$  (in  $d = 1$ ). This universal quantity takes non-trivial values in higher dimension, e.g.  $X^\infty \approx 0.34$  is measured in  $d = 2$  [128]. This shows that the study of global rather than local quantities makes the measurement of  $X^\infty$  much easier. Finally, having a non-trivial value of  $X^\infty$  for global observables suggests that the possibility to define an effective temperature remains valid, but it has become a more complicated object, related to global fluctuations on large lengthscales.

Kinetically constrained spin models represent a second class of non-mean-field systems whose off-equilibrium has been thoroughly studied recently [125]. This is quite a natural thing to do since these systems have local, finite ranged interactions, and they combine the interesting features of being defined in terms of (effective) microscopic degrees of freedom, having local dynamical rules, and displaying thermally activated and heterogeneous dynamics.

The case of the Fredrickson–Andersen model, described in Sect. “Some Theory and Models”, has been studied in great detail [125], and we summarize the main results. Here, the relevant dynamic variables are the Fourier

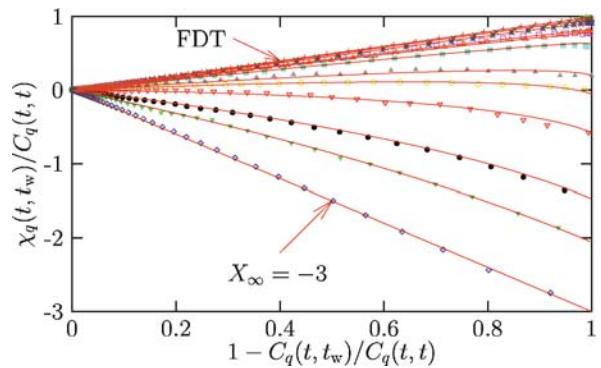
components of the mobility field, which also correspond in that case to the fluctuations of the energy density. Surprisingly, the structure of the generalized fluctuation-dissipation relation remains once more very simple. In particular, in dimension  $d > 2$ , one finds a scaling form similar to (27),  $X_q(t, t_w) = X(q^2 t_w)$ , with a well-defined limit at large distance  $X_{q=0}(t, t_w) \equiv X^\infty$ . The deep analogy with critical Ising models stems from the fact that mobility defects in KCMs diffuse in a way similar to domain walls in coarsening Ising models. It is in fact by exploiting this analogy that analytic results are obtained in the aging regime of the Fredrickson-Andersen model [130].

There is however a major qualitative difference between the two families of model. The (big!) surprise lies in the sign of the asymptotic FDR, since calculations show that [129]

$$X^\infty = -3, \quad d > 2.$$

In dimension  $d = 1$ , one finds  $X_{q=0}(t, t_w) = f(t/t_w)$  with  $X_{q=0}(t \rightarrow \infty, t_w) = (3\pi)(16 - 6\pi) \approx -3.307$ . Numerical simulations confirm these calculations. In Fig. 17, we show such a comparison between simulations (symbols) and theory (lines) in the case of the  $d = 3$  Fredrickson-Andersen model [129]. Fourier components of the mobility field yield parametric FD plots that follow scaling with the variable  $q^2 t_w$ , as a direct result of the presence of a growing lengthscale for dynamic heterogeneity,  $\xi(t_w) \sim \sqrt{t_w}$ . Again, generalized fluctuation-dissipation relations explicitly depend on the spatial lengthscale considered, unlike in mean-field studies. In Fig. 17, the limit  $q = 0$  corresponding to global observables is also very interesting since the plot is a pure straight line, as in equilibrium. Unlike equilibrium, however, the slope is not 1 but -3. A negative slope in this plot means a negative FDR, and therefore suggests a negative effective temperature, a very non-intuitive result at first sight.

Negative response functions in fact directly follow from the thermally activated nature of the dynamics of these models [129]. First, one should note that the global observable shown in Fig. 17 corresponds to fluctuations of the energy,  $e(t_w)$ , whose conjugated field is temperature. In the aging regime the system slowly drifts towards equilibrium. Microscopic moves result from thermally activated processes, corresponding to the local crossing of energy barriers. An infinitesimal change in temperature,  $T \rightarrow T + \delta T$  with  $\delta T > 0$ , accelerates these barrier crossings and makes the relaxation dynamics faster. The energy response to a positive temperature pulse is therefore negative,  $\delta e < 0$ , which directly yields  $\delta e / \delta T < 0$ , which explains the negative sign of the FDR. This result does not



Glasses and Aging, A Statistical Mechanics Perspective on, Figure 17

Parametric response-correlation plots for the Fourier components of the mobility field in the  $d = 3$  Fredrickson-Andersen model. Symbols are from simulations, lines from analytic calculations, and wavevectors decrease from top to bottom. The FDT is close to being satisfied at large  $q$  corresponding to local equilibrium. At larger distance deviations from the FDT are seen, with an asymptotic FDR which becomes negative. Finally, for energy fluctuations at  $q = 0$  (bottom curve), the plot becomes a pure straight line of (negative!) slope  $-3$ , as a result of thermally activated dynamics

hold in mean-field glasses, where thermal activation plays no role.

Finally, another scenario holds for local observables in some KCMs when kinetic constraints are stronger, such as the East model [125] or a bidimensional triangular plaquette model [100]. Here, relaxation is governed by a hierarchy of energy barriers that endow the systems with specific dynamic properties. In the aging regime following a quench, in particular, the hierarchy yields an energy relaxation that arises in discrete steps which take place on very different timescales, reminiscent of the ‘time sectors’ encountered in mean-field spin glasses. Surprisingly, it is found that to each of these discrete relaxations one can associate a well-defined (positive) value of the fluctuation-dissipation ratio, again reminiscent of the dynamics of mean-field spin glass models. Therefore, even in models that are very far from the mean-field limit the physical picture of a slow relaxation taking place on multiple timescales with each timescale characterized by an effective temperature seems to have some validity.

## Driven Glassy Materials

We have introduced aging phenomena with the argument that in a glass phase, the timescale to equilibrate becomes so long that the system always remembers its complete history. This is true in general, but one can wonder whether it

is possible to invent a protocol where the material history could be erased, and the system ‘rejuvenated’ [132]. This concept has been known for decades in the field of polymer glasses, where complex thermo-mechanical histories are often used.

Let us consider an aging protocol where the system is quenched to low temperature at time  $t_w = 0$ , but the system is simultaneously forced by an external mechanical constraint. Experimentally one finds that a stationary state can be reached, which explicitly depends on the strength of the forcing: a system which is forced more strongly relaxes faster than a less solicited material, a phenomenon called ‘shear-thinning’. The material has therefore entered a driven steady state, where memory of its age is no longer present and dynamics has become stationary: aging is stopped.

Many studies of these driven glassy states have been performed in recent years. In the language of the jamming phase diagram in Fig. 4, these correspond to studies of the (Temperature, Load) plane for molecular liquids, or the (1/Density, Load) plane for colloidal systems. The former studies are relevant for the rheology of supercooled liquids and glasses, and the  $T \ll T_g$  limit corresponds to studies of the plasticity of amorphous solids, a broad field in itself. In the colloidal world, such studies are also relevant for the newly-defined field of the rheology of ‘soft glassy materials’. These materials are (somewhat tautologically) defined as those for which the non-linear rheological behaviour is believed to result precisely from the competition between intrinsically slow relaxation processes and an external forcing [150]. It is believed that the rheology of dense colloidal suspensions, foams, emulsions, binary mixtures, or even biophysical systems are ruled by such a competition, quite a broad field of application indeed.

From the point of view of statmech modeling, soft glassy rheology can be naturally studied from the very same angles as the glass transition itself. As such trap models [150,151], mean-field spin glasses [18] and the related mode-coupling theory approach [83,134] have been explicitly extended to include an external mechanical forcing. In all these cases, one finds that a driven steady state can be reached and aging is indeed expected to stop at a level that depends on the strength of the forcing. Many of the results obtained in aging systems about the properties of an effective temperature are also shown to apply in the driven case, as shown both theoretically [18] and numerically [16]. A most interesting aspect is that the broad relaxation spectra predicted to occur in glassy materials close to a glass transition directly translate into ‘anomalous’ laws both for the linear rheological behaviour (seen

experimentally in the broad spectrum of elastic,  $G'(\omega)$ , and loss,  $G''(\omega)$ , moduli), and the non-linear rheological behaviour (a strong dependence of the viscosity  $\eta$  upon the shear rate  $\dot{\gamma}$ ).

## Future Directions

The problem of the glass transition, already very exciting in itself, has ramifications well beyond the physics of supercooled liquids. Glassy systems figures among the even larger class of ‘complex systems’. These are formed by a set of interacting degrees of freedom that show an emergent behaviour: as a whole they exhibit properties not obvious from the properties of the individual parts. As a consequence the study of glass-formers as statistical mechanics model characterized by frustrated interactions is a fertile ground to develop new concepts and techniques that will likely be applied to other physical, and more generally, scientific situations.

An example, already cited in this review, are the recent progress obtained in computer science and information theory [55] using techniques originally developed for spin glasses and structural glasses. More progress is certainly expected in the future along these interdisciplinary routes. Concerning physics, glassiness is such an ubiquitous and, yet as we showed, rather poorly understood problem that many developments are very likely to take place in the next decade.

Instead of guessing future developments of the field (and then very likely be proven wrong) we prefer to list a few problems we would like to see solved in the next years.

- Are the jamming transitions of granular media and colloids related to the glass transition of supercooled liquids? If yes, what is the common physical mechanism behind the dramatic slowing down?
- Is the glass transition related to a true phase transition? If yes, a static or a dynamic one? A finite or zero temperature one?
- Do RFOT, defects models, or frustration-based theory form the correct starting points of ‘the’ theory of the glass transition?
- Is MCT really a useful theory for the first decades of slowing down of the dynamics? Can one find direct evidence that an avoided MCT transition exists and controls the dynamics?
- What is the correct physical picture for the low temperature phase of glass-forming liquids and spin glasses?
- Are there general principles governing off-equilibrium equilibrium dynamics, and in particular aging and sheared materials?

- Do non-disordered, finite-dimensional, finite-range statmech model exist that display a thermodynamically stable amorphous phase at low temperature?

Finally, notice that we did not discuss possible interplays between glassiness and quantum fluctuations. This is a very fascinating topic. Quantum glassiness, and more generally, slow quantum dynamics are research subjects which are still in their infancies but that will likely undergo exciting developments in the near future.

### Acknowledgments

We thank C. Marchetti for inviting us to write this review and the collaborators who worked with us on glass physics. We thank J.-P. Bouchaud, A. Lefèvre, T. Sarlat for a careful reading of our manuscript and suggestions. Our work is supported by ANR Grants CHEF, TSANET and DYNHET.

### Bibliography

1. Abou B, Gallet F (2004) Phys Rev Lett 93:160603
2. Adam G, Gibbs JH (1958) J Chem Phys 43:139
3. Adhikari AN, Capurso NA, Bingemann D (2007) J Chem Phys 127:114508
4. Allen M, Tildesley D (1987) Computer Simulation of Liquids. Oxford University Press, Oxford
5. Angell CA (1997) J Res NIST 102:171
6. Angell CA (1995) Science 267:1924
7. Appignanesi GA, Rodriguez JA, Fris, Montani RA, Kob W (2006) Phys Rev Lett 96:057801
8. Barrat A (1998) Phys Rev E 57:3629
9. Barrat J-L, Dalibard J, Feigelman M, Kurchan J (eds) (2003) Slow relaxations and nonequilibrium dynamics in condensed matter. Springer, Berlin
10. Bassler H (1987) Phys Rev Lett 58:767
11. Bellon L, Ciliberto S, Laroche C (2001) Europhys Lett 53:511
12. Bellon L, Ciliberto S (2002) Physica D 168:325
13. Bengtzelius U, Götzte W, Sjölander A (1984) J Phys C 17:5915
14. Bennemann C, Donati C, Baschnagel J, Glotzer SC (1999) Nature 399:246; Lacevic N, Starr FW, Schroder TB, Glotzer SC (2003) J Chem Phys 119:7372
15. Bert F, Dupuis V, Vincent E, Hammann J, Bouchaud J-P (2004) Phys Rev Lett 92:167203
16. Berthier L, Barrat J-L (2002) J Chem Phys 116:6228
17. Berthier L, Barrat J-L, Kurchan J (1999) Eur Phys J B 11:635
18. Berthier L, Barrat J-L, Kurchan J (2000) Phys Rev E 61:5464
19. Berthier L, Barrat J-L (2002) Phys Rev Lett 89:095702
20. Berthier L, Biroli G, Bouchaud J-P, Cipelletti L, D El Masri, L'Hôte D, Ladieu F, Pierno M (2005) Science 310:1797
21. Berthier L, Biroli G, Bouchaud J-P, Kob W, Miyazaki K, Reichman DR (2007) J Chem Phys 126:184503
22. Berthier L, Biroli G, Bouchaud J-P, Kob W, Miyazaki K, Reichman DR (2007) J Chem Phys 126:184504
23. Berthier L, Bouchaud J-P (2002) Phys Rev B 66:054404
24. Berthier L, Chandler D, Garrahan JP (2005) Europhys Lett 69:320
25. Berthier L, Garrahan JP (2003) J Chem Phys 119:4367
26. Berthier L, Garrahan JP (2005) J Phys Chem B 109:3578
27. Berthier L, Kob W (2007) J Phys: Condens Matter 19:205130
28. Berthier L (2004) Phys Rev E 69:020201(R)
29. Berthier L (2007) Phys Rev E 76:011507
30. Berthier L (2007) Phys Rev Lett 98:220601
31. Berthier L, Viasnoff V, White O, Orlyanchik V, Krzakala F, in Reference [9]
32. Berthier L, Young AP (2005) Phys Rev B 71:214429
33. Binder K, Kob W (2005) Glassy materials and disordered solids. World Scientific, Singapore
34. Biroli G, Bouchaud J-P, Miyazaki K, Reichman DR (2006) Phys Rev Lett 97:195701
35. Biroli G, Bouchaud J-P, Tarjus G (2005) J Chem Phys 123:044510
36. Biroli G, Bouchaud JP (2004) Europhys Lett 67:21
37. Biroli G, Mézard M (2001) Phys Rev Lett 88:025501
38. Bouchaud J-P, Biroli G (2004) J Chem Phys 121:7347
39. Bouchaud J-P, Dupuis V, Hammann J, Vincent E (2001) Phys Rev B 65:024439
40. Bouchaud JP (1992) J Phys I France 2:1705
41. Bray AJ (1994) Adv Phys 43:357
42. Bray AJ, Moore MA (1984) J Phys C 17:L463; and in (1987) Heidelberg Colloquium on Glassy Dynamics In: van Hemmen JL, Morgenstern I (eds), Lectures Notes in Physics, vol 275. Springer, Berlin
43. Bray AJ, Moore MA (1987) Phys Rev Lett 58:57
44. Buisson L, Bellon L, Ciliberto S (2003) J Phys Condens Matter 15:S1163
45. Buisson L, Ciliberto S, Garcimartin A (2003) Europhys Lett 63:603
46. Butler S, Harrowell P (1991) J Chem Phys 95:4454 ((1991) J Chem Phys 95:4466)
47. Calabrese P, Gambassi A (2005) J Phys A 38:R133
48. Castellani T, Cavagna A (2005) J Stat Mech P05012
49. Cavagna A, Grigera TS, Verrocchio P (2007) Phys Rev Lett 98:187801
50. Chandler D, Garrahan JP, Jack RL, Maibaum L, Pan AC (2006) Phys Rev E 74:051501
51. Chaudhuri P, Berthier L, Kob W (2007) Phys Rev Lett 99:060604
52. Cheng Z, J, Chaikin PM, Phan S, Russel WB (2002) Phys Rev E 65:041405
53. Cohen MH, Grest GS (1982) Phys Rev B 26:6313
54. Coluzzi B, Verrocchio P (2002) J Chem Phys 116:3789
55. Mézard M, Bouchaud J-P, Dalibard J (eds) (2007) Complex systems. Springer, Berlin
56. Coslovich D, Pastore G (2007) J Chem Phys 127:124504
57. Crisanti A, Ritort F (2003) J Phys A 36:R181
58. Cugliandolo LF in Reference [9]
59. Cugliandolo LF, Kurchan J, Peliti L (1997) Phys Rev E 55:3898
60. Cugliandolo LF, Kurchan J (1993) Phys Rev Lett 71:173
61. Cugliandolo LF, Kurchan J (1994) J Phys A 27:5749
62. D'Anna G, Gremaud G (2001) Nature 413:407
63. Dalle-Ferrier C, Thibierge C, Alba-Simionescu C, Berthier L, Biroli G, Bouchaud J-P, Ladieu F, L'Hôte D, Tarjus G (2007) Phys Rev E 76:041510
64. Das SP, Mazenko GF (1986) Phys Rev A 34:2265
65. Dauchot O, Marty G, Biroli G (2005) Phys Rev Lett 95:265701
66. Debenedetti PG (1996) Metastable liquids. Princeton University Press, Princeton

67. Debenedetti PG, Stillinger FH (2001) *Nature* 410:259
68. Depken M, Stinchcombe R (2005) *Phys Rev E* 71:065102
69. Donati C, Douglas J, Kob W, Plimpton SJ, Poole PH, Glotzer SC (1998) *Phys Rev Lett* 80:2338
70. Dowd MT, Kennett MP (2007) *Phys Rev E* 76:031502
71. Dzero M, Schmalian J, Wolynes PG (2005) *Phys Rev B* 72:100201
72. Ediger MD (2000) *Annu Rev Phys Chem* 51:99
73. Fielding S, Sollich P (2002) *Phys Rev Lett* 88:050603
74. Fisher DS, Huse DA (1986) *Phys Rev Lett* 56:1601
75. Franck FC (1952) *Proc R Soc London* 215:43
76. Franz S, Donati C, Parisi G, Glotzer SC (1999) *Philos Mag B* 79:1827
77. Franz S, Mulet R, Parisi G (2002) *Phys Rev E* 65:021506
78. Franz S, Mézard M, Parisi G, Peliti L (1998) *Phys Rev Lett* 81:1758
79. Franz S, Parisi G (2000) *J Phys Condens Matter* 12:6335
80. Franz S, Stat J (2005) *Mech* P04001; (2006) *Europhys Lett* 73:492
81. Fredrickson GH, Andersen HC (1984) *Phys Rev Lett* 53:1244
82. Fredrickson GH, Brawer SA (1986) *J Chem Phys* 84:3351
83. Fuchs M, Cates ME (2002) *Phys Rev Lett* 89:248304
84. Garrahan JP, Chandler D (2002) *Phys Rev Lett* 89:035704
85. Garrahan JP, Chandler D (2003) *Proc Natl Acad Sc USA* 100:9710
86. Garrahan JP (2002) *J Phys Condens Matter* 14:1571
87. Glarum SH (1960) *J Chem Phys* 33:639
88. Gleim T, Kob W, Binder K (1998) *Phys Rev Lett* 81:4404
89. Godrèche C, Luck J-M (2000) *J Phys A* 33:1151; Lippiello E, Zannetti M (2000) *Phys Rev E* 61:3369
90. Godrèche C, Luck J-M (2000) *J Phys A* 33:9141
91. Goldstein M (1969) *J Chem Phys* 51:3728
92. Grigera TS, Israeloff NE (1999) *Phys Rev Lett* 83:5038
93. Gross J, Mézard M (1984) *Nucl Phys B* 240:431
94. Götze W (1999) *J Phys Condens Matter* 11:A1
95. Hansen JP, McDonald IR (1986) *Theory of Simple Liquids*. Academic, London
96. Hodge I (1997) *J Res NIST* 102:195
97. Holyst R (2001) *Physica A* 292:255
98. Horbach J, Kob W (2001) *Phys Rev E* 64:041503
99. Hurley MM, Harrowell P (1995) *Phys Rev E* 52:1694
100. Jack RL, Berthier L, Garrahan JP (2006) *J Stat Mech P12005*
101. Jack RL, Berthier L, Garrahan JP (2005) *Phys Rev E* 72:016103
102. Jack RL, Garrahan JP (2005) *J Chem Phys* 123:164508
103. Jack RL, Mayer P, Sollich P (2006) *J Stat Mech Theory Exp* P03006
104. Jaeger HM, Nagel SR, Behringer RP (1996) *Rev Mod Phys* 68:1259
105. Johari GP (2000) *J Chem Phys* 112:8958
106. Jung Y, Garrahan JP, Chandler D (2004) *Phys Rev E* 69: 061205
107. Jönsson PE, Mathieu R, Nordblad P, Yoshino H, H Aruga Katori, Ito A (2004) *Phys Rev B* 70:174402
108. Kauzmann AW (1948) *Chem Rev* 43:219
109. Kegel WK, van Blaaderen A (2000) *Science* 287:290
110. Keys AS, Abate AR, Glotzer SC, Durian DJ (2007) *Nat Phys* 3:260
111. Kirkpatrick TR, Thirumalai D (1987) *Phys Rev Lett* 58:2091; Kirkpatrick TR, Wolynes PG (1987) *Phys Rev A* 35:3072
112. Kirkpatrick TR, Thirumalai D, Wolynes PG (1989) *Phys Rev A* 40:1045
113. Kisker J, Santen L, Schreckenberg M, Rieger H (1996) *Phys Rev B* 53:6418
114. Kivelson D, Kivelson SA, Zhao X-L, Nussinov Z, Tarjus G (1995) *Physica A* 219:27
115. Kob W, Andersen HC (1993) *Phys Rev E* 48:4364
116. Kob W, Donati C, Plimpton SJ, Poole PH, Glotzer SC (1997) *Phys Rev Lett* 79:2827
117. Krzakala F, Montanari A, F Ricci-Tersenghi, Semerjian G, Zdeborová L (2007) *Proc Natl Acad Sci* 104:10318
118. Krzakala F (2005) *Phys Rev Lett* 94:077204
119. Kurchan J, Laloux L (1996) *J Phys A* 29:1929
120. Kurchan J (2005) *Nature* 433:222
121. Larson RG (1999) *The Structure and Rheology of Complex Fluids*. Oxford University Press, New York
122. Leutheusser E (1984) *Phys Rev A* 29:2765
123. Lipowski A, Johnston D, Espriu D (2000) *Phys Rev E* 62:3404
124. Liu AJ, Nagel SR (1998) *Nature* 396:21
125. Léonard S, Mayer P, Sollich P, Berthier L, Garrahan JP (2007) *J Stat Mech P07017*
126. Mapes MK, Swallen SF, Ediger MD (2006) *J Chem Phys* 124:054710
127. Marty G, Dauchot O (2005) *Phys Rev Lett* 94:015701
128. Mayer P, Berthier L, Garrahan JP, Sollich P (2003) *Phys Rev E* 68:016116
129. Mayer P, Léonard S, Berthier L, Garrahan JP, Sollich P (2006) *Phys Rev Lett* 96:030602
130. Mayer P, Sollich P (2007) *J Phys A* 40:5823
131. McCullagh GD, Cellai D, Lawlor A, Dawson KA, *Phys. Rev. E* 71:030102 (2005)
132. McKenna GB, Kovacs AJ (1984) *Polym Eng Sci* 24:1131
133. Menon N, Nagel SR, *Phys. Rev. Lett* 74:1230 (1995); Fernandez LA, Martin V-Mayor, Verrocchio P (2006) *Phys Rev E* 73:020501; and [49]
134. Miyazaki K, Reichman DR (2002) *Phys Rev E* 66:050501(R)
135. Monasson R (1995) *Phys Rev Lett* 75:2847
136. Mézard M, Parisi G (1999) *Phys Rev Lett* 82:747
137. Mézard M, Parisi G, Virasoro M (1988) *Spin Glass Theory and Beyond*. World Scientific, Singapore
138. Nauroth M, Kob W (1997) *Phys Rev E* 55:657
139. Nelson DR (2002) *Defects and Geometry in Condensed Matter Physics*. Cambridge University Press, Cambridge
140. Nicodemi M (1999) *Phys Rev Lett* 82:3734
141. Pan AC, Garrahan JP, Chandler D (2005) *Phys Rev E* 72:041106
142. Pardo LC, Lunkenheimer P, Loidl A (2007) *Phys Rev E* 76:030502(R)
143. Parisi G, Zamponi F (2005) *J Chem Phys* 123:144501
144. Pusey PN and van Meegen W (1986) *Nature* 320:340
145. Richert R, Angell CA (1998) *J Chem Phys* 108:9016
146. Ritort F, Sollich P (2003) *Adv Phys* 52:219
147. Rivoire O, Birol G, Martin OC, Mézard M (2004) *Eur Phys J B* 37:55
148. Réfrégier P, Vincent E, Hammam J, Ocio M (1987) *J Phys France* 48:1533
149. Sethna JP, Shore JD, Huang M (1991) *Phys Rev B* 44:4943
150. Sollich P, Lequeux F, Hebraud P, Cates ME (1997) *Phys Rev Lett* 78:2020
151. Sollich P (1998) *Phys Rev E* 58:738
152. Struik LCE (1978) *Physical aging in amorphous polymers and other materials*. Elsevier, Amsterdam
153. Szamel G, Flennery E (2004) *Europhys Lett* 67:779
154. Tarjus G, Kivelson D (1995) *J Chem Phys* 103:3071

155. Tarjus G, Kivelson SA, Nussinov Z, Viot P (2005) J Phys Condens Matter 17:R1143
156. Thalmann F (2002) J Chem Phys 116:3378
157. Toninelli C, Biroli G, Fisher DS (2004) Phys Rev Lett 92:185504
158. Toninelli C, Biroli G, Fisher DS (2006) Phys Rev Lett 96:035702
159. Toninelli C, Wyart M, Berthier L, Biroli G, Bouchaud J-P (2005) Phys Rev E 71:041505
160. Tracht U, Wilhelm M, Heuer A, Feng H, K Schmidt-Rohr, Spiess HW (1998) Phys Rev Lett 81:2727; Reinsberg SA, Qiu XH, Wilhelm M, Spiess HW, Ediger MD (2001) J Chem Phys 114:7299
161. Vidal Russell E, Israeloff NE (2000) Nature 408:695
162. Viot P, Talbot J, Tarjus G (2003) Fractals 11:185
163. Vogel M, Glotzer SC (2004) Phys Rev E 70:061504
164. Wang P, Song CM, Makse HA (2006) Nat Phys 2:526
165. Weeks E, Crocker JC, Levitt AC, Schofield A, Weitz DA (2000) Science 287:627
166. Weeks ER, Crocker JC, Weitz DA (2007) J Phys Condens Matter 19:205131
167. Whitelam S, Berthier L, Garrahan JP (2005) Phys Rev E 71:026128
168. Whitelam S, Berthier L, Garrahan JP (2004) Phys Rev Lett 92:185705
169. Whitelam S, Garrahan JP (2004) J Phys Chem B 108:6611
170. Wuttke J, Petry W, Pouget S (1996) J Chem Phys 105:5177
171. Xia XY, Wolynes PG (2000) Proc Natl Acad Sci USA 97:2990
172. Yamamoto R, Onuki A (1998) Phys Rev E 58:3515
173. Young AP (ed) (1998) Spin glasses and random fields. World Scientific, Singapore

## Gliders in Cellular Automata

CARTER BAYS

Department of Computer Science and Engineering,  
University of South Carolina, Columbia, USA

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Other GL Rules in the Square Grid](#)

[Why Treat All Neighbors the Same?](#)

[Gliders in One Dimension](#)

[Two Dimensional Gliders in Non-Square Grids](#)

[Three and Four Dimensional Gliders](#)

[Future Directions](#)

[Bibliography](#)

### Glossary

**Game of life** A particular cellular automaton (CA) discovered by John Conway in 1968.

**Neighbor** A neighbor of cell  $x$  is typically a cell that is in close proximity to (frequently touching) cell  $x$ .

**Oscillator** A periodic shape within a specific CA rule.

**Glider** A translating oscillator that moves across the grid of a CA.

**Generation** The discrete time unit which depicts the evolution of a CA.

**Rule** Determines how each individual cell within a CA evolves.

### Definition of the Subject

A cellular automaton is a structure comprising a grid with individual cells that can have two or more states; these cells evolve in discrete time units and according to a rule, which usually involves neighbors of each cell.

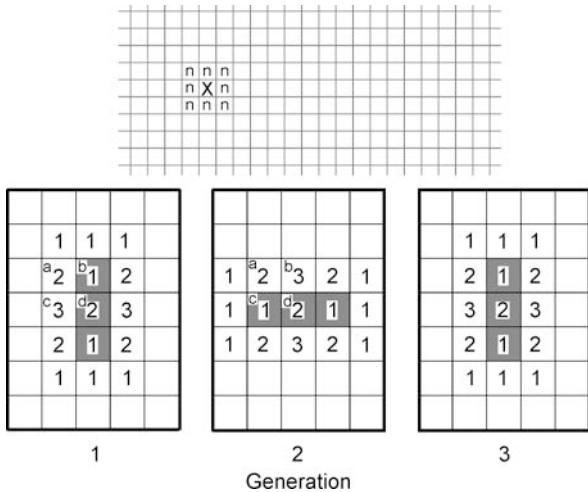
### Introduction

Although cellular automata has origins dating from the 1950s, interest in that topic was given a boost during the 1980s by the research of Stephan Wolfram, which culminated in 2002 with his publication of the massive tome, “A New Kind of Science” [11]. And widespread popular interest was created when John Conway’s “game of life” cellular automaton was initially revealed to the public in a 1970 Scientific American article [8]. The single feature of his game that probably caused this intensive interest was undoubtedly the discovery of “gliders” (translating oscillators). Not surprisingly, gliders are present in many other cellular automata rules; the purpose of this article is to examine some of these rules and their associated gliders.

Cellular automata (CA) can be constructed in one, two, three or more dimensions and can best be explained by giving a two dimensional example. Start with an infinite grid of squares. Each individual square has eight touching neighbors; typically these neighbors are treated the same (a *Moore neighborhood*), whether they touch a candidate square on a side or at a corner. (An exception is one dimensional CA, where position usually plays a role). We now fill in some of the squares; we shall say that these squares are alive. Discrete time units called generations evolve; at each generation we apply a rule to the current configuration in order to arrive at the configuration for the next generation; in our example we shall use the rule below.

- (a) If a live cell is touching two or three live cells (called neighbors), then it remains alive next generation, otherwise it dies.
- (b) If a non-living cell is touching exactly three live cells, it comes to life next generation.

Figure 1 depicts the evolution of a simple configuration of filled-in (live) cells for the above rule.



Gliders in Cellular Automata, Figure 1

**Top:** Each cell in a grid has eight neighbors. The cells containing  $n$  are neighbors of the cell containing the  $X$ . Any cell in the grid can be either **dead** or **alive**. **Bottom:** Here we have outlined a specific area of what is presumably a much larger grid. At the left we have installed an initial shape. Shaded cells are alive; all others are dead. The number within each cell gives the quantity of live neighbors for that cell. (Cells containing no numbers have zero live neighbors). Depicted are three generations, starting with the configuration at generation one. Generations two then three show the result when we apply the following cellular automata rule: *Live cells with exactly two or three live neighbors remain alive (otherwise they die); dead cells with exactly three live neighbors come to life (otherwise they remain dead)*. Let us now evaluate the transition from generation one to generation two. In our diagram, cell  $a$  is dead. Since it does not have exactly three live neighbors, it remains dead. Cell  $b$  is alive, but it needs exactly two or three live neighbors to remain alive; since it only has one, it dies. Cell  $c$  is dead; since it has exactly three live neighbors, it comes to life. And cell  $d$  has two live neighbors; hence it will remain alive. And so on. Notice that the form repeats every two generations. Such forms are called oscillators

There are many notations for describing CA rules; these can differ depending upon the type of CA. For CA of more than one dimension, and in our present discussion, we shall utilize the following notation, which is standard for describing CA in two dimensions with Moore neighborhoods. Later we shall deal with one dimension.

We write a rule as

$$E_1, E_2, \dots / F_1, F_2 \dots$$

where the  $E_i$  ("environment") specify the number of live neighbors required to keep a living cell alive, and the  $F_i$  ("fertility") give the number required to bring a non-living cell to life. The  $E_i$  and  $F_i$  will be listed in ascending order; hence if  $i > j$  then  $E_i > E_j$  etc.

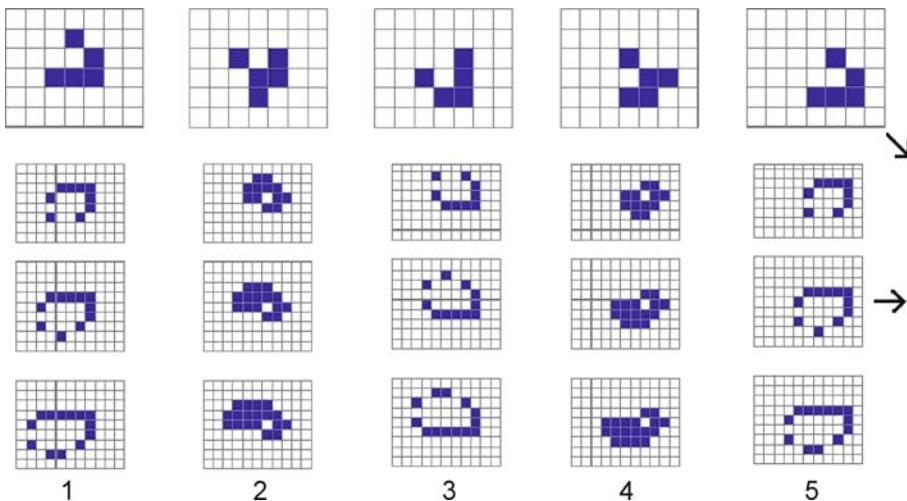
Thus the rule for the CA given above is 2,3/3. This rule, discovered by John Horton Conway, was examined in several articles in Scientific American and elsewhere, beginning with the seminal article in 1970 [8]. It is popularly known as Conway's game of life. Of course it is not really a game in the usual sense, as the outcome is determined as soon as we pick a starting configuration.

Note that the shape in Fig. 1 repeats, with a period of two. A repeating form such as this is called an oscillator. Stationary forms can be considered oscillators with a period of one. In Figs. 2 and 3 we show several oscillators that move across the grid as they change from generation to generation. Such forms are called translating oscillators, or more commonly, gliders. Conway's rule popularized the term; in fact a flurry of activity began during which a great many shapes were discovered and exploited. These shapes were named whimsically – "blinker" (Fig. 1), "boat", "beehive" and an unbelievable myriad of others. Most translating oscillators were given names other than the simple moniker glider – there were "lightweight spaceships", "puffer trains", etc. For this article, we shall call all translating oscillators gliders.

Of course rule 2,3/3 is not the only CA rule (even though it is the most interesting). Configurations under some rules always die out, and other rules lead to explosive growth. (We say that rules with expansive growth are unstable). We can easily find gliders for many unstable rules; for example Fig. 4 illustrates some simple constructs for rule 2/2. Note that it is practically impossible NOT to create gliders with this rule! Hence we shall only look at gliders for rules that stabilize (i. e. exhibit bounded growth) and eventually yield only zero or more oscillators. We call such rules GL (game of life) rules. Stability can be a rather murky concept, since there may be some carefully constructed forms within a GL rule that grow without bounds. Typically, such forms would never appear in random configurations. Hence, we shall informally define a GL rule as follows:

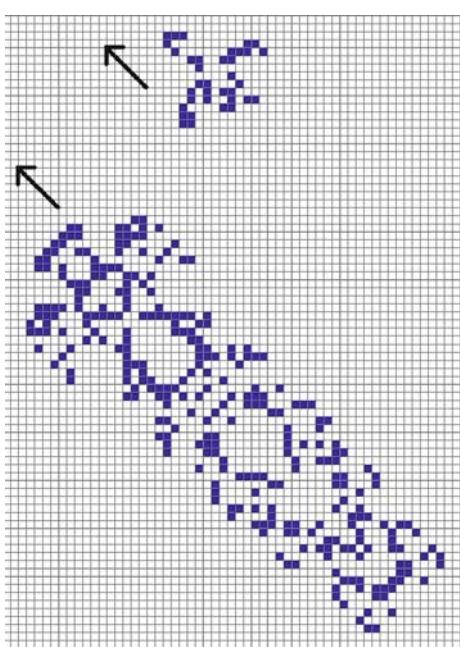
- All neighbors must be touching the candidate cell and all are treated the same (a Moore neighborhood).
- there must exist at least one translating oscillator (a glider).
- Random configurations must eventually stabilize.

This definition is a bit simplistic; for a more formal definition of a GL rule refer to [5]. Conway's rule 2,3/3 is the original GL rule and is unquestionably the most famous CA rule known. A challenge put forth by Conway was to create a configuration that would generate an ever increasing quantity of live cells. This challenge was met by William Gosper in 1970 – back when computing time was



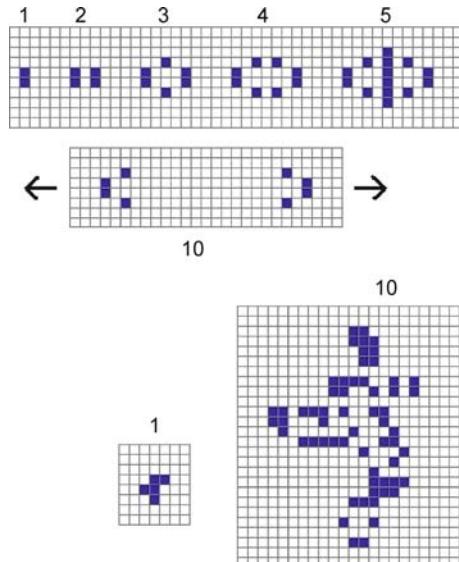
**Gliders in Cellular Automata, Figure 2**

Here we see a few of the small gliders that exist for 2,3/2. The form at the top – the original glider – was discovered by John Conway in 1968. The remaining forms were found shortly thereafter. Soon after Conway discovered rule 2,3/2 he started to give his various shapes rather whimsical names. That practice continues to this day. Hence, the name glider was given only to the simple shape at the top; the other gliders illustrated were called (from top to bottom) lightweight spaceship, middleweight spaceship and heavyweight spaceship. The numbers give the generation; each of the gliders shown has a period of four. The exact movement of each is depicted by its shifting position in the various small enclosing grids



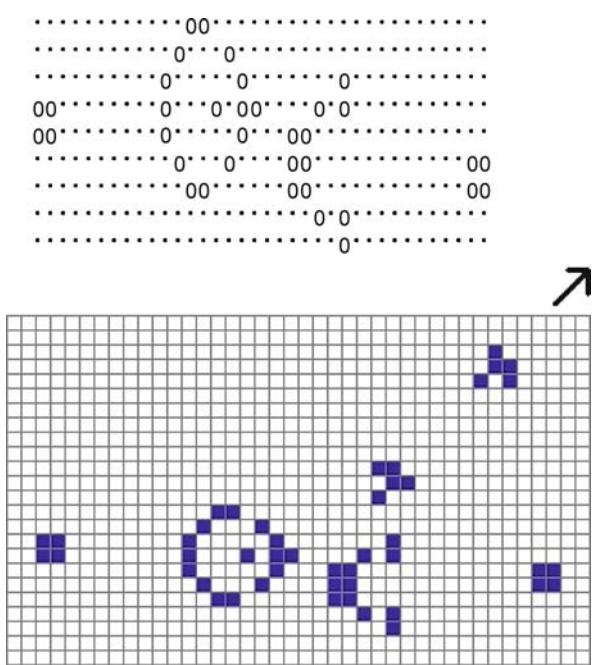
**Gliders in Cellular Automata, Figure 3**

The rule 2,3/2 is rich with oscillators – both stationary and translating (i.e. gliders). Here are but two of many hundreds of gliders that exist under this rule. The top form has a period of five and the bottom conglomeration, a period of four



**Gliders in Cellular Automata, Figure 4**

Gliders exist under a large number of rules, but almost all such rules are unstable. For example the rule 2/2 exhibits rapid unbounded growth, and almost any starting configuration will yield gliders; e.g. just two live cells will produce two gliders going off in opposite directions. But almost any small form will quickly grow without bounds. The form at the bottom left expands to the shape at the right after only 10 generations. The generation is given with each form



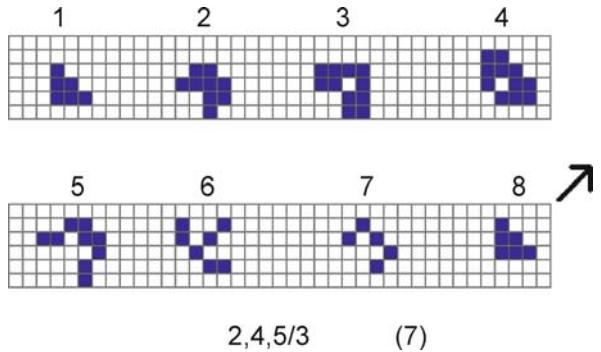
**Gliders in Cellular Automata, Figure 5**

A fascinating challenge was proposed by Conway in 1970 – he offered \$50 to the first person who could devise a form for 2,3/2 that would generate an infinite number of living cells. One such form could be a glider gun – a construct that would create an endless stream of gliders. The challenge was soon met by William Gosper, then a student at MIT. His glider gun is illustrated here. At the top, testifying to the primitive computational power of the time, is an early illustration of Gosper's gun. At the bottom we see the gun in action, sending out a new glider every thirty generations (here it has sent out two gliders). Since 1970 there have been numerous such guns that generate all kinds of forms – some gliders and some stationary oscillators. Naturally in the latter case the generator must translate across the grid, leaving its intended stationary debris behind

expensive and computers were slow by today's standards. He devised a form that spit out a continuous stream of gliders – a “glider gun”, so to speak. Interestingly, his gun configuration was displayed not as nice little squares, but as a rather primitive typewritten output (Fig. 5); this emphasizes the limited resources available in 1970 for seeking out such complex structures. Soon a cottage industry developed – all kinds of intricate initial configurations were discovered and exploited; such research continues to this day.

### Other GL Rules in the Square Grid

The rule 2, 4, 5/3 is also a GL rule and sports the glider shown in Fig. 6. It has not been seriously investigated and



**Gliders in Cellular Automata, Figure 6**

There are a large number of interesting rules that can be written for the square grid and Rule 2, 3/2 is undoubtedly the most fascinating – but it is not the only GL rule. Here we depict a glider that has been found for the rule 2, 4, 5/3. And since that rule stabilizes, it is a valid GL rule. Unfortunately it is not as interesting as 2, 3/2 because its glider is not as likely to appear in random (and other) configurations – hence limiting the ability of 2, 4, 5/3 to produce interesting moving configurations. Note that the period is seven, indicated in parentheses

will probably not reveal the vast array of interesting forms that exist under 2, 3/3. Interestingly, 2, 3/3, 8 appears to be a GL rule which not unsurprisingly supports many of the constructs of 2, 3/3. This ability to add terms of high neighbor counts onto known GL rules, obtaining other GL rules, seems to be easy to implement – particularly in higher dimensions or in grids with large neighbor counts such as the triangular grid, which has a neighbor count of 12.

### Why Treat All Neighbors the Same?

By allowing only Moore neighborhoods in two (and higher) dimensions we greatly restrict the number of rules that can be written. And certainly we could consider specialized neighborhoods – e. g. treat as neighbors only those cells that touch on sides, or touch only the left two corners and nowhere else, or touch anywhere, but state in our rule that two or more live neighbors of a subject cell must not touch each other, etc. But here we are only exploring gliders. Consider the following rule for finding the next generation.

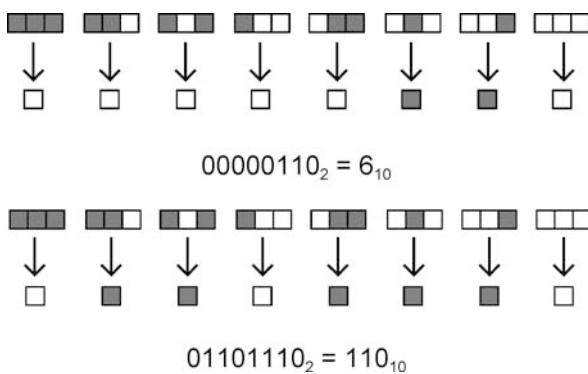
- 1) A living cell dies.
- 2) A dead cell comes to life if and only if its left side touches a live cell.

If we start, say, with a single cell we will obtain a glider of one cell that moves to the right one cell each generation! Such rules are easy to construct, as are more complex

glider-producing positional rules. So we shall not investigate them further. Yet as we shall see, the neighbor position is an important consideration in one dimensional CA.

### Gliders in One Dimension

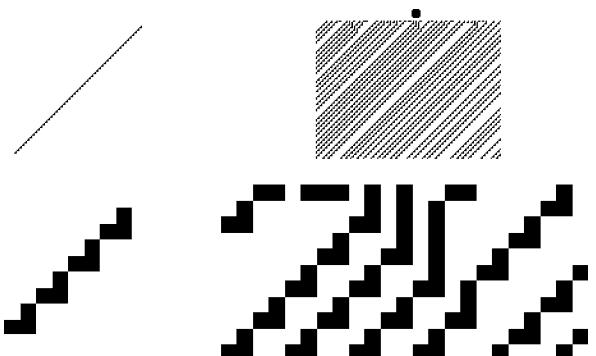
One dimensional cellular automata differ from CA in higher dimensions in that the restrictive grid (essentially a single line of cells) limits the number of rules that can be applied. Hence, many 1D CA involve neighborhoods that extend beyond the immediate two touching neighbors of a cell whose next generation status we wish to evaluate. Or more than the two states (alive, dead) may be utilized. For our discussion about gliders, we shall only look at the simplest rules – those involving just the two adjacent neighbors and two states. Unlike 2D (and higher) dimensions, we usually consider the relative position of the neighbors when giving a rule. Since three cells (center, left, right) are involved in determining the state for the next generation of the central cell, we have  $2^3 = 8$  possible initial states, with each state leading to a particular outcome. And since each initial state causes a particular outcome (i.e. the cell in the middle lives or dies next generation) we thus we have  $2^8$  possible rules. The behavior of these 256 rules has been extensively studied by Wolfram [11] who also introduced a very convenient shorthand that completely describes each rule (Fig. 7).



Gliders in Cellular Automata, Figure 7

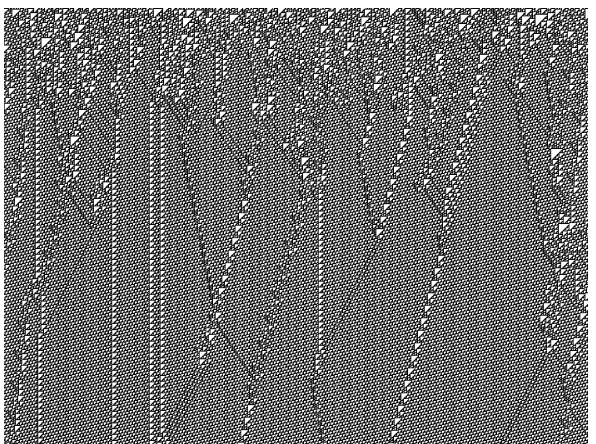
The one dimensional rules six and 110 are depicted by the diagram shown. There are eight possible states involving a center cell and its two immediate neighbors. The next generation state for the center cell depends upon the current configuration; each possible current state is given. The rule is specified by the binary number depicted by the next generation state of the center cell. This notation is standard for the simplest 1D CA and was introduced by Wolfram (see [11]), who also converts the binary representation to its decimal equivalent. There are 256 possible rules, but most are not as interesting as rule 110. Rule six is one of many that generate nothing but gliders (see Fig. 8).

As we add to the complexity of defining 1D CA we greatly increase the number of possible rules. For example, just by having three states instead of two, we note that now, instead of  $2^3$  possible initial states, there are  $3^3$  (Fig. 12). This leads to 27 possible initial states, and we now can create  $3^{27}$  unique rules – more than six trillion! Wolfram observed that even with more complex 1D rules, the fundamental behavior for all rules is typified by the simplest rules [11].



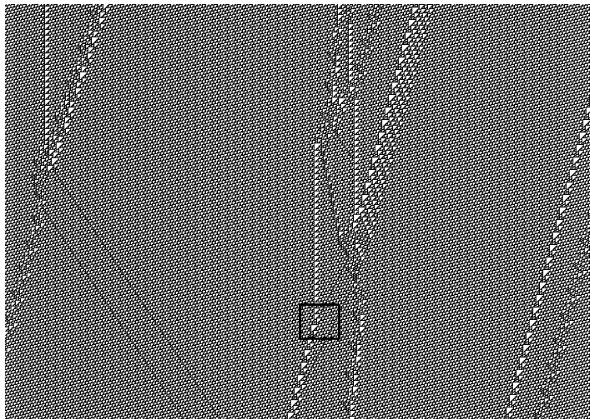
Gliders in Cellular Automata, Figure 8

Rule six (along with many others) creates nothing but gliders. At the *upper left*, we have several generations starting with a single live cell (*top*). (For 1D CA each successive generation moves vertically down one level on the page.) At the *lower left* is an enlargement of the first few generations. By following the diagram for rule six in Fig. 7, the reader can see exactly how this configuration evolves. At the *top right*, we start with a random configuration; at the *lower right* we have enlarged the small area directly under the large dot. Very quickly, all initial random configurations lead solely to gliders heading west



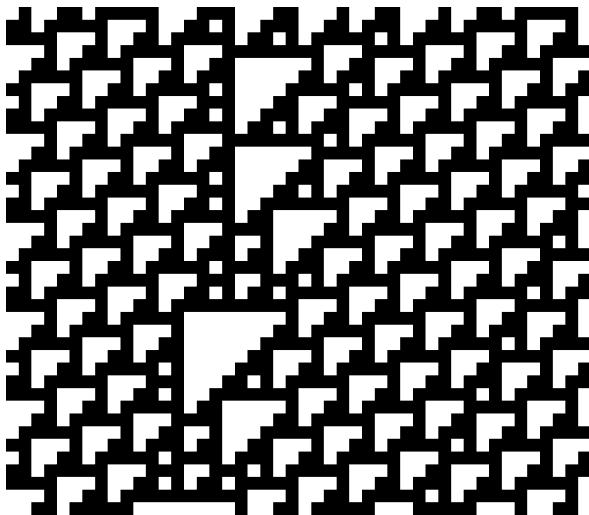
Gliders in Cellular Automata, Figure 9

Evolution of rule 110 for the first 500 generations, given a random starting configuration. With 1D CA, we can depict a great many generations on a 2D display screen



**Gliders in Cellular Automata, Figure 10**

Rule 110 at generations 2000–2500. The structures that move vertically are stationary oscillators; slanted structures can be considered gliders. Unlike higher dimensions, where gliders move in an unobstructed grid with no other live cells in the immediate vicinity, many 1D gliders reside in an environment of oscillating cells (the background pattern). The black square outlines an area depicted in the next figure



**Gliders in Cellular Automata, Figure 11**

An area from the previous figure enlarged. One can carefully trace the evolution from one generation to the next. The background pattern repeats every seven generations

Gliders in 1D CA are very common (Figs. 8 and 9) but true GL rules are not, because most gliders for stable rules exist against a uniform patterned background (Figs. 9 through 11) instead of a grid of non-living cells.

### Two Dimensional Gliders in Non-Square Grids

Although most 2D CA research involves a square grid, the triangular tessellation has been investigated somewhat.

■	■■	■■■	■■■■	■■■■■	■■■■■■	■■■■■■■	■■■■■■■■	■■■■■■■■■	■■■■■■■■■■
8	7	6	5	4	3	2	1	0	
■■	■■■	■■■■	■■■■■	■■■■■■	■■■■■■■	■■■■■■■■	■■■■■■■■■	■■■■■■■■■■	
17	16	15	14	13	12	11	10	9	
■■■	■■■■	■■■■■	■■■■■■	■■■■■■■	■■■■■■■■	■■■■■■■■■	■■■■■■■■■■	■■■■■■■■■■■	
26	25	24	23	22	21	20	19	18	

**Gliders in Cellular Automata, Figure 12**

There are 27 possible configurations when we have three states instead of two. Each configuration would yield some specific outcome as in Fig. 7; thus there would be three possible outcomes for each state, and hence  $3^{27}$  distinct rules

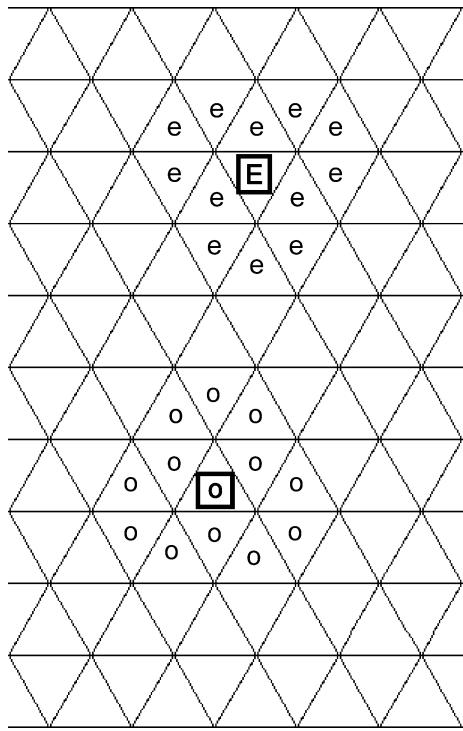
Here we have 12 touching neighbors; as with the square grid, they are all treated equally (Fig. 13). The increased number of neighbors allows for the possibility of more GL rules (and hence several gliders). Figure 14 shows many of these gliders and their various GL rules. The GL rule 2, 7, 8/3 supports two rather unusual gliders (Figs. 15 and 16) and to date is the only known GL rule other than Conway's original 2, 3/3 game of life that exhibits glider guns. Figure 17 shows starting configurations for two of these guns and Fig. 18 exhibits evolution of the two guns after 800 generations. Due to the extremely unusual behavior of the period 80 2, 7, 8/3 glider (Fig. 16), it is highly likely that other guns exist.

The hexagonal grid supports the GL rule 3/2, along with GL rules 3, 5/2, 3, 5, 6/2 and 3, 6/2, which all behave in a manner very similar to 3/2. The glider for these three rules is shown in Fig. 19. It is possible that no other distinct hexagonal GL rules exist, because with only six touching neighbors, the set of interesting rules is quite limited. Moreover the fertility portion of the rule must start with two and rules of the form  $*/2, 3$  are unstable. Thus, any other hexagonal GL rules must be of the form  $*/2, 4$ ;  $*/2, 4, 5$ ; etc. (i. e. only seven other fertility combinations).

A valid GL rule has also been found for at least one pentagonal grid (Fig. 19). Since there are several topologically unique pentagonal tessellations (see [10]), probably other pentagonal gliders will be found, especially when all the variants of the pentagonal grid are investigated.

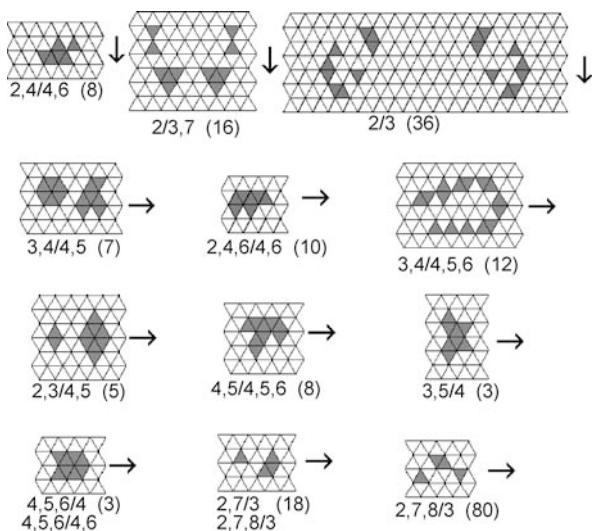
### Three and Four Dimensional Gliders

In 1987, the first GL rules in three dimensions were discovered [1,7]. The initially found gliders and their rules are depicted in Fig. 20. It turns out that the 2D rule 2, 3/3 is in many ways contained in the 3D GL rule 5, 6, 7/6. (Note the similarity between the glider at the bottom of Fig. 20 and at the top of Fig. 2). During the ensuing years, several other 3D gliders were found (Figs. 21 and 22). Most of these gliders were unveiled by employing random but



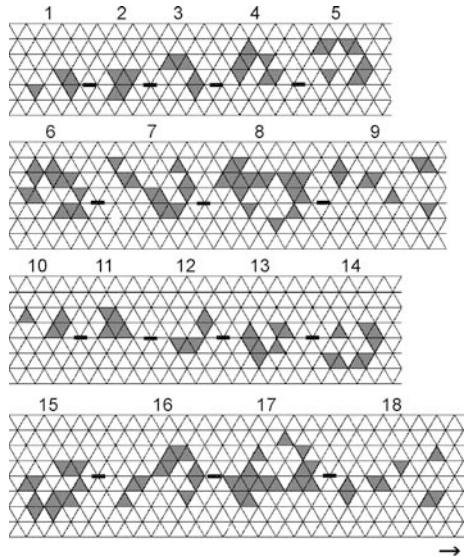
**Gliders in Cellular Automata, Figure 13**

Each cell in the triangular grid has 12 touching neighbors. The subject central cells can have two orientations, E and O



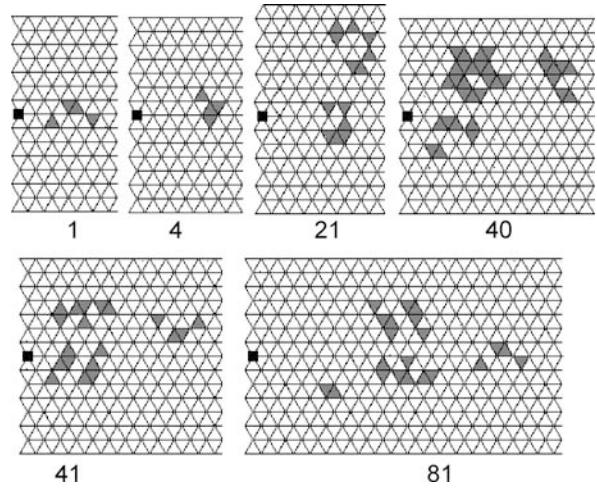
**Gliders in Cellular Automata, Figure 14**

Most of the known GL rules and their gliders are illustrated. The period for each is given in parentheses



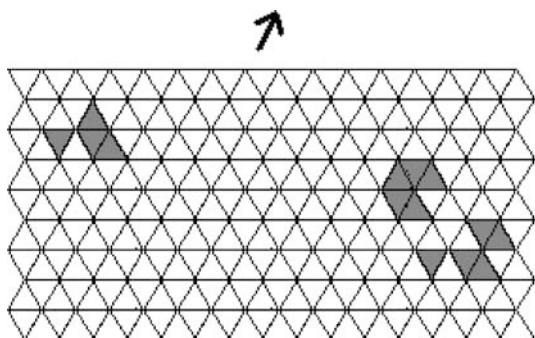
**Gliders in Cellular Automata, Figure 15**

The small 2, 7, 8/3 glider is shown. This glider also exists for the GL rule 2, 7/3. The small horizontal dash is for positional reference



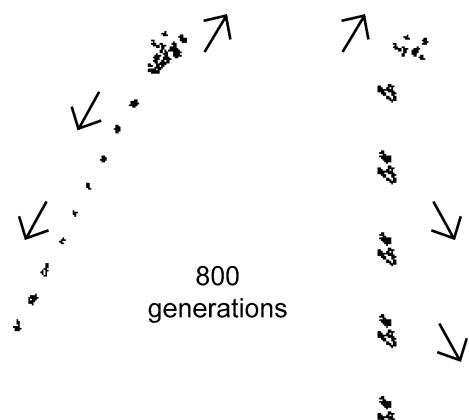
**Gliders in Cellular Automata, Figure 16**

Here we depict the large 2, 7, 8/3 glider. Perhaps *flamboyant* would be a better description, for this glider spews out much debris as it moves along. It has a period of 80 and its exact motion can be traced by observing its position relative to the black dot. Note that the debris tossed behind does not interfere with the 81st generation, where the entire process repeats 12 cells to the right. By carefully positioning two of these gliders, one can (without too much effort) construct a situation where the debris from both gliders interact in a manner that produces another glider. This was the method used to discover the two guns illustrated in Figs. 17 and 18



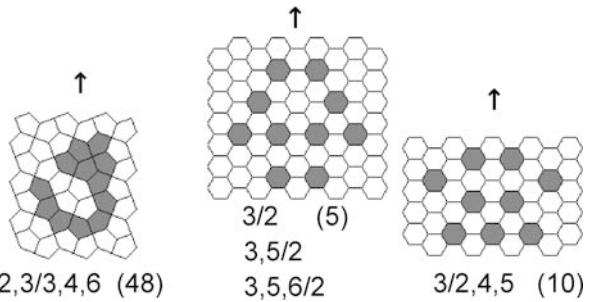
Gliders in Cellular Automata, Figure 17

The GL rule  $2, 7, 8/3$  is of special interest in that it is the only known GL rule besides Conway's rule that supports glider guns – configurations that spew out an endless stream of gliders. In fact, there are probably several such configurations under that rule. Here we illustrate two guns; the top one generates period 18 (small) gliders and the bottom one creates period 80 (large) gliders. Unlike Gosper's  $2, 3/3$  gun, these guns translate across the grid in the direction indicated. In keeping with the fanciful jargon for names, translating glider guns are also called "rakes"



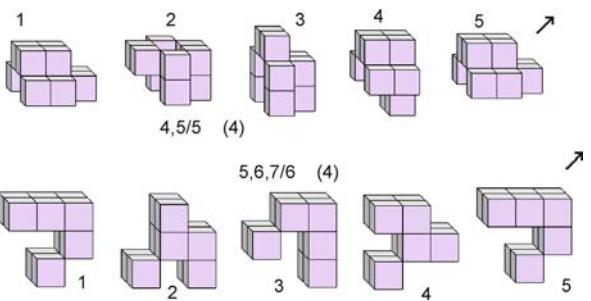
Gliders in Cellular Automata, Figure 18

After 800 generations, the two guns from Fig. 17 will have produced the output shown. Motion is in the direction given by the arrows. The gun at the left yields period 18 gliders, one every 80 generations, and the gun at the right produces a period 80 glider every 160 generations



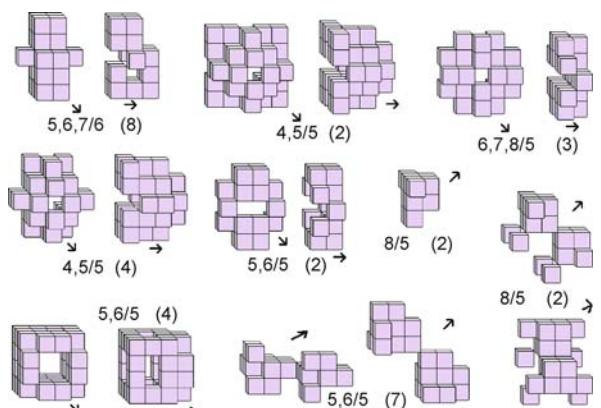
Gliders in Cellular Automata, Figure 19

GL rules are supported in pentagonal and hexagonal grids. The pentagonal grid (left) is called the Cairo Tiling, supposedly named after some paving tiles in that city. There are many different topologically distinct pentagonal grids; the Cairo Tiling is but one. At the right are gliders for the hexagonal rules  $3/2$  and  $3/2, 4, 5$ . The  $3/2$  glider also works for  $3, 5/2; 3, 5, 6/2$  and  $3, 6/2$ . All four of these rules are GL rules. The rule  $3/2, 4, 5$  is unfortunately disqualified (barely) as a GL rule because very large random blobs will grow without bounds. The periods of each glider are given in parentheses



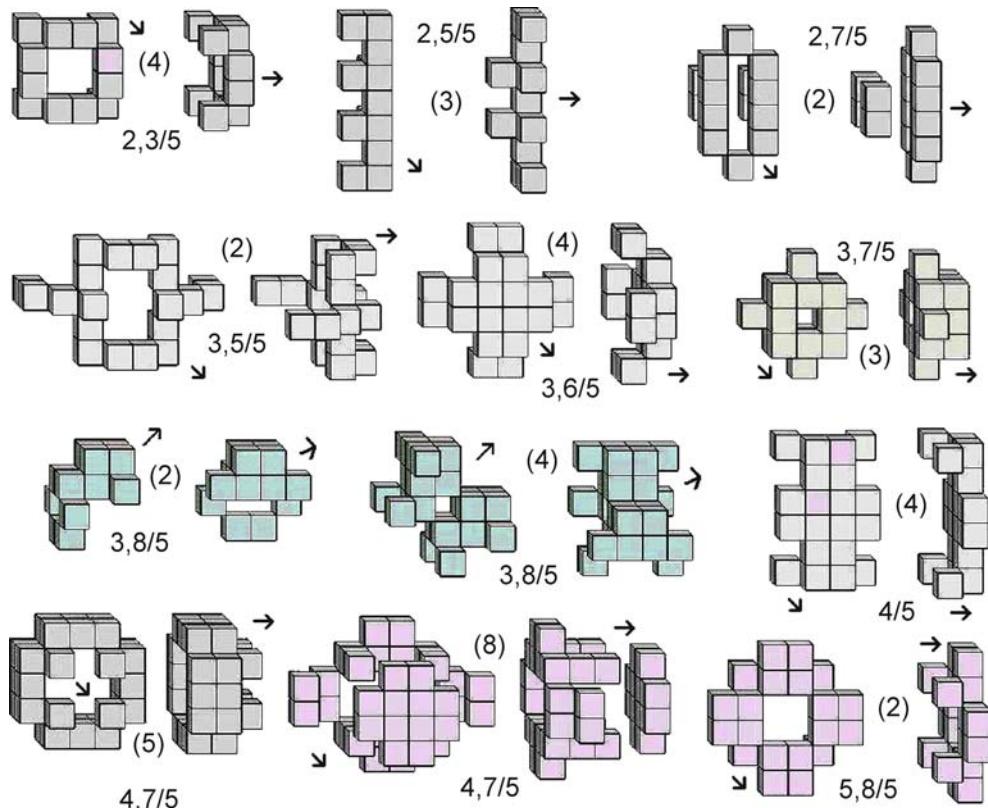
Gliders in Cellular Automata, Figure 20

The first three dimensional GL rules were found in 1987; these are the original gliders that were discovered. The rule  $5, 6, 7/6$  is analogous to the 2D rule  $2, 3/3$  (see [1]). Note the similarity between this glider and the one at the top of Fig. 2



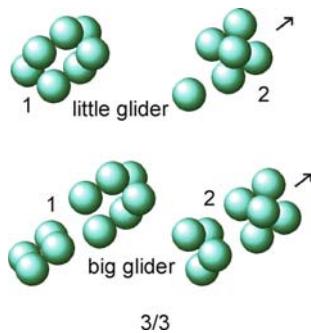
Gliders in Cellular Automata, Figure 21

Several more 3D GL rules were discovered between 1990–1994. They are illustrated here. The  $8/5$  gliders were originally investigated under the rule  $6, 7, 8/5$



**Gliders in Cellular Automata, Figure 22**

By 2004, computational speed had greatly increased, so another effort was made to find 3D gliders under GL rules; these latest discoveries are illustrated here



**Gliders in Cellular Automata, Figure 23**

Some work has been done with the 3D grid of dense packed spheres. Two gliders have been discovered for the rule 3/3, which almost qualifies as a GL rule

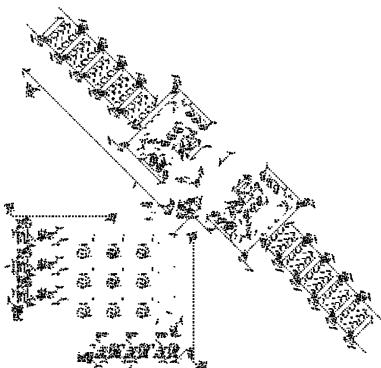
symmetric small initial configurations. The large number of live cells in these 3D gliders implies that they are uncommon random occurrences in their respective GL rules; hence it is highly improbable that the plethora of interest-

ing forms (e.g. glider guns) such as those for 2D rule 2, 3/3 exist in three dimensions.

The 3D grid of dense packed spheres has also been investigated somewhat; here each sphere touches exactly 12 neighbors. What is pleasing about this configuration is that each neighbor is identical in the manner that it touches the subject cell, unlike the square and cubic grids, where some neighbors touch on their sides and others at their corners. The gliders for spherical rule 3/3 are shown in Fig. 23. This rule is a borderline GL rule, as random finite configurations appear to stabilize, but infinite ones apparently do not.

### Future Directions

Gliders are an important by-product of many cellular automata rules. They have made possible the construction of extremely complicated forms – most notably within the universe of Conway's rule, 2, 3/3. (Figs. 24 and 25 illustrate a remarkable example of this complexity). Needless to say many questions remain unanswered. Can a glider gun be



**Gliders in Cellular Automata, Figure 24**

The discovery of the glider in 2, 3/3, along with the development of several glider guns, has made possible the construction of many extremely complex forms. Here we see a Turing machine, developed in 2001 by Paul Rendell. Figure 25 enlarges a small portion of this structure

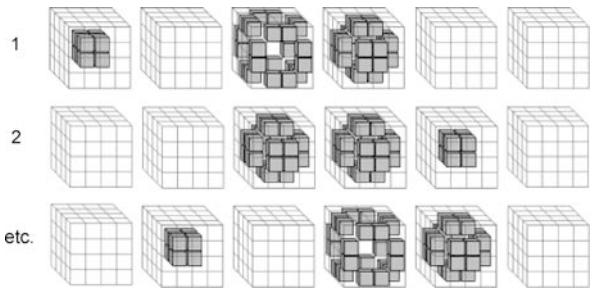


**Gliders in Cellular Automata, Figure 25**

We have enlarged a tiny portion at the upper left of the Turing machine shown in Fig. 24. One can see the complex interplay of gliders, glider guns, and various other stabilizing forms

constructed for some three dimensional rule? This would most likely be rule 5, 6, 7/6, which is the three dimensional analog of 2, 3/3 [7], but so far no example has been found.

The area of cellular automata research is more-or-less in its infancy – especially when we look beyond the square grid. Even higher dimensions have been given a glance; Fig. 26 shows just one of several gliders that are known to exist in four dimensions. Since each cell has 80 touching neighbors, it will come as no surprise that there are a large number of 4D GL rules. But there remains much work to be done in lower dimensions as well. Consider simple one dimensional cellular automata with four possible states. It will be a long time before all  $10^{38}$  possible rules have been investigated!



**Gliders in Cellular Automata, Figure 26**

Some work (not much) has been done in four dimensions. Here is an example of a glider for the GL rule 11, 12/12, 13. Many more 4D gliders exist

## Bibliography

1. Bays C (1987) Candidates for the Game of Life in Three Dimensions. *Complex Syst* 1:373–400
2. Bays C (1987) Patterns for Simple Cellular Automata in a Universe of Dense Packed Spheres. *Complex Syst* 1:853–875
3. Bays C (1994) Cellular Automata in the Triangular Tessellation. *Complex Syst* 8:127–150
4. Bays C (1994) Further Notes on the Game of Three Dimensional Life. *Complex Syst* 8:67–73
5. Bays C (2005) A Note on the Game of Life in Hexagonal and Pentagonal Tessellations. *Complex Syst* 15:245–252
6. Bays C (2007) The Discovery of Glider Guns in a Game of Life for the Triangular Tessellation. *J Cell Autom* 2(4):345–350
7. Dewdney AK (1987) The game Life acquires some successors in three dimensions. *Sci Am* 286:16–22
8. Gardner M (1970) The fantastic combinations of John Conway's new solitaire game 'Life'. *Sci Am* 223:120–123
9. Preston K Jr, Duff MJB (1984) Modern Cellular Automata. Plenum Press, New York
10. Sugimoto T, Ogawa T (2000) Tiling problem of convex pentagon[s]. *Forma* 15:75–79
11. Wolfram S (2002) A New Kind of Science. Wolfram Media, Champaign IL

## GPS: Applications in Crustal Deformation Monitoring

JESSICA MURRAY-MORALEDA  
US Geological Survey, Menlo Park, USA

### Article Outline

- Glossary
- Definition of the Subject
- Introduction
- Global Positioning System Measurements
- Applications of GPS Data to the Study of Seismic and Volcanic Hazards
- Future Directions

## Acknowledgments

## Bibliography

## Glossary

**Confidence ellipse** As applied to a vector representing a displacement or velocity estimate, the confidence ellipse defines the region within which the value is estimated at or above a specified confidence level (e.g., 95%). Confidence ellipses are computed by propagation of errors when computing the position. The ellipse is usually plotted at the tip of a GPS vector (e.g., Fig. 9).

**GNSS** Global Navigation Satellite System. Although this acronym stands for the same phrase as GLONASS, GNSS is a generic term referring to space-based navigation systems like the Global Positioning System (GPS) operated by the U. S., Russia's GLONASS, and the Galileo system under development by the European Union.

**Kinematic GPS** A method of collecting GPS data in which the receiver is continuously or intermittently in motion. This receiver, called the rover, can receive corrections for ambiguity resolution and common errors from a nearby stationary receiver.

### Interferometric synthetic aperture radar (InSAR)

A satellite-based imaging technique in which the satellite emits a radar signal and measures the phase of the returning signal after it has been scattered off the surface of the Earth. The difference in phase of the scattered waves measured during two passes of the same satellite can be used to produce a map of deformation, called an interferogram, that occurred during the time between the two satellite passes.

**International GNSS service (IGS)** An international consortium of agencies worldwide that provide data from permanent GPS and GLONASS sites in order to generate precise orbital and satellite clock parameters.

**Ionosphere** The electrically charged portion of the atmosphere from ~60 km to ~400 km above sea level. The ionosphere is dispersive, meaning that the degree to which it delays signal propagation depends on the signal's frequency and the electron content of the ionosphere.

**Mega-thrust earthquake** A type of earthquake which causes rupture of a long portion of the interface between a subducting plate and the over-riding plate. These earthquakes involve slip on a huge surface area, making them among the largest on Earth.

**Moment magnitude ( $M_w$ )** A magnitude scale used to compare the energy released in earthquakes. The mo-

ment magnitude is computed from the seismic moment. Therefore, because  $M_w$  accounts for the full rupture length of the earthquake, the moment magnitude scale does not saturate for large events in the way that other magnitude scales do.

**Reference frame** A terrestrial reference frame is defined by a set of points on Earth whose coordinates are precisely determined in a coordinate system with a specified origin and orientation of the axes. In order to compare GPS site positions, displacements, or velocities they must all be transformed into the same reference frame. For GPS, the most commonly used reference frame is the International Terrestrial Reference Frame (ITRF) which is updated periodically.

**Rupture** The slip that occurs during an earthquake. This term is often used in discussing the way in which the slip progresses with time over the fault surface, as in "the rupture front propagated southeast."

**Satellite laser ranging (SLR)** A geodetic technique for measuring the position of points on the surface of the Earth. Observation stations emit pulses of light that bounce off retroreflectors on satellites and return to the stations. The stations record the travel time of the light which is used to calculate a range measurement.

**Slip** The distance that material on one side of a fault moves relative to that on the other side.

**Stable North America** The stable interior portion of the North American continent that is not affected by plate boundary deformation. Often this term is used in the context of a "stable North American" reference frame, meaning that GPS velocities are transformed so that the velocities at stations considered to be in the stable interior of the continent are essentially zero. Because of factors such as Glacial Isostatic Adjustment (GIA), even some GPS sites in the continental interior have nonzero velocities. These sites are typically omitted when defining a stable North American reference frame.

**Strainmeter** An instrument that is capable of measuring change in distance over short baselines. These instruments typically come in two forms. The first is installed at the Earth's surface and uses a laser interferometer to measure the changes in distance over baseline lengths of 100s of meters. The second type is installed in a borehole 100s of meters deep to measure subtle changes in the diameter of the borehole. Some borehole strainmeters measure volumetric strain (e.g. the Sacks-Evertson strainmeter) and others measure three independent components of horizontal strain (e.g. the Gladwin tensor strainmeter).

**Strong motion seismograph** Seismic instrument de-

signed to record high-amplitude shaking near an earthquake rupture. These instruments typically record acceleration, and are sometimes called accelerometers. Data recording is often triggered by the arrival of the first seismic waves, and these instruments can record acceleration several times that of gravity.

**Telescismic** Refers to seismic waves recorded at distances greater than 3000 km from the epicenter.

**Troposphere** The portion of the atmosphere from the Earth's surface to ~15 km which delays GPS signal propagation. The degree to which the GPS signal is delayed depends on the spatially and temporally varying atmospheric pressure and water vapor content.

**Very long baseline interferometry (VLBI)** A geodetic positioning technique in which radio signals from distant sources such as quasars received at an array of antennas are used to calculate precise positions.

## Definition of the Subject

The Global Positioning System (GPS) is a space-based Global Navigation Satellite System (GNSS). Using signals transmitted by GPS satellites, the positions of ground-based receivers can be calculated to high precision, making it possible to track the movement of points on the Earth's surface over time. Unlike older geodetic surveying methods which involved periodically measuring angles, distances, or elevations between points, GPS can provide three-component (latitude, longitude, and altitude) position information at a range of sampling rates and on a global scale. GPS equipment is easy to use and can be set up to collect data continuously. Since its early geophysical applications in the mid-1980s, this versatile tool, which can be used to track displacements over time periods of seconds to decades, has become indispensable for crustal deformation studies, leading to many important insights and some surprising discoveries.

## Introduction

This article focuses on applications of GPS data to the study of tectonic, seismic, and volcanic processes. GPS has become a valuable tool for investigating other types of crustal deformation as well, including landslides (e.g., [25,50,100,128,151]), global sea-level change [150], and the ongoing rebound (termed Glacial Isostatic Adjustment or GIA) of the Earth's crust since the retreat of the ice sheets which covered much of North America and northern Europe during the last ice age (e.g., [19,80,88,95,117,141]), but these topics are beyond the scope of this article. The discussion presented here begins with an overview of how GPS works and how it is

used to collect data for geophysical studies. The rest of the paper describes a variety of ways in which GPS data have been used to measure crustal deformation and investigate the underlying processes, as illustrated by examples from the literature. Since GPS is so widely used in geophysical studies, examples of many more applications exist, and the reader is encouraged to explore the literature for more information.

## Global Positioning System Measurements

### How GPS works

The US Department of Defense developed GPS to provide positioning and timing information, primarily for military purposes, that would be available any time of day, anywhere on Earth, regardless of weather conditions. The first GPS satellites were launched in 1978. Soon afterward the Soviet Union developed a similar system, called GLONASS (which, like the generic acronym GNSS, also stands for Global Navigation Satellite System), and more recently the European Space Agency has designed a satellite navigation system called Galileo which, unlike its predecessors, is dedicated to civilian and commercial, rather than military, use. The rest of this article will focus on GPS. The scope of this article permits only a brief overview. Dzurisin [40] gives a broader discussion with a focus on applications in volcanic investigations. Hofmann-Wellenhof et al. [56] give a thorough treatment of the technical details.

The GPS satellite constellation nominally consists of 24 satellites, as well as several spares. The satellites orbit 20,200 km above the Earth with orbital periods of nearly 12 hours, and each passes over a given point on the Earth's surface once per sidereal day (which is about four minutes shorter than a solar day). From any point on the Earth's surface, at any given time, from four to ten satellites are above the horizon (and thus potentially visible). Each satellite remains visible for approximately five out of every 12 hours [40].

The idea behind satellite positioning is that one can determine the distance between a receiver on the ground and an orbiting satellite from the time it takes a signal to travel from the satellite to the receiver. This calculation, therefore, requires a means for precise time-keeping according to a universally accepted standard. The US Naval Observatory defines "GPS time," and GPS specifications require GPS time to be within one microsecond of Coordinated Universal Time (UTC). The difference between a satellite's or receiver's internal clock and GPS time (due to clock drift) is termed "clock bias" and is accounted for in processing GPS data.

GPS satellites broadcast signals on two carrier frequencies termed L1 and L2 in the microwave band. The “coarse acquisition” (C/A) code is modulated on the L1 carrier, and the precise (P) code is modulated on both L1 and L2. A navigation message, containing information about the satellite orbits, clocks (the time given by the satellite’s clock and information about the difference between that satellite’s time and GPS time), and state of health, as well as ionospheric conditions, is modulated on both carriers.

The receiver “locks on” to a satellite by generating a replica of one or more of the codes modulated on the satellite signal it receives and continually cross-correlating the internally generated code with that received from the satellite until the two match. Once it has locked on to the satellite it can obtain the navigation message, determine the signal travel time, and measure the carrier signal. The apparent distance, or “pseudorange,” between the GPS antenna and the satellite antenna is then calculated by multiplying the time it takes for the signal transmitted by the satellite to travel to the receiver by the speed of light. The term pseudorange emphasizes that the travel time used in this calculation includes the effects of satellite and receiver clock biases as well as a variety of other error sources, some of which may be mitigated during processing, and is therefore not equivalent to the true geometric range between the satellite and receiver. It is possible for the receiver to measure the code to a precision of about 1% of its length (293 meters for C/A-code and 29.3 meters for P-code), which results in 3 meter and 30 cm precision in calculated pseudorange. However, the P-code is generally encrypted by the military, called “anti-spoofing” or A-S, and therefore a civilian receiver cannot use this code for positioning.

The C/A code is modulated on L1, and thus this carrier can be easily measured once the receiver has locked on to the satellite using the C/A code. Because of A-S, in order to obtain the L2 carrier on which the encrypted P-code is modulated, civilian users must have receivers that apply more sophisticated signal processing techniques (see pp. 81–85 in [56]). The L1 carrier has a wavelength of 19 cm and the L2, 24.4 cm. Since the receiver can measure the carrier signal to 1% of a cycle length, much greater resolution can be achieved using the carrier phase measurements, rather than the code information, to calculate the pseudorange.

When the carrier signal is used, the satellite-receiver distance is calculated by multiplying the number of carrier cycles (which is generally not an integer) between the satellite and the receiver by the wavelength of the carrier signal. However, when a receiver locks on to a satellite, it

only can measure the initial fraction of a carrier cycle that it receives. Although it measures the number of full cycles thereafter (which changes as the satellite moves overhead), the receiver has no way of knowing how many full cycles in addition to the initial fraction were between it and the satellite to begin with. This unknown number of cycles is often called the integer ambiguity and will be different for each satellite–receiver pair. In order to take advantage of the more precise positioning that can be achieved using the carrier signal, processing techniques have been developed to address the problem of integer ambiguities (see [40] for an overview and [56] for more detail).

The positions of the satellites at any given time are given by their orbital parameters. This information is transmitted as part of the navigation message, however more precise orbital information, available from the International GNSS Service (IGS), is used in scientific applications. With the satellite positions assumed known, once the distance from the receiver to at least four satellites is measured, the position in three coordinate dimensions (e.g., north, east, and vertical) of a GPS antenna on the ground can be found. Four satellites are necessary in order to solve for the three coordinate positions and the satellite and receiver clock bias. However, positioning accuracy is greatly improved with data from additional satellites as it is then possible to estimate some unknown noise sources.

Although GPS receivers are capable of determining a position in real time using internal software, for scientific applications the data are generally downloaded, and one of several processing software packages is used to obtain much more precise positions. This type of software allows the user more control over the way in which the data are processed, for instance by enabling the use of precise orbital parameters, the fixing of ambiguities, and the application of sophisticated models for atmospheric delay and variations in the antenna phase center (the part of the antenna that actually receives the GPS signal).

It is possible to reduce or eliminate certain error sources by differencing data. These errors include atmospheric delays to signal propagation that affect neighboring GPS stations (e.g. within 10 s of km of each other) in a similar way, satellite orbital and clock errors that will be common to data from the same satellite recorded by more than one station at the same time, and receiver clock errors that are common to measurements to multiple satellites made by that receiver at the same time.

Other means of addressing error sources exist as well. For example, the ionosphere is dispersive, meaning the delay in signal propagation that it causes depends on the frequency of the signal. Because GPS uses two frequen-

cies, the effect of this delay can be eliminated from the data. Tropospheric delay is addressed during data processing through a combination of models for the “dry” component (which is dependent on atmospheric pressure, temperature, and elevation), and treatment of the “wet” component (which shows large variability depending on water vapor content) as a stochastic parameter to be estimated. Multipath (when the GPS signal bounces off something before reaching the ground antenna) and set-up error (commonly, operator error in measuring the height of the antenna) are mitigated by antenna design, choice of station location, and surveying technique.

The military has used two means of limiting civilian access to GPS positioning capabilities. The first, called selective availability (SA), involved degrading the accuracy of the satellite orbit information broadcast by the satellites and introducing noise into the satellite clock information. This caused an approximately ten-fold increase in positioning error. Using information from a global network of continuously operating GPS stations it is possible to calculate more accurate satellite orbits for use in post-processing of GPS data. During processing, clock errors can be eliminated by differencing data from multiple stations or estimated along with station positions. Therefore, SA did not pose an insurmountable obstacle to scientific users of GPS. In May 2000 the US government discontinued SA, although it reserves the right to reinstate it if deemed necessary. The other means by which the military can reduce civilian access to GPS data is through A-S. A-S is implemented by encrypting the P-code such that only military users can decipher it, thus preventing outside parties from sending phony “GPS” signals that would prevent accurate positioning. A-S began in 1994 and continues to the present, but its effect on the accuracy of GPS positioning for scientific purposes is generally small.

### Methods of GPS Data Collection for Crustal Deformation Studies

For some geological and geophysical applications, typically those which require mapping or recording of sample locations, kinematic GPS methods are used. In these cases, a base station is used to generate corrections that are applied to measurements made by a “roving” antenna which is moved to each site of interest [40]. However, for the crustal deformation applications described in this article, the GPS antenna remains at one location for an extended period of time (usually at least eight hours). Measurements of this type are often classified as “campaign” or “continuous,” depending on the frequency of measurement and the way in which the receiver and antenna are installed. Here

I give a brief description of each; Blewitt et al. [13] give a detailed discussion as well.

Before the advent of GPS as a tool for monitoring crustal deformation, other types of geodetic observations were made by periodically measuring triangulation, trilateration and leveling networks using traditional land surveying techniques. The objective was to see how the movement of the Earth’s surface changed over time, since this reflects a variety of geophysical processes. However, these techniques did not produce 3-component position data, but rather measurements of angles, distances, or elevation changes.

The surveying approach in the early years of GPS monitoring (e.g. mid-1980s to mid-1990s) was similar to that of its predecessors in that instruments were deployed temporarily, for several hours at a time, once a year or so. This approach is called “campaign” or “survey-mode” GPS (SGPS). At that time receivers were prohibitively expensive, limiting their wide-spread use. Also, there were fewer satellites in orbit, and one had to schedule surveying to coincide with the time of day for which there would be satellite coverage in the area of interest.

When collecting SGPS measurements one sets up a tripod which will hold the antenna over a marker, typically a benchmark set in the ground (Fig. 1a,b). The use of such a marker enables one to find the correct spot for future measurements months or years later. The antenna is centered using an optical plummet over a point imprinted on the marker to indicate its center (Fig. 1b). One must then measure the distance between the ground (where the benchmark is) and the antenna (where the GPS signal is actually received), and this distance will be used to convert the positions obtained from data processing to the positions of the benchmarks. As you might guess, there is considerable room for error when having to set up the antenna anew every time observations are made. Height measurement errors are particularly common. The repeated measurement of the same benchmark over time produces a time series like that in Fig. 2a.

As GPS receivers became less expensive, the possibility of permanently installing continuously recording GPS (CGPS) receivers became a real option. Some of the earliest continuous GPS stations for geophysical studies were installed in Japan [144]. The first continuous GPS sites in southern California were installed in 1990 at Pasadena, Pinyon Flat, and La Jolla [77]. Today several countries have extensive CGPS networks for geophysical monitoring purposes, most notably GEONET which consists of over 1000 stations across Japan. Regional CGPS networks have existed in the US for many years covering southern California (Southern California Integrated GPS Network,

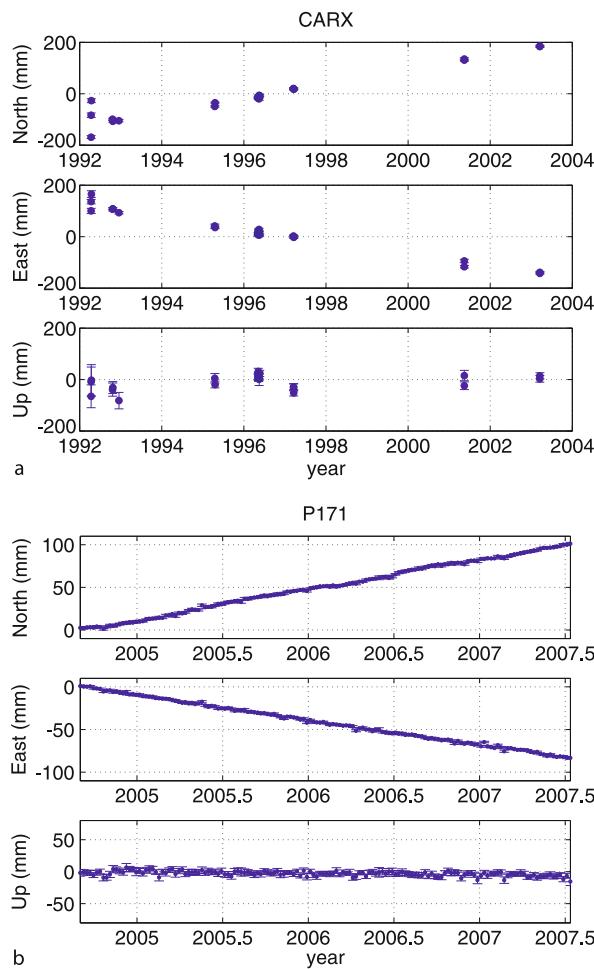


**GPS: Applications in Crustal Deformation Monitoring, Figure 1**

a Campaign GPS station located in western Nevada. In the foreground the tripod holds the antenna. The receiver will be stored in one of the boxes in the background while data are collected. Rocks are stacked on the feet of the tripod to prevent it from moving during data collection. (USGS photo) b Benchmark that is being observed in a. A benchmark often has a stamping telling which agency installed it and when. The cross in the center is the point over which the tripod is centered every time that benchmark is observed. (USGS photo) c Continuous GPS station located in Fremont, California. This station, P222, is part of the Plate Boundary Observatory. The antenna, protected from the elements by a domed cover, is supported by a monument with legs driven several meters into the ground. The antenna sends data to the receiver, located along with batteries in the box in the background, via a buried cable. Also visible in the background is the solar panel which provides supplemental power to the site. (UNAVCO photo, reprinted with permission)

or SCIGN [60]), the San Francisco Bay Area (Bay Area Regional Deformation, or BARD [72]), the Pacific Northwest (Pacific Northwest Geodetic Array, or PANGA [92]), and the Basin and Range (Basin and Range Geodetic Network, or BARGEN [10]; and Eastern Basin-Range and Yellowstone hotspot, or EBRY, <http://www.mines.utah.edu/~rbsmith/RESEARCH/UUGPS.html>). A much larger network consisting of ~850 sites is now underway. This new network, called the Plate Boundary Observatory (PBO), will provide CGPS coverage for seismically and volcanically active areas throughout the western continental US and Alaska [147].

In addition to providing daily positions, CGPS sites have the advantage of permanent monumentation, thus eliminating set-up error (Fig. 1c). Having daily positions (e.g., Fig. 2b) enables much better estimates of site velocities, but the more frequent data and added precision means it has become necessary to address additional error sources that were hidden in the noise of the less frequent measurements. These include seasonal signals and time-correlated noise likely due to monument instability, atmospheric effects, reference frame errors, or mismodeled orbits or antenna phase centers [15,33,74,86,166,170]. Despite the proliferation of CGPS stations in recent years, de-



**GPS: Applications in Crustal Deformation Monitoring, Figure 2**  
**Time series of changes in station positions.** Both these sites are located in central California, west of the San Andreas fault. The position changes are plotted relative to the center of the North American continent. Because of the relative motion between the North American and Pacific tectonic plates, these sites are moving to the northwest over time, relative to the stable interior of the continent which is not affected by strain accumulation due to the interaction of the two plates. **a** Campaign GPS site CARX near Parkfield, California. **b** Continuous GPS station P171 of the Plate Boundary Observatory network. This station is located just south of Monterey Bay

pending on factors such as site access and site condition, financial resources, and scientific goals, campaign-style GPS measurements are still frequently collected.

#### Relative Precision

Daily repeatability (or scatter) in the time series for CGPS sites is typically 0.5 to 1.5 mm in the horizontal and  $\sim 3.5$

to 6.5 mm in the vertical. If noise that is common to all stations in the network (due, for example, to errors in satellite orbits or atmospheric delay) has been eliminated, short-term repeatabilities of  $\sim 0.5$  mm in the horizontal and  $\sim 2$  mm in the vertical can be achieved. With 1.5 to 2.5 years of CGPS data, velocity estimates with  $\sim 1$  mm/yr uncertainties are possible [13]. In contrast, SGPS position repeatability is  $\sim 3$  to 5 mm in the horizontal and  $\sim 10$  mm in the vertical [40], and  $\sim 10$  years of SGPS data would be required to reach velocity uncertainty of  $\sim 1$  mm/yr [166].

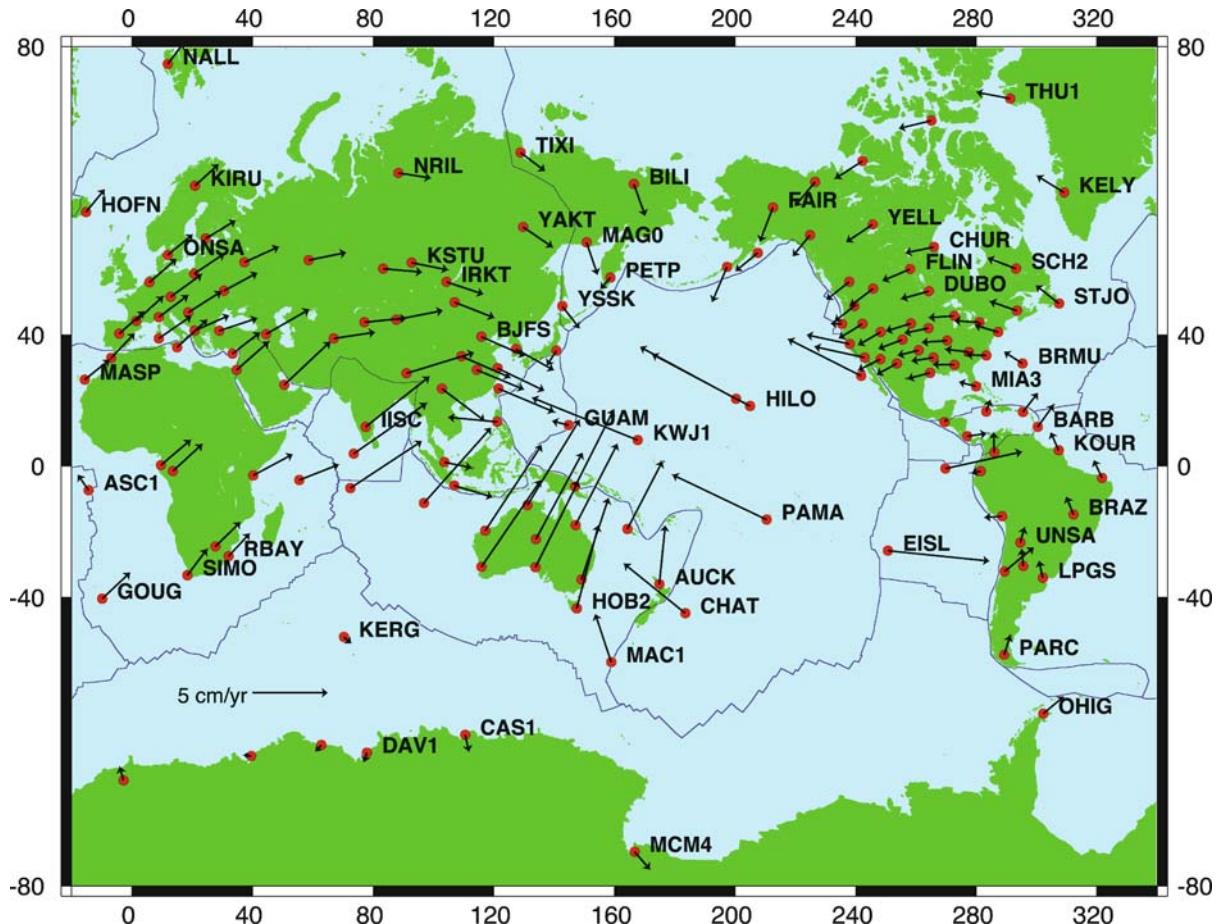
One way to improve the accuracy of GPS measurements is to extend the length of an observation session, for example from six hours to 24 hours, when estimating daily positions. The additional data mean that multipath noise (which varies throughout the day) averages out better, and it is easier to fix ambiguities and estimate tropospheric delay parameters. The accuracy of GPS measurements has also improved over time. A major reason for this has been the expansion of the global network of tracking stations that are used to calculate precise orbital information for the satellites and to define reference frames for GPS positions. Models for various noise sources have also been refined over time, and some noise sources have been mitigated, for instance by improving the stability of geodetic monuments. It should be noted that the vertical signal is generally much noisier than the horizontal because it is not possible to track satellites below the horizon, and therefore there is no position control from below.

#### Applications of GPS Data to the Study of Seismic and Volcanic Hazards

GPS data provide important constraints on the underlying processes that lead to observed deformation, especially when used in combination with other data types. One of the major strengths of GPS is its ability to track positions over time spans of seconds to decades; another is that it enables measurement of position changes over continental-scale baselines. This section will discuss several broad areas of study using GPS data and provide specific examples of each. Background on modeling methods will be included as necessary.

#### Plate Motions

Earthquakes occur in response to stresses in the Earth's crust, and these stresses are largely due to the motion of the Earth's tectonic plates. Knowing the velocity at which the plates move gives us insight into the amount of deformation that must be accommodated on plate boundary faults, and what type of earthquakes might be expected. The Earth's plates are generally thought to be rigid, at



GPS: Applications in Crustal Deformation Monitoring, Figure 3

Velocities for a globally distributed selection of GPS sites. Figure courtesy NASA/JPL-Caltech

least in their interiors (away from plate boundaries where plates interact with one another). One approach to measuring global plate motions combines rate information obtained from the age of magnetic reversals recorded in the basalt of the oceanic crust formed at mid-ocean ridges with information regarding the direction of plate motion gleaned from the orientation of transform faults and direction of slip in earthquakes. Because most such studies use a magnetic anomaly that is  $\sim 3$  million years old, the rates of plate motion inferred in this manner represent an average over the time since the Pliocene. The most commonly used plate motion model of this type is called NUVEL-1A [29,30]. This model gives relative velocities for pairs of plates. Under the assumption that the lithosphere has no net rotation with respect to the mantle below, Argus and Gordon [2] developed the NNR-NUVEL-1A model of absolute plate motions, where NNR refers to “no net rotation.”

In contrast to approaches based on geologic data, geodetic techniques enable the estimation of present-day plate motions using essentially instantaneous measurements. Two such geodetic methods, Satellite Laser Ranging (SLR) and Very Long Baseline Interferometry (VLBI), were used in the 1980s to measure the positions of points on the Earth’s surface. Given a good spatial distribution of sites world-wide, it is possible to use such data to estimate plate velocities. However, expense and practical considerations limited the number of sites that could be observed with SLR and VLBI. The advent of GPS provided a cost-effective alternative enabling precise three-component positioning with dense spatial coverage (e.g., Fig. 3).

Several studies have used GPS data, either alone or in combination with other observations, to estimate global plate motions (e.g., [4,79,124,140]), and numerous other studies have used such data in analyses focused on subgroups of plates. The plates are modeled as rigid, rotating,

spherical caps. Their motions are expressed as Euler vectors, defined by a location of the Euler pole (point E in Fig. 4a) and the angular velocity of the plate around that pole. These parameters are related to the velocities of the GPS sites by

$$\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{r} \quad (1)$$

where  $\mathbf{v}$  is the GPS velocity vector,  $\boldsymbol{\Omega}$  is the Euler vector for the plate in question, and  $\mathbf{r}$  is the position vector of the GPS site in Cartesian geocentric coordinates (e.g., [79]). The relationship given in (1) may be understood from Fig. 4. The Euler vector and position vector may be expressed as unit vectors ( $\boldsymbol{\Omega}_u$  and  $\mathbf{r}_u$ ) multiplied by their magnitudes ( $\omega$  and  $R$ )

$$\boldsymbol{\Omega} = \omega \boldsymbol{\Omega}_u \quad (2)$$

$$\mathbf{r} = R \mathbf{r}_u \quad (3)$$

where  $\omega$  is the velocity of rotation (generally in degrees or radians per million years), and  $R$  is the radius of the Earth (Fig. 4a). As shown in Fig. 4b, the distance,  $\mathbf{d}$ , that a GPS station travels during a time period  $t$  due to rotation about the Euler pole is given by

$$\mathbf{d} = \omega t R \sin \delta \quad (4)$$

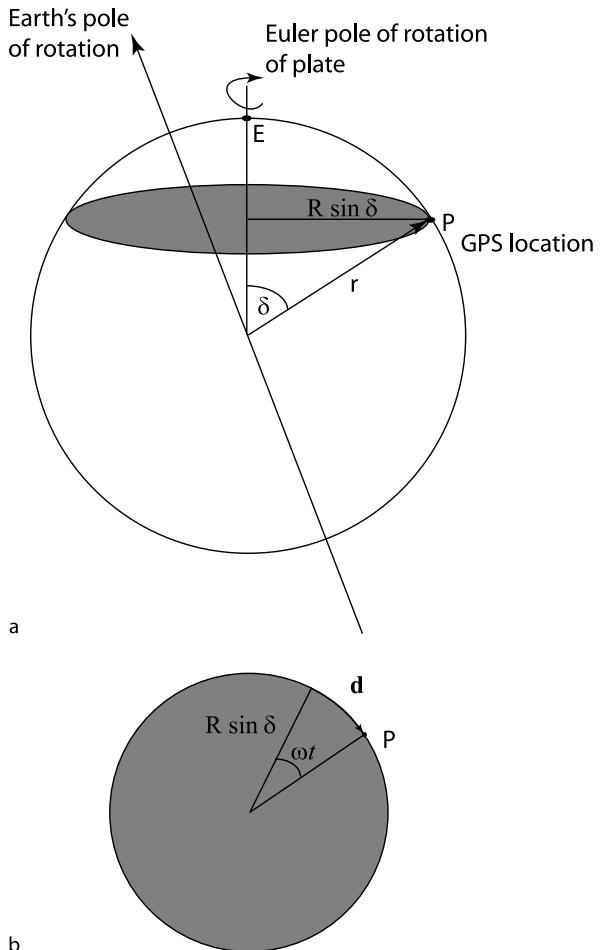
where  $\delta$  is the angular distance between  $\boldsymbol{\Omega}$  and  $\mathbf{r}$  (Fig. 4a). The quantity  $\sin \delta$  is simply the cross product of the unit vectors for the Euler vector and the position vector,

$$\sin \delta = \boldsymbol{\Omega}_u \times \mathbf{r}_u . \quad (5)$$

Therefore, plugging (2), (3), and (5) into (4), and dividing by time gives (1). The system of equations given in (1) can be solved to estimate the Euler vector that best fits the GPS data for each plate.

The early studies (e.g., [4,79]) were able to estimate velocities for only a handful of plates (e.g., six and eight, respectively). In recent years, however, the global distribution of CGPS sites has grown substantially such that Prawirodirjo and Bock [124] estimated the velocities of 17 “major and minor tectonic plates” and Sella et al. [140] considered 19 “plates and continental blocks.” The increasing precision of GPS measurements has made it possible to more rigorously test plate rigidity and the existence of purported plate boundaries, as well as address the potential systematic velocity error introduced by GIA in North America and Eurasia [124,140].

Velocities of GPS sites used in plate motion studies have repeatedly been found to be consistent with the assumption that plates behave rigidly, as evidenced by the



**GPS: Applications in Crustal Deformation Monitoring, Figure 4**  
**Relationship between the velocity of a GPS station and an Euler pole of rotation.** Point P is the location of a GPS station, and point E is the location of the Euler pole of rotation,  $\boldsymbol{\Omega}$ , that describes the motion of the plate on which point P sits. The Euler vector points from the center of the Earth to point E, and its magnitude is the rate of rotation,  $\omega$ . Point P is located at an angular distance of  $\delta$  from point E. The velocity recorded at point P due to rotation about the Euler pole is given by  $\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{r}$ , where  $\mathbf{r}$  is the position vector of the GPS station, and its origin is the center of the earth. Therefore, the length of  $\mathbf{r}$  is the radius of the earth,  $R$ . The gray circle that passes through the point P has radius  $R \sin \delta$ .  
**b** View of the gray circle in a looking along the direction of the Euler vector. The distance traveled by point P during a time period  $t$  due to the rotation given by  $\boldsymbol{\Omega}$  is given by  $\mathbf{d}$ . See text for details

fact that in general a single Euler pole fits the GPS velocities for a given plate well. Moreover, estimates of global plate velocities inferred from these data have agreed with plate motions inferred from geologic and other data (e.g., NUVEL-1A). However, the improvement in size and quality of GPS datasets has highlighted some discrepancies be-

tween NUVEL-1A and the models arrived at geodetically (e.g., [140]). These differences may be due in some cases to systematic errors in the rates estimated from magnetic anomaly data in tectonically complex areas where relative plate motion is not localized at the spreading center. In other cases the difference may reflect an actual change in the rate of relative plate motion over the past 3 million years. Refining both the geologic and geodetic plate motion models continues to be an area of active research (e.g., [3,28]).

### Earthquake and Volcano Source Modeling

**Source Potency and Geometry** Although the Earth's plates behave rigidly, as evidenced by the velocities of GPS sites in the stable interiors of plates, at their edges neighboring plates interact, colliding, diverging, or sliding past each other. These processes, as well as others such as the movement of magma underground, impart stress to the Earth's crust, distorting the shape of a volume of crustal material. Some of this deformation is permanent, leading for example to mountain building. However, the brittle upper portion of the Earth's crust deforms elastically, meaning that a large portion of crustal deformation is recoverable; once the stress is relieved, the crustal material returns to its pre-stress shape. The discussion presented here will focus primarily on modeling which assumes an elastic crust, however several studies have considered the effects of nonelastic material properties (e.g., [37,65,135]).

Mathematical expressions from continuum mechanics that describe the stress, strain, and displacement in an elastic solid due, for example, to a point source or to movement on a planar dislocation can be applied to the study of crustal deformation sources using surface displacement data like those provided by GPS. Okada [109,110] presented concise analytic expressions that are used in many such studies today. Mogi [99] discussed the special case of surface displacement due to three orthogonal, equal-amplitude point sources of inflation which could represent an inflating or deflating magma body at depth. Dislocation sources are described by their dimensions, orientation, location in the Earth's crust, amount of movement (e.g. slip) which takes place across them, and the direction that the material on one side of the dislocation moves relative to that on the other (sometimes termed the "sense of slip"). Volcanic sources such as dikes and sills can be modeled by dislocations with opening rather than slip. Magma chambers are often represented by inflation sources defined by their locations, amount of inflation, and, in the case of more complicated geometries such as ellipsoidal sources, their shapes. Dzurisin [40] gives a good discussion of ap-

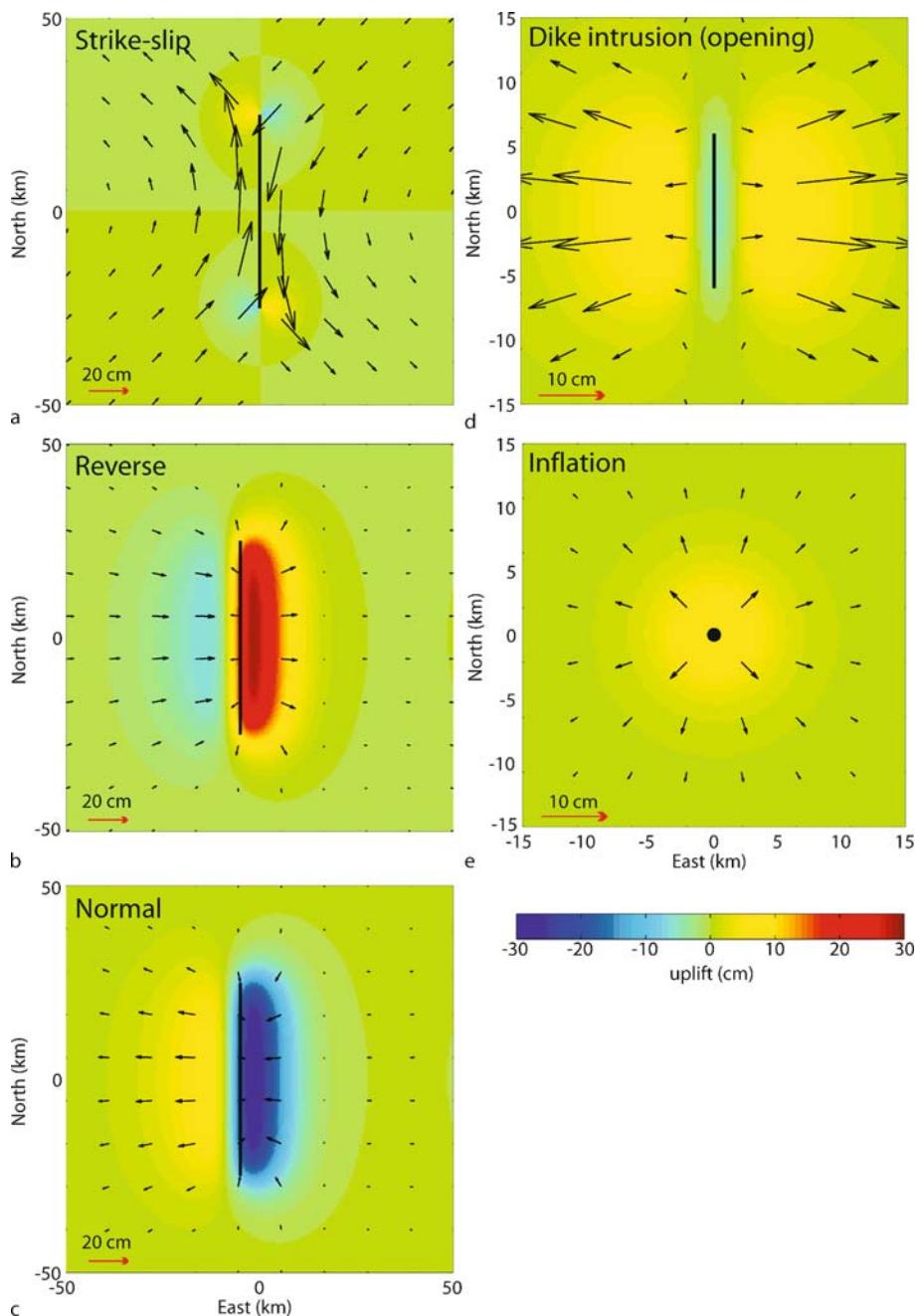
proaches for modeling a variety of volcanic deformation sources.

The term "source geometry" refers to all source characteristics except the amount of slip, opening, or inflation. These latter three parameters, which describe the strength of the source, are sometimes referred to as the "source potency." The surface displacement field produced in any deformation event reflects not only the source potency but also the source geometry and characteristics of the crustal material. Figure 5 presents the expected horizontal and vertical displacement due to different modes of shear slip and opening on a planar dislocation and to a point source of inflation. As can be seen, each deformation source produces characteristic surface displacement patterns. As with the estimation of plate motions described earlier, a system of equations can be written that relates a deformation source such as a dislocation in the crust to the displacements measured with GPS at the Earth's surface. In simplified terms this system of equations can be written as

$$\mathbf{d} = \mathbf{Gs} \quad (6)$$

where  $\mathbf{d}$  is a vector of station displacements measured by GPS,  $\mathbf{s}$  is a vector of source potency (e.g., fault slip), and  $\mathbf{G}$  is a matrix which embodies the mathematical expressions relating potency to displacements for an assumed fault geometry and elastic properties. This system of equations can be solved (or "inverted") to estimate the unknown potency that best fits the known displacements. The displacements at the Earth's surface are nonlinearly related to the source geometry, but are linearly related to the source potency. Therefore, when the source geometry is known, inverting for the potency is a linear inverse problem. In the simplest case the potency can be assumed to be uniform for the deformation source. In the case of a fault, this implies that the same amount of slip occurred everywhere on the fault, and the vector  $\mathbf{s}$  would have just one element. However, in the presence of multiple sources, the total displacement at the surface is simply the sum of the contributions from all the sources (e.g., Fig. 6). This means that spatially variable fault slip can be estimated by dividing the model fault into a grid of subfaults, each of which contribute to the observed displacements, and estimating the slip on each subfault. In this case the length of the vector  $\mathbf{s}$  would be the number of subfaults. The number of subfaults used is generally dictated by how much data are available.

Inversion of geodetic data for characteristics of deformation sources is underdetermined, meaning that a large number of source models can fit the data within errors. Having vertical and horizontal displacement measurements improves the ability to distinguish among different

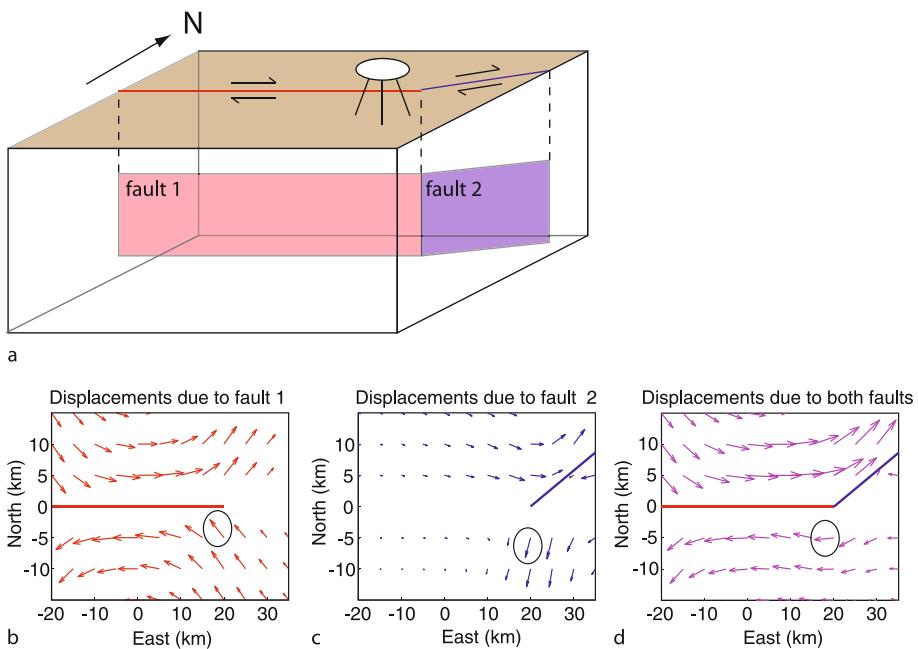


GPS: Applications in Crustal Deformation Monitoring, Figure 5

Predicted displacement due to dislocation and inflation sources assuming a homogeneous elastic half space. Vectors show horizontal displacement and colored background shows vertical motion. Note change of vector and spatial scales from a – c to d, e. Heavy black line in a – d is surface projection of upper edge of source dislocation. Black circle in e is surface projection of point source of inflation. See Table 1 for source parameters

**GPS: Applications in Crustal Deformation Monitoring, Table 1**  
Source parameters for deformation sources depicted in Fig. 5

Panel in Figure 5	Description	Length (km)	Width (km)	Depth to top (km)	Dip (degrees)	Slip or opening (m)	Sense of slip	Inflation ( $10^6 \text{ m}^3$ )
a	Dislocation	50	14	0	90	1	Right lateral, strike slip	
b	Dislocation	50	10	6.34	60	1	Reverse	
c	Dislocation	50	10	6.34	60	1	Normal	
d	Dislocation	12	5	3	90	1	Opening	
e	Inflation point source			5				10



**GPS: Applications in Crustal Deformation Monitoring, Figure 6**

Predicted displacements due to two sources. **a** Source geometry consisting of two vertical strike-slip faults and location of a GPS receiver. **b** Map view of displacement due to slip on the fault shown in red in **a**. **c** Map view of displacement due to slip on the fault shown in blue in **a**. **d** Map view of displacement due to slip on both faults. Note how the displacement at the circled locations in **b** and **c** due to the individual faults is very different from that which would be recorded by GPS in **d** due to slip on both faults

possible source geometries. For instance, several types of volcanic sources will produce similar patterns of vertical deformation, but with the inclusion of horizontal displacement measurements it is possible to differentiate among them. When using GPS data to estimate the spatial distribution of slip on a fault (by dividing the model fault into subfaults and estimating the slip on each), spatial smoothing is often used to provide added constraints in the inversion. The justification for this is that abrupt changes in the amplitude of slip would result in high stresses on the fault surface, which is physically unlikely. The relative weight given to fitting the data and to spatial smoothness is often determined empirically (e.g., [161]). Non-negativity may

also be applied, for instance to include prior knowledge about the sense of slip on a fault as a constraint on fault slip estimates.

Regardless of these means for regularizing inversions, because geodetic measurements are collected at the surface, their sensitivity to the details of a deformation source decreases with depth. GPS measurements will be most sensitive to source processes occurring in the upper few kilometers of the Earth's crust near the GPS receiver's location. Deeper sources will affect GPS sites over a broader region, but the recorded deformation signal will lack detail about the source. This can be best understood if one thinks of an earthquake that causes rupture of the Earth's surface.

A GPS receiver near the fault will record data which primarily reflect the shallow slip and surface rupture close to that receiver. A receiver 15 km away from the fault will not be sensitive to the shallow, near-fault deformation, and the recorded signal will be due to large-scale features of slip on deeper parts of the fault. Likewise, a receiver at the summit of a volcano can record the movement of magma that is collecting near the crater, whereas receivers lower on the flanks of the volcano will likely not record that signal but could be expected to track deformation due to movement of magma at greater depths.

The structure of major faults such as the San Andreas has been studied extensively by mapping, imaging the spatial distribution of background seismicity, and applying geophysical techniques that use, for example, seismic reflection and refraction, gravity, or magnetic data to highlight contrasts in rock properties. Therefore, if an earthquake hypocenter is found to be located on a major fault, the source geometry may be well-known *a priori*. In the absence of such information, or to refine the source geometry used in inversions, the spatial distribution of aftershocks and the location and extent of any surface rupture are also used.

Traditionally, spatially sparse geodetic measurements were assumed to be insensitive to the details of the model fault geometry used in inversions. With the recent growth of spatially dense GPS networks, however, more physically realistic fault geometries have been required. Maerten et al. [84] showed that using non-planar fault geometries that better represented independent information, for example from surface rupture mapping, led to significant improvement in fits to the GPS displacements for the 1999  $M_w$  7.1 Hector Mine earthquake. Methods for precisely relocating seismicity (e.g., [163]) have illuminated fine-scale fault structures that were previously obscured in less-precise catalog locations. This has enabled the development of more realistic model fault geometries.

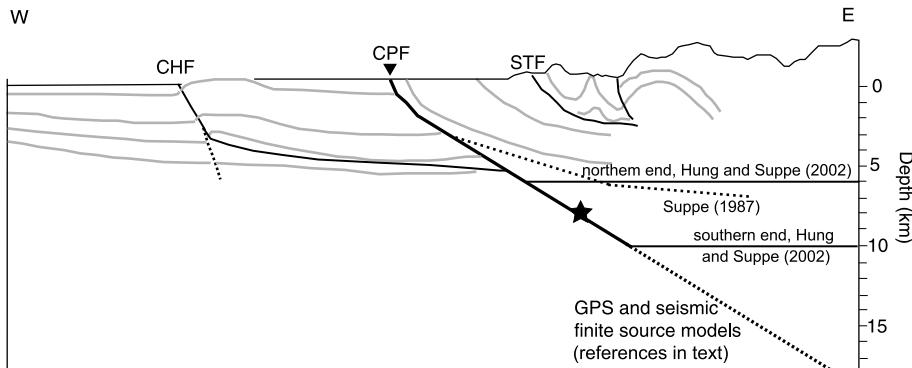
For example, Murray and Langbein [101] used displacements measured with GPS to estimate the slip distribution of the 2004  $M_6$  Parkfield, California earthquake. This event took place on the well-studied San Andreas fault in central California. Earlier work (e.g., [41,149]) had suggested that the fault was essentially vertical, with a strike of  $149^\circ$ . During the 2004 event, the coseismic displacement recorded for one GPS station, CARH, located within about 500 meters of the fault was in the opposite direction to what is predicted for a right lateral strike-slip fault (e.g., Fig. 5a). In order to fit the data for this station, an additional fault structure was needed. A sub-parallel fault called the Southwest Fracture Zone (SWFZ) which had exhibited movement in a previous

earthquake at this locale was a likely candidate. Using relocated aftershocks of the 2004 event [159] as a guide, Murray and Langbein [101] developed a non-planar fault geometry that consisted of the primary San Andreas fault and a subsidiary SWFZ. These two model fault surfaces passed through the relocated seismicity and intersected the mapped surface traces of the faults at the Earth's surface. Using this fault geometry, Murray and Langbein [101] inverted the GPS data to image the coseismic and postseismic slip associated with this earthquake.

Unlike in the case of Parkfield described above, often very little is known *a priori* about the geometry of a deformation source. This is especially true for sources that lie completely underground such as magmatic intrusions and earthquakes that do not cause any surface rupture (e.g., "blind" thrust events). However, as discussed earlier (Figs. 5 and 6), the surface displacement field produced in any deformation event reflects not only the source potency but also the source geometry. Therefore, by observing the spatial pattern of displacement using GPS, it is possible to infer what type of source lies underground by finding the source model that best predicts the observed data, for instance by using the expressions for deformation in an elastic material. Since surface displacements are nonlinearly related to source geometry, parameters describing the geometry cannot be estimated using linear inversion techniques but rather must be found through nonlinear optimization. Cervelli et al. [21] give a good overview of several approaches to this type of problem. When estimating the geometry and potency of an earthquake or opening source, often a two-step approach is employed: first the source geometry is estimated assuming uniform fault slip or dike opening, and then the inferred geometry is held fixed and the spatial distribution of slip or opening is estimated. As described in the following examples, the ability to infer the source geometry can help answer important questions about the underlying processes driving deformation such as whether a fault terminates in a décollement or how magma sources interact.

(a) *Deep Structure of the Chelungpu Fault, Taiwan* Johnson and Segall [65] used GPS-measured displacements caused by the 1999  $M_w$  7.6 Chi-Chi Taiwan earthquake to constrain the geometry of its rupture surface. They then went on to answer fundamental questions about the seismotectonics of Taiwan, which has formed due to the collision between the Philippine Sea plate and the Eurasian plate.

The extensive surface rupture that accompanied the Chi-Chi earthquake suggested that oblique reverse / left-lateral slip occurred on the previously known Chelungpu

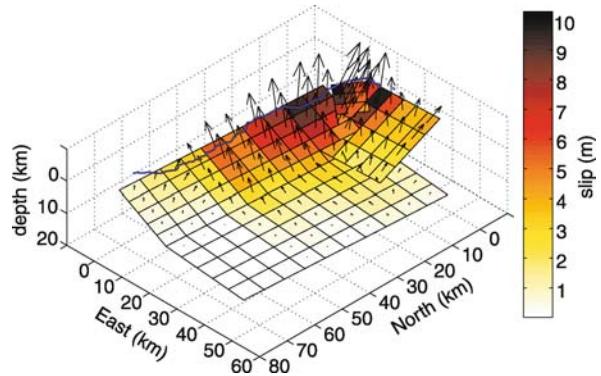


**GPS: Applications in Crustal Deformation Monitoring, Figure 7**

Cross section of the Chelungpu fault in Taiwan. The solid, near-horizontal black lines and the upper dotted line represent the ramp and décollement structure of the “thin-skinned” model, while the steeper dotted line below 10 km is the deeper extension of the Chelungpu fault envisioned by the “thick-skinned” model. CHF, Changhua fault; CPF, Chelungpu fault; STF, Shuangtung fault; star, hypocenter of 1999 Chi-Chi earthquake. Reprinted from [65] with permission from Elsevier

thrust fault which strikes north-south and dips  $\sim 30^\circ$  east. A nearly horizontal décollement structure has been interpreted to exist at depths of 6 to 10 km beneath much of Taiwan based on geologic mapping, seismicity locations, and seismic reflection profiles. In one proposed deformation model, termed the “thin-skinned” model, thrust faults like the Chelungpu fault intersect the décollement at depths of  $< 10$  km (Fig. 7). The alternative “thick-skinned” model interprets seismic and gravity data to suggest that, although a décollement may have controlled the long-term tectonic evolution of the island, more recent deformation has taken place on down-dip extensions of thrust faults with the same dip as their shallower portions (Fig. 7). In the case of the Chi-Chi earthquake, aftershocks occurred both at depths which might coincide with a décollement and considerably deeper, potentially on the down-dip extension of the Chelungpu thrust fault.

Johnson and Segall [65] optimized the source geometry of the Chi-Chi earthquake using a half-space with laterally and vertically varying shear modulus. They showed that the GPS data required slip on a thrust plane that transitions into an essentially horizontal décollement structure at  $\sim 8$  km depth. Moreover, they showed that displacements at the north end of the fault, where the surface rupture changed to a more east-west orientation, could only be fit by an additional thrust fault, which they term a “lateral ramp,” extending to the depth of the décollement. The estimated slip distribution of the earthquake using this geometry is shown in Fig. 8. These authors use the results for the Chi-Chi event as the basis for a conceptual model in which deformation follows the thin-skinned model, and lateral ramps form north of the Chelungpu fault due to this



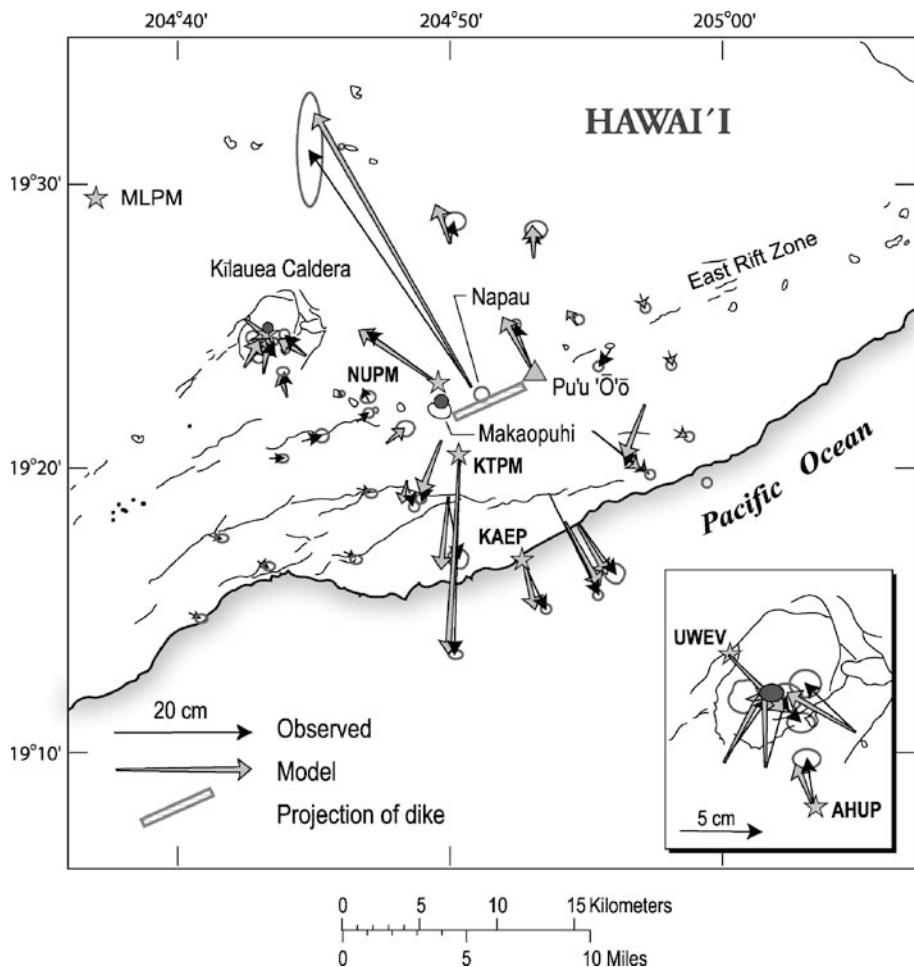
**GPS: Applications in Crustal Deformation Monitoring, Figure 8**

Slip distribution of the 1999  $M_w$  7.6 Chi-Chi Taiwan earthquake and optimized fault geometry inferred from GPS data. Colors indicate magnitude of slip and vectors show the direction that the hanging wall moved relative to the foot wall. The blue curve shows the trace of the earthquake rupture at the Earth’s surface. Reprinted from [65] with permission from Elsevier

fault’s orientation oblique to the direction of plate convergence.

(b) *Magma Plumbing System at Kilauea Volcano, Hawaii*  
Kilauea volcano, on the Big Island of Hawaii, consists of a summit crater as well as two rift zones, themselves consisting of several craters, extending from the summit down the flanks of the volcano (Fig. 9). Since 1983 Pu’u O’o, a collection of volcanic vents in the East Rift Zone, has been the center of eruptive activity, apparently fed by magma flowing through lava tubes from Kilauea summit.

In January 1997 a fissure eruption occurred on the East Rift Zone at Napau crater,  $\sim 3$  km closer to the summit



**GPS: Applications in Crustal Deformation Monitoring, Figure 9**

Displacements measured using GPS during the 1997 fissure eruption at Kilauea volcano. On the basis of the observations (black vectors with 95% confidence ellipses) Owen et al. [112] inferred the source of the deformation to be a combination of deflation at the summit of Kilauea (inward pointing arrows; see inset) and the intrusion of a dike along the east rift zone culminating in the eruption. The displacements predicted by this source geometry are shown by the gray arrows. CGPS sites are indicated by stars with four-character station codes. The deflation sources are shown as gray circles and the dike location by the hatched rectangle. Thin black lines are faults, fractures, and fissures. Adapted with permission from [112] (copyright 2000, American Geophysical Union)

than Pu'u O'o. In addition to uplift, the GPS instruments recorded horizontal displacements around the eruptive fissure at Napau Crater that were directed outward from the rift zone except at the ends of the fissure where they pointed inwards, parallel to the rift (Fig. 9). The magnitude of the displacements died off quickly with distance from the fissure. This displacement pattern is characteristic of a shallow dike intrusion within the rift zone. Prior to the eruption, sites at Kilauea summit showed subsidence and a radially inward pattern due to horizontal shortening across the summit, and this accelerated drastically during the eruption. This pattern suggests deflation, andulti-

mately emptying, of a magmatic source beneath the summit.

Using nonlinear optimization, Owen et al. [112] found that a source geometry consisting of a steeply dipping dike aligned with the rift and the fissures, combined with deflation both at the summit and at Makaopuhi crater, best fit the GPS observations (Fig. 9). The volume change at the two deflation sources was an order of magnitude less than the volume of the inferred dike intrusion. However, this discrepancy can be remedied if magma previously stored in a lava lake at Pu'u O'o which was seen to drain during this event, along with magma in a conduit thought to con-

nect the summit magma chamber to Pu'u O'o, also contributed to the intrusion. That the period of time leading up to the eruption had been characterized by steady deflation of the summit indicates that dike formation was not in response to magma overpressurization at the summit, but rather some other process. Owen et al. [112] suggest that ongoing southeastward movement of the south flank of the volcano (e.g., [111]) created tensile stresses that encouraged the dike intrusion and fissure eruption at the rift. Following the eruption, inflation resumed at the summit and the lava lake refilled, further evidence of the connection between the two magma reservoirs.

**Combined Use of Multiple Data Types** Wherever possible, multiple data types are used together to infer crustal deformation source characteristics. For example, GPS data are frequently used in combination with other geodetic measurements such as Interferometric Synthetic Aperture Radar (InSAR) [18] and leveling [38] data. GPS measurements are also often used in combination with seismic records to estimate fault slip.

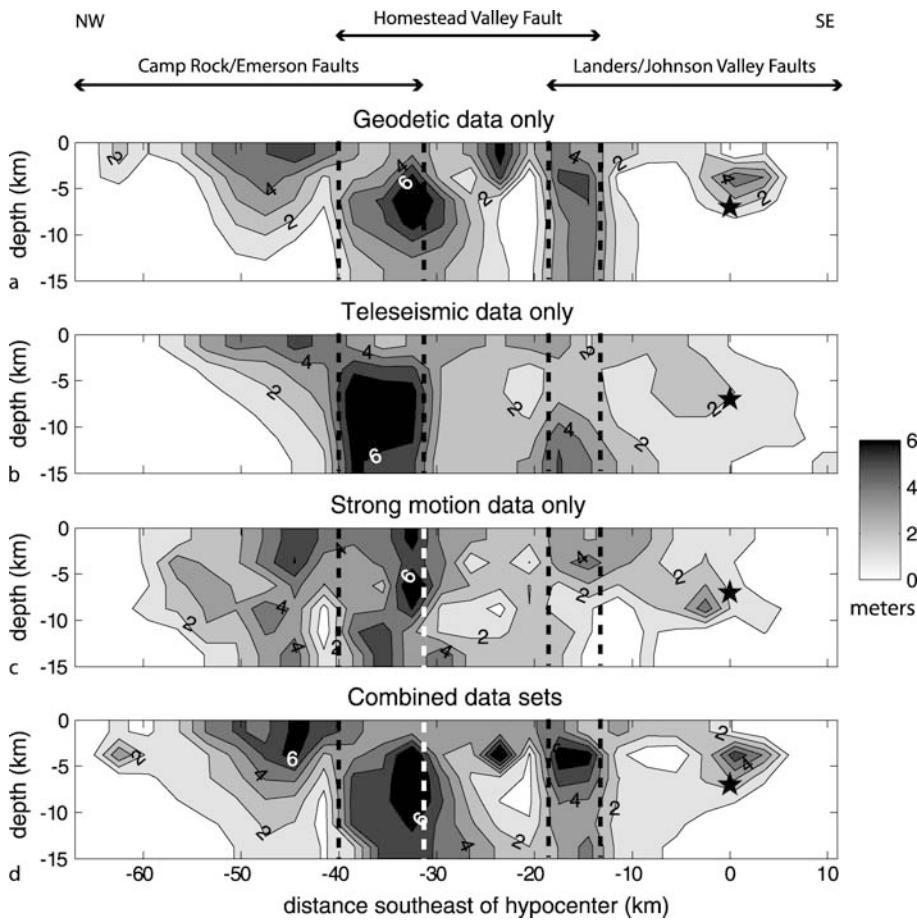
(a) *GPS and InSAR* GPS data are commonly used in combination with InSAR data because of the complementary nature of these two data types. GPS observations provide three-component displacements, good horizontal precision, and (in the case of CGPS) good temporal coverage. InSAR, on the other hand, has exceptional spatial coverage, is more sensitive to vertical deformation than GPS, and does not require the deployment of instruments on the ground (thus enabling data collection from otherwise hazardous areas such as volcanoes). InSAR, jointly with GPS where possible, has been widely used to study volcanic deformation (see for example [39,40,119]). Likewise, GPS and InSAR observations, in some cases in combination with seismic data, have been used to infer the slip distribution and rupture history of numerous earthquakes including the 1992 Landers earthquake [54], the 1995 Kobe earthquake [113], the 1999 Hector Mine earthquake (e.g., [70,132,148]), and the 1999 Izmit earthquake [16,27]. Studies have explored the relative constraints on source parameters provided by each data-type, as well as approaches for optimally weighting different data types in inversions, especially when one method, like InSAR, produces many more data points than another, such as GPS (e.g., [71,126,148]). Wright et al. [168] used a combination of InSAR and GPS data to infer the slip distribution of the 2002 Denali earthquake. Because of the remote location of this event, GPS sites were clustered along roads [57,58], thus resulting in poor spatial coverage along some portions of the fault. In this situation InSAR ob-

servations helped reduce uncertainty in the slip estimates compared to estimates obtained from GPS data alone.

(b) *GPS and Seismic Data* It has long been recognized that GPS data are a useful complement to seismic records for estimating fault slip [162]. While seismic data are sensitive to the rupture process of an earthquake (the amount and temporal progression of slip), when using these data to characterize the rupture process, trade-offs exist between the time-history of slip and its spatial distribution. GPS offsets due to an earthquake reflect only the final (or “static”) slip distribution, and thus the slip history inferred from the seismic data can be constrained to produce a static slip distribution that fits the geodetic displacements. The combined use of these two data types tends to have the added advantage of improved instrumental coverage over the study area.

Wald and Heaton [162] conducted a comparison of slip distributions for the 1992  $M_w$  7.2 Landers earthquake inferred from strong motion, teleseismic, and geodetic data (GPS displacements and displacements calculated from trilateration measurements) individually and in a combined inversion. For all inversions they used a consistent fault geometry parametrization comprised of three fault segments based on the aftershock locations and the extensive ground surface rupture that was mapped following the event. They corrected the geodetic data for the effects of the  $M_w$  6.2 Big Bear event and assumed that post-Landers measurements were made soon enough to avoid contamination by postseismic signals.

On inspection of the final slip distributions (Fig. 10) obtained from inversions of each of the three datasets independently, the authors identified several features that were common to all three and thus appeared to be robust regardless of dataset. For example, slip at the hypocenter was moderate and limited to a small depth range. Peak slip at depth occurred along the central portion of the fault, while slip became shallower at the ends (to the northwest and southeast). The greatest near-surface slip was on the Camp Rock / Emerson faults at the northwest end of the rupture. The slip distribution obtained through combined inversion of the three datasets was most similar to that from the geodetic data alone because the timing of slip is an additional degree of freedom in the inversion of teleseismic and strong motion waveform data that is not available in the inversion of static offsets. In addition to the final slip distribution, these authors imaged the temporal progression of slip on the fault surface. The rupture appears to slow as it nears the surface, as well as when it approaches the two step-over regions between fault segments. Furthermore, the authors infer that although the



GPS: Applications in Crustal Deformation Monitoring, Figure 10

Slip distributions estimated by Wald and Heaton [162] for the 1992  $M_w$  7.2 Landers earthquake using different data sets. Contour interval is 1 meter. Star marks hypocenter of earthquake. Dashed lines indicate along-strike boundaries of faults named at top. **a** Geodetic data only. **b** Teleseismic data only. **c** Strong motion data only. **d** Combined inversion of all three data sets. Adapted from [162]

rupture generally propagated unilaterally northwest, each time the rupture jumped northwest to a subsequent segment it also propagated backwards down the abandoned portion of this segment southeast of the fault intersection. Thus, the combined use of geodetic and seismic data provided a more complete and robust understanding of the rupture dynamics of this earthquake, which could then be used to investigate, for example, the spatial patterns in the strength of ground shaking due to this event.

Another example of the combined use of geodetic and seismic data comes from the Dec. 26, 2004 Sumatra–Andaman earthquake, a subduction zone megathrust event which ruptured a  $\sim$ 1200 km length of the plate boundary between the Indo-Australian and Eurasian plates [11,81]. This earthquake, which produced peak-to-peak surface wave motions greater than 1 cm worldwide [115] and measurable static offsets at GPS sites at

least 4500 km away [6] and probably farther [73], resulted in more than 283,000 deaths, largely due to its triggering of a major tsunami. This is the largest earthquake to have been recorded since the establishment of digital seismic networks and GPS, and both seismic and geodetic data have been critical in describing the rupture process of this earthquake.

The moment magnitude of an earthquake,  $M_w$ , is a measure of its size and may be derived from the seismic moment ( $M_o$ ), which is defined as

$$M_o = \mu sA \quad (7)$$

where  $\mu$  is the shear modulus of the faulted rock (in units of Pascals),  $s$  is the amount of fault slip during the earthquake (meters),  $A$  is the area of the surface that slipped in the earthquake (meters $^2$ ), and  $M_o$  has units of New-

tons × meters. “Tsunami earthquakes” [120] are a subclass of subduction zone earthquakes defined as producing tsunamis larger than would have been predicted from their moment magnitude. Tsunami earthquakes have been observed to have slow rupture velocity with relatively little seismic energy release at high frequencies. Although the moment magnitude should represent the net amount of static slip that occurred in the earthquake, it is typically estimated from seismic data at periods of 100 to 300 seconds [81]. If significant seismic energy is released at longer periods, this method will underestimate moment magnitude. The Harvard CMT (Centroid Moment Tensor) solution for the Sumatra–Andaman event, computed using surface wave data with periods of 300 to 500 seconds, had  $M_o = 4.0 \times 10^{22}$  Nm which corresponds to  $M_w 9.0$  [81]. However, researchers quickly began to see evidence that this earthquake was in fact larger.

Ammon et al. [1] analyzed seismic data from a broad frequency range (including data with periods up to 54 minutes) and concluded that these observations could be fit by a model in which the majority of the rupture occurred over a 10 minute time span and slip was concentrated south of  $8^\circ$  N (on the southernmost  $\sim 800$  km of the fault surface, Fig. 11). However, they note that the large GPS displacements reported in the Nicobar and Andaman islands would require 2- to 3-times more slip north of  $8^\circ$  N. Studies of the Earth’s seismic free oscillations [116,152] used these very low-frequency data (at periods up to one hour) to estimate the moment of the earthquake and found that moment increased with the period of the oscillations. These data suggest that the rupture required  $\sim 10$  minutes to travel from south to north, had more slip in the Nicobar and Andaman islands region than originally thought, and led to  $M_w$  estimates of 9.13 to 9.3. Park et al. [116] also point out that even slower slip (e.g. over a time span of  $\sim 1$  hour) could have occurred and would be difficult to detect in the free oscillation data.

Banerjee et al. [7] compiled GPS data from several studies and used these observations to infer the static slip distribution on the rupture surface. The total moment release associated with their slip estimate is  $7.62 \times 10^{22}$  Nm (corresponding to an  $M_w 9.22$ ), smaller than that estimated by Stein and Okal [152], but greater than the  $6.11 \times 10^{22}$  Nm ( $M_w 9.13$ ) inferred from GPS data by Kreemer et al. [73]. Banerjee et al. [7] conclude that from  $\sim 2$  to  $\sim 16$  meters of slip on portions of the fault north of  $8^\circ$  N is needed to fit the GPS data from the Andaman and Nicobar islands. However, they argue that this slip did not occur slowly over a time span of an hour or more because the continuous GPS site at Phuket, Thailand showed little movement after  $\sim 10$  minutes following

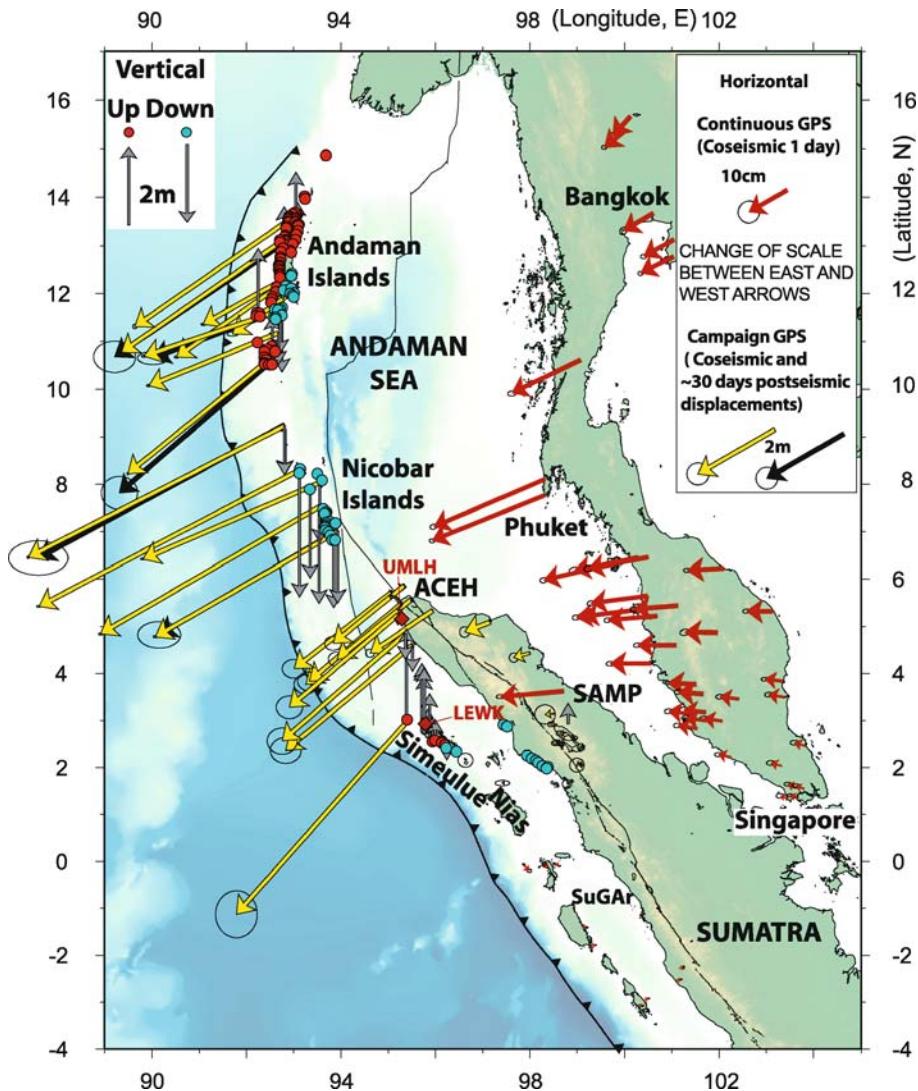
the earthquake [160], and other sites in Thailand which should be particularly sensitive to slip on the Andaman segment given their location also do not show movement more than 10 to 20 minutes following the earthquake [53]. Although some slip estimates based on geodetic data alone have higher moment than those from seismic data alone, suggesting the occurrence of aseismic slip, Chlieh et al. [24] and Rhie et al. [129] conducted joint inversions of GPS and seismic data and were able to fit both datasets satisfactorily with a single rupture model.

Thus, although seismic data provide information on the details of fault rupture that cannot be obtained from GPS, in the case of the Sumatra–Andaman event the GPS observations provided needed constraints on the extent and duration of fault rupture, both of which had important implications for tsunami generation.

(c) *GPS and Gravity Data* GPS data have become widely used to study volcanic deformation processes, including the long-term uplift observed at calderas such as Yellowstone (Wyoming), Campi Flegrei (Italy), and Long Valley (California). Geodetic observations can constrain the source geometry and volume change [8]. They cannot, however, discriminate if the deformation is due to an influx of hydrothermal fluids or the intrusion of magma. Battaglia et al. [8,9] address this problem through the combined use of geodetic and gravity data recorded during a period of uplift in Long Valley caldera. Modeling the observed uplift using a point source will produce biased results if the true source does not possess spherical symmetry. Furthermore, the uplift signal of a range of source geometries can be similar, but the horizontal deformation signal can help in distinguishing among different models (e.g., [32]). Battaglia et al. [8] use a combination of vertical and horizontal geodetic measurements (GPS, leveling, and line-length data) to find the best-fitting source geometry, in their case a vertical prolate ellipsoid. Then, with that source geometry uniquely determined, they perform a joint inversion of the uplift and gravity data to infer the volume and mass of the source, from which they obtain a density range of 1180 to  $2330 \text{ kg/m}^3$  [9]. Since this density range is too high for hydrothermal fluids to be the sole source of uplift at Long Valley, these authors conclude that a silicic magma body or combination of magma and hydrothermal fluids is required to produce the observed deformation.

## Deformation over Time Scales of Decades to Seconds

We have seen that GPS data are useful for constraining models of deformation that happens rapidly, such



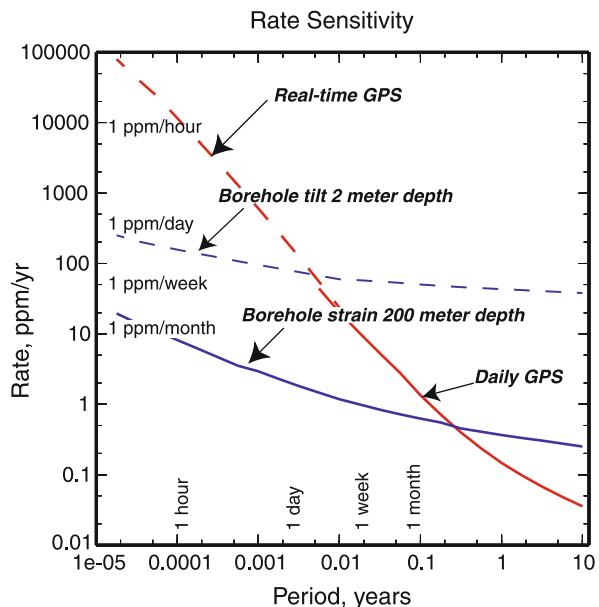
**GPS: Applications in Crustal Deformation Monitoring, Figure 11**

Summary of geodetic data recorded in the vicinity of the 2004 Sumatra-Andaman earthquake. The campaign GPS (yellow and black) vectors are compiled from [49,153] and contain approximately one month of postseismic deformation. The continuous GPS data come from Vigny et al. [160]. Note that the near-field vectors (those in the western part of the mapped region) and the far-field vectors use a different scale. Dots represent measurements of vertical deformation from satellite imagery [91]. The gray arrows indicate uplift and subsidence from GPS data, measurements of the vertical movement of coral heads, and mapping of shoreline changes [12,49,153]. Figure adapted from [24]

as slip in an earthquake. However, perhaps the greatest strength of GPS is its ability to record deformation that occurs over a wide range of time periods. GPS can measure fault slip that happens too slowly to generate seismic waves (termed aseismic slip), volcanic deformation that occurs over several days or several years, and long-term interseismic strain accumulation. GPS can provide much more temporally dense measurements than InSAR and, al-

though less sensitive, is stable to longer time periods than strainmeter data as shown in Fig. 12 [68].

**Interseismic Deformation** Interseismic deformation refers to the gradual straining of the Earth's crust that occurs during the time between moderate to large earthquakes. This strain can be caused by the build up of stress that will eventually be released during earthquakes, as well



**GPS: Applications in Crustal Deformation Monitoring, Figure 12**

Comparison of rate sensitivity for GPS, borehole strainmeters at 200 meters depth, and borehole tiltmeters (which measure the gradient in vertical deformation) at 2 meters depth. The x-axis indicates the time period that may be spanned by the different data types. For example, daily GPS measurements may span one day or longer. The y-axis indicates the strain rate that can be resolved as a function of the period. Strain is the change in length (area or volume) per a unit length (area or volume), and thus is unitless and can be expressed as parts per million (ppm). As can be seen from the plot, the borehole tiltmeters and strainmeters are more sensitive than GPS at shorter periods, but at periods longer than a few days and a few months, respectively, GPS measurements provide better resolution of strain rates. Figure courtesy of John Langbein

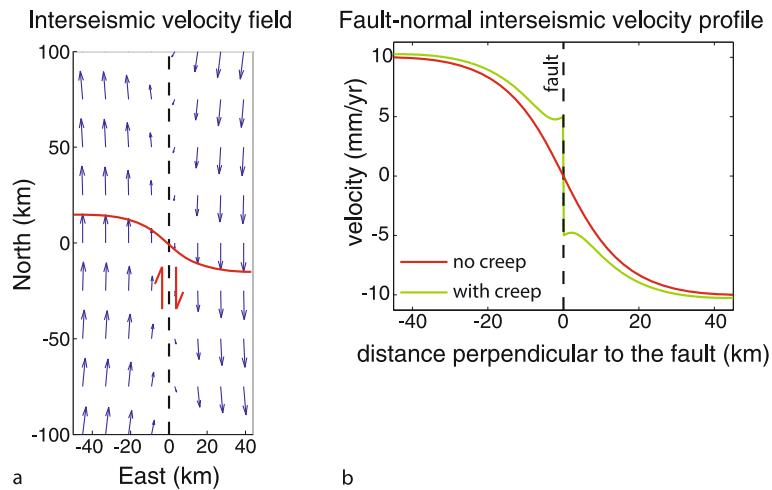
as reflect the broader-scale patterns of deformation in response to tectonic plate motion.

**(a) Block Versus Continuum Models for Deformation** As one moves from the global scale of tectonic plates to the continental scale, a major question is whether continents deform through the movement of many small rigid blocks (like plates on a smaller scale) or through more continuous deformation [157]. GPS data have been used in numerous studies to try to elucidate this hotly debated issue, and one region that has been a focus of study is the Tibetan plateau which accommodates strain due to the collision of India with Eurasia. GPS velocities for a profile of stations spanning central Tibet roughly parallel to the direction of maximum convergence between India and Eurasia show a linear gradient in the component of velocity parallel to the convergence direction. This observation has been cited as

evidence for continuously distributed deformation across the region (e.g. [171]) or distributed strain combined with the movement of a small number of crustal blocks [23]. In contrast, Thatcher [158] showed that the velocities predicted by a model in which the Tibetan plateau was divided into a set of rigid, rotating crustal blocks defined by faults and other geologic features could also fit the data since the difference between the observed GPS velocities and those predicted by the block model were relatively small and did not exhibit widespread systematic spatial patterns. Although the rigid block model does not predict a linear velocity gradient across central Tibet, Thatcher [158] notes that both the block and distributed models fit the velocity profile within errors. The spatial sampling provided by GPS remains sparse across large portions of the Tibetan plateau, and it is likely that additional data will be necessary to resolve the outstanding question as to whether continental deformation in this region is primarily block-like or continuous.

**(b) Estimating Interseismic Fault Slip Rates** Following the 1906 San Francisco earthquake, H. F. Reid [127] used his observations of deformation in that event to formulate a description of the earthquake cycle which he termed “elastic rebound.” Reid recognized that strain builds up in the Earth’s crust around faults during the interseismic period, and that strain is eventually released in earthquakes. Reid’s hypothesis predated the theory of plate tectonics, but we now understand that the source of the ongoing stress affecting the faults is the motion of the Earth’s plates. The strain build-up and release causes measurable deformation of the Earth’s surface. Areas where the crustal strain measured using GPS is large are likely to have earthquakes to relieve that strain. These events may take place on faults that are not visible on the earth’s surface, and in this case the geodetic data can provide an important clue to the existence of seismic hazard. GPS enables measurement of deformation during all phases of the earthquake cycle.

One of the most important pieces of information needed to characterize a region’s seismic hazard is an estimate of the slip rates on the major active faults that could affect that region. Most seismically active regions are at plate boundaries. The Earth’s rigid plates are always moving at rates that are essentially constant over time periods comparable to earthquake recurrence intervals (e.g., hundreds to thousands of years). At plate boundaries the relative motion between two neighboring plates is accommodated on faults. The portion of the rate of relative motion that is accommodated across a given fault is that fault’s slip rate. Slip rates are often estimated from geo-



**GPS: Applications in Crustal Deformation Monitoring, Figure 13**

a Expected velocity field due to interseismic slip below 15 kilometers on a vertical strike slip fault. The red curve highlights the difference in velocity near and far from the fault. b Expected velocity profiles for a locked fault and one that exhibits shallow creep. The locked fault is as in a. The creeping fault slips below 15 km depth and creeps in the uppermost 4 km but is locked from 4 to 15 km depth. Note that in the case of a creeping fault, there is an offset in the velocity profile close to the fault. The near-fault inflection in the green curve shows the combined effect of the strain due to slip below the locked zone and creep that reaches the Earth's surface

logic data, for example by dating samples collected from a location in which a measurable offset of a stream channel has occurred due to movement on a fault that crosses the stream [146]. However, slip rates can be inferred from geodetic data as well.

Below a certain depth in the Earth's crust (e.g., ~15 km for many strike-slip faults in the San Andreas system) the temperature and pressure are sufficiently high for the crustal rock to behave plastically in response to stress, rather than experiencing brittle failure. Earthquakes occur above this depth, which is termed the "brittle-ductile transition," the "locking depth," or the "transition depth," but not below. In the time between moderate to large earthquakes most faults are largely locked above the transition depth, meaning no movement occurs across them. (It is true that very small earthquakes occur frequently on most faults, but these events affect a relatively small portion of the fault's surface area and release a tiny fraction of the energy, or moment, released in moderate and large events.) The material below the transition depth deforms gradually and continuously in response to plate motion, and a fault that is locked above the transition depth may exist as a zone of distributed shear below that depth. In the vicinity of a locked fault, the constant movement of the material below the transition depth strains the elastic crust above. This is manifest by a characteristic pattern of interseismic velocities for points on the Earth's surface near the fault. For example, the interseismic velocity profile per-

pendicular to a strike slip fault like the San Andreas will have a characteristic sigmoidal shape as shown in Fig. 13, the details of which reflect the locking depth and slip rate of the fault.

In regions dominated by a small number of major faults of known geometry (e.g., the San Andreas fault system in northern California, or subduction zones of Cascadia and Japan), interseismic slip rates can be inferred in much the same way as was done for earthquake slip (Eq. (6)) using relatively simple dislocation models in which the fault is prescribed to be locked above the transition depth and freely slipping below that (simulating the motion of the tectonic plates). A number of studies (e.g. [47,125]) have taken this approach to model the GPS velocities of sites in northern California as the superposition of interseismic velocity signals due to slip on the multiple sub-parallel faults that make up the San Andreas system. When the slip rate on a fault is estimated from interseismic GPS velocities, the resulting value is generally called the interseismic slip rate to emphasize that it has been estimated from data collected over a time period entirely within the interval between two earthquakes on the fault in question. When slip rates are estimated from geologic data they are often called "long-term average" rates to emphasize that they represent the rate over many earthquake cycles. In the absence of post-seismic effects or other transient deformation (described in more detail in a later section) these two slip rate estimates for a given

fault should be the same. In subduction zones, interseismic velocities are often used to estimate the degree of plate “coupling,” which reflects the size of the locked zone that may rupture in a large earthquake [17,44,93,108,154,164].

As discussed earlier in the context of the Tibetan plateau, one interpretation of continental deformation patterns is that they arise from the rotation of fault-bounded blocks, and the GPS data can be used to estimate the Euler poles of rotation for each block. Data from sites near known faults may be discarded from the analysis because these observations will likely reflect the interseismic elastic strain accumulation due to the faults rather than the long-term rigid behavior of the blocks. Slip rates on the faults that bound the blocks can be calculated from the relative rates of block rotation (e.g. [158]). An alternative block modeling approach retains all the data, and the estimated block rotation rates must result in slip rates on block-bounding faults that are compatible with the patterns of strain accumulation recorded in the GPS velocity field. As with the dislocation models, a transition depth is assumed for each fault. The portion of each fault above the transition depth is treated as locked, and slip on the portion below the transition depth drives the observed strain accumulation.

Block modeling of this type lends itself to the many seismically active regions, such as southern California, that are characterized by numerous faults with complex geometries. Unlike models which represent individual faults by separate dislocations, block models are required to be self-consistent in that the rotation rates must be compatible for all blocks, slip rates have to be consistent at fault intersections, and the total slip rate across the region is made to match the relative plate rate. A drawback of this approach, however, is that it is difficult to accommodate dipping faults and to constrain fault-perpendicular motion in a realistic way.

Both block models and dislocation models suffer from trade-offs between slip rate estimates on different faults and sensitivity to poor data coverage. Furthermore, it is difficult to resolve the contribution to the observed GPS velocity from strain accumulation on closely spaced faults (e.g. within two locking depths of each other), and thus slip rate estimates on neighboring faults tend to trade-off with each other and with the assumed locking depth. It will always be difficult to resolve slip rates on faults that are close together given that displacement measurements are confined to the Earth’s surface.

Several studies have applied the technique of block modeling with inclusion of elastic strain accumulation to the western United States [26,89,90]. For many faults the slip rates estimated by these studies agree to within er-

rors with those estimated from geologic studies. However, there are some discrepancies. For example, the ~5 mm/yr slip rate Meade and Hager [90] estimated for the San Bernardino segment of the San Andreas fault is considerably lower than the geologic estimate of ~25 mm/yr. Discrepancies between geodetic and geologic slip rates have been found in several locations world wide, but it is not the case that one data type tends to produce consistently higher rate estimates than the other. The differences likely result from a combination of factors, including assumptions made in interpreting the data, the localized nature of geologic estimates, and the different time periods spanned by the two data types. Consistency between geodetic and geologic slip rates is an area of ongoing study.

Time varying deformation is often observed following large earthquakes. One source of this signal in many cases is the viscoelastic response of the material below the transition depth. When a large earthquake occurs it imparts stress to this material, which then deforms slowly, restressing the elastic crust above. The rate of the resulting strain would be expected to be high directly after the earthquake and die off with time. Likewise, the strain accumulation rate on the affected fault would vary throughout the earthquake cycle. These processes have been incorporated into another category of interseismic deformation models called “viscoelastic cycle” models [66,67,133,137,156], which can be used to estimate fault slip rates and earthquake recurrence times. These models account for the strain rate maxima on faults during the interseismic period as well as temporal variations in the months to years following a large earthquake.

(c) *Fault Creep* An interesting interseismic phenomenon, termed “fault creep,” is observed on several faults of the San Andreas system in central California and the San Francisco Bay Area, as well as faults in Taiwan, the Philippines, and Turkey. Faults exhibiting creep slip steadily or episodically at low average rates (e.g. 10 mm/yr). Some creep events are confined to the uppermost ~500 meters of faults, but in other cases creep occurs deeper at what are generally considered to be seismogenic depths (e.g. to a depth of ~15 km on faults of the San Andreas system). However, this fault slip is too slow to generate seismic waves. The ongoing slip of creeping faults is a nuisance, offsetting cultural features like curbs and buildings. However, ongoing creep constantly relieves stress which would otherwise be released in an earthquake and thus reduces the seismic hazard due to that fault. For example, a 2003 study of earthquake probabilities for the San Francisco Bay area found that accounting for creep on the Calaveras fault reduced the predicted rate of  $M \geq 6.7$

earthquakes on this fault by a factor of three [167]. Therefore, knowing the extent of creep is valuable in hazard assessments.

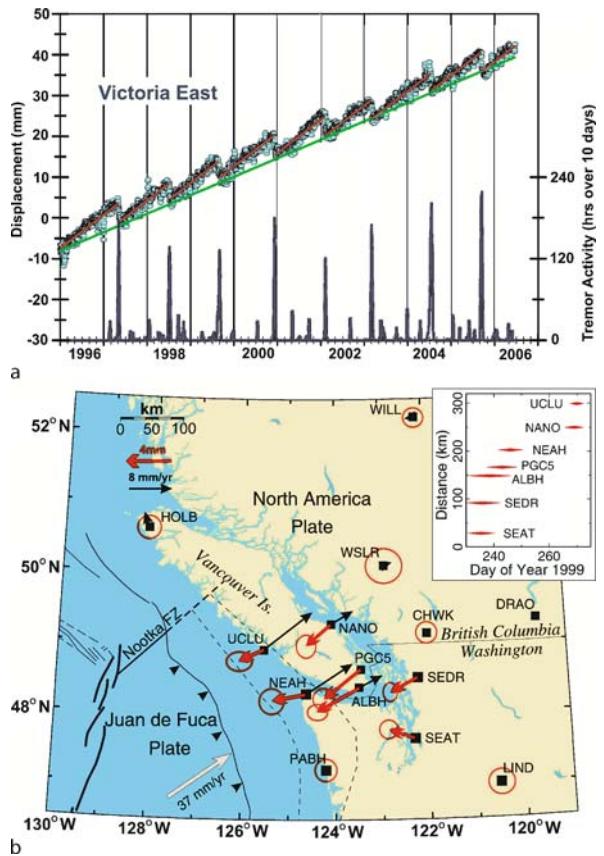
Fault creep will have a different interseismic signature than that for a fault that is locked in the seismogenic zone (Fig. 13b). Instead of the smooth transition seen across a locked fault, there will be a step because there is fault offset near the surface. Therefore, GPS data can be used to infer the depth-extent of fault creep. Several studies have used GPS data, in some cases in combination with InSAR, creepmeter, and microseismicity observations, to estimate the spatial distribution of fault creep (e.g. [64,85, 103,136]).

**Temporally Varying Deformation** Continuous GPS measurements are particularly useful for observing transient, or time-varying, deformation. Sources of transient deformation include slow slip events, postseismic response, and volcanic processes.

(a) *Slow Slip Events* Slow slip events, sometimes also referred to as “slow earthquakes” or “silent earthquakes” (the latter emphasizing the lack of a seismic signature to the event), are a phenomenon in which fault slip occurs at too slow a rate to generate seismic waves. The duration of slow slip events that have been observed geodetically world-wide ranges from days (e.g. [22,106]) to years (e.g. [98,102]). Precursory transient slip with duration of minutes has been observed prior to earthquakes on mid-oceanic ridge transform faults using seismic data (e.g. [62]).

Shallow creep events and a multi-year transient increase in slip rate have been observed along the San Andreas fault where continuous or frequent monitoring using creepmeters, strainmeters, and two-color electronic distance measuring instruments provided temporally dense measurements [51,52,75,82]. In 1996, a couple of years after the establishment of the Japanese CGPS network, evidence for slow slip in the Boso Peninsula region near Tokyo became apparent in the data [131]. As the spatial coverage of CGPS networks worldwide improved it became clear that slow slip events were much more frequent than previously thought, occurring over a variety of spatial and temporal timescales and tectonic settings.

The majority of large slow slip events observed to date have occurred in subduction zones, including those of Japan, the Pacific Northwest of the United States, Mexico, New Zealand, and Alaska. In 2001 a surprising pattern was recognized in time series for several CGPS sites in the Cascadia region in the Pacific Northwest of the United States (Fig. 14a). The subduction interface is thought to



**GPS: Applications in Crustal Deformation Monitoring, Figure 14**

a Time series for station ALBH located in Victoria British Columbia, Canada relative to stable North America. The blue dots are station positions. Although the overall interseismic movement of this site is eastward due to the ongoing strain caused by subduction (green line), every ~14 months this site moves westward (steps in time series). The red line represents the average velocity in the time between slow slip events, which is a higher rate than the long-term interseismic movement (green line). The blue curve represents the time series of nonvolcanic seismic tremor. Periods of increased tremor activity coincide with the times at which the GPS site shows anomalous westward movement [130]. b Portion of the Cascadia subduction zone. Black vectors are interseismic velocities of continuous GPS sites relative to stable North America, and red vectors are anomalous displacements during the 1999 slow slip event. The inset shows relative timing of transient displacements among different sites. From [35]. Copyright, Her Majesty the Queen in right of Canada (2001)

be locked near the Earth’s surface and freely slipping at greater depths. The ongoing deep slip results in interseismic motion of GPS sites toward the over-riding plate (Fig. 14b). In the case of the Cascadia sites, this means that the long-term average interseismic movement is eastward. However, it was observed that occasionally the sites briefly

moved in the opposite direction, causing a step-like pattern in the position time series. An intriguing feature of the observed reversals in station velocities is that the pattern was found to repeat approximately every 14 months [94].

That several CGPS sites in the region showed a coherent reversal in the direction of motion at the same time (Fig. 14b) suggested that the source of this signal could be slip on the interface between the down-going slab and over-riding plate of the subduction zone. Modeling of the GPS data [35] indicated that the source region was a portion of the subduction interface between 30 km and 40 km depth that is a transitional zone linking the shallow fully locked and deeper fully slipping parts of the interface.

In 2002 Obara [105] reported a very low frequency seismic signal called tremor, typically observed at active volcanoes, emanating from ~30 km depth in the subduction zone of southwest Japan. A similar signal was soon discovered in Cascadia, and researchers quickly realized that the tremor occurred simultaneously with slow slip events in those locales (Fig. 14a). Furthermore, they identified the source regions of the tremor and found that it coincides spatially with the inferred source region of the slow slip [106,130]. In volcanic settings tremor is thought to be caused by the movement of fluids through conduits underground. Recent studies have found that slow slip events in subduction zones are accompanied by very low frequency earthquakes and that the tremor in these locales is actually made up of many low amplitude low frequency events [142,143]. One interpretation is that shear slip on the subduction interface, rather than fluid flow, causes nonvolcanic tremor, implying that tremor and slow slip both arise from the same underlying process of shear slip. Although the tremor may not be directly caused by fluid flow, the low frequency earthquakes and slow slip appear to coincide spatially with areas of high fluid pressure in the pore spaces of subduction zone rocks. The high pore pressure, perhaps resulting from metamorphic reactions that release fluid, may encourage shear slip [143].

Like Cascadia, other subduction zones, for example the Guerrero region of Mexico [83] and the Shikoku [106] and Tokai [55] regions of Japan, have also experienced, to varying degrees, quasi-periodic slow slip events. Slow slip, in some cases multiple events, has been observed in other subduction zones such as Alaska [107], New Zealand [34], and the Tokai [98] and Bungo channel [114] regions of Japan, but it remains to be seen if the transient slip is periodic. Even in Cascadia, which shows clear periodicity, there is variation in the periodicity along the strike of the subduction zone. For example, slow slip events have been observed in both northern and southern Vancouver Island with a ~14 month periodicity, but the events in these two

locations are 6 months out of phase with each other. Slow slip events in northern California, also part of the Cascadia subduction zone, have been found to have an ~11 month recurrence interval [155]. In southwest Japan, the Shikoku region experiences short duration (on the order of a week), small amplitude slow slip events (in fact only detectable in the tiltmeter data) at six-month intervals coincident with tremor [106]. In contrast, the subduction zone beneath the Bungo channel region, which abuts the Shikoku region directly to the southwest, experiences infrequent large slow slip events lasting ~1.5 years [114]. Similarly, in the Tokai region, short-term and long-term slow slip events seem to occur on nearly overlapping parts of the subduction zone [55,98].

Similar to fault creep, the cumulative effect of several slow slip events may be to relieve stress on the transition zone along the whole length of the subduction interface without a large earthquake. However, slow earthquakes in the subduction zones of Cascadia, Japan, and elsewhere may impart stress to the locked subduction interface up dip and thus increase the likelihood of a large earthquake [35,123]. Understanding both the mechanism of these events and what it means for seismic hazard continues to be a focus of intense study.

Transient slip has also been observed in non-subduction zone settings. GPS data have shown that the south flank of Kilauea volcano in Hawaii moves seaward at a rate of several cm/yr [111], perhaps a manifestation of the instability of the volcanic edifice. However, in 2001 it was observed that this motion sped up for a few days. Cervelli et al. [22] modeled the GPS observables during the period of increased station velocity and concluded that the source was a ~M 6 slow earthquake lasting 36 hours and generating an average of almost 9 cm of slip on a nearly horizontal thrust fault about 4.5 km underground. This observed transient signal occurred nine days after a storm caused 1 meter of rainfall on this part of the Big Island. Cervelli et al. [22] calculated that given reasonable values for the porosity and permeability of the rocks in the study area, the slow slip could have been triggered by an increase in pore fluid pressure as the rain water penetrated into faults on the volcanic edifice. Subsequently several more very similar transient events have been observed in this locale by GPS [139], not accompanied by anomalously high rainfall. Moreover, it has been recognized that these events are accompanied by increased seismicity in an adjacent area following the onset of the deformation signal. This lends credence to the interpretation that the displacement signal is due to fault slip and suggests that the transient slip triggers the seismicity. In order for slow slip to trigger the observed seismicity, the slip must occur deeper

than originally thought. A model in which slip occurs at the depth of the interface between the volcano and the underlying ocean floor, ~8 km, fits the GPS displacements as well as one in which slip occurs at more shallow depths (~4.5 km) [139].

*(b) Postseismic Deformation* Often following a moderate or large earthquake, continued aseismic deformation is observed. This may be due to several sources, including continued slip on the fault surface, diffusion of pore fluid, and the viscoelastic response of the lower crust (below the transition depth) and upper mantle. All of these processes are triggered by the stress changes imparted to the surrounding crust and mantle by the earthquake. More than one source can be active simultaneously, and the effects change and decay with time. The spatial and temporal evolution of postseismic signals can provide important insights into the frictional, hydrological, and rheological characteristics of the crust and upper mantle, and enable a better understanding of the way stress is redistributed in the crust, which is an important consideration in assessing seismic hazard.

Afterslip is continued slip on the fault plane after the rapid slip which generates seismic waves has ceased. The effects can start immediately after the event (e.g. [76]) and can last for months or years, typically showing a logarithmic decay with time (e.g. [134]). Afterslip is thought to arise from the stress changes imparted by the earthquake to parts of the fault which have frictional properties that allow slow slip. Typically the total afterslip following an event is a fraction of the amount of slip that occurred coseismically. However, some events such as the Sanriku-haruka-oki [169] and Tokachi-Oki [97] earthquakes in Japan have had afterslip with moment release approaching or, in the case of the 2004  $M_w$  6 Parkfield earthquake [76], exceeding the coseismic moment. Although the factors controlling the amount of afterslip are not fully understood, the fact that the San Andreas fault near Parkfield, as well as many subduction zone faults, are known to exhibit fault creep may be a factor.

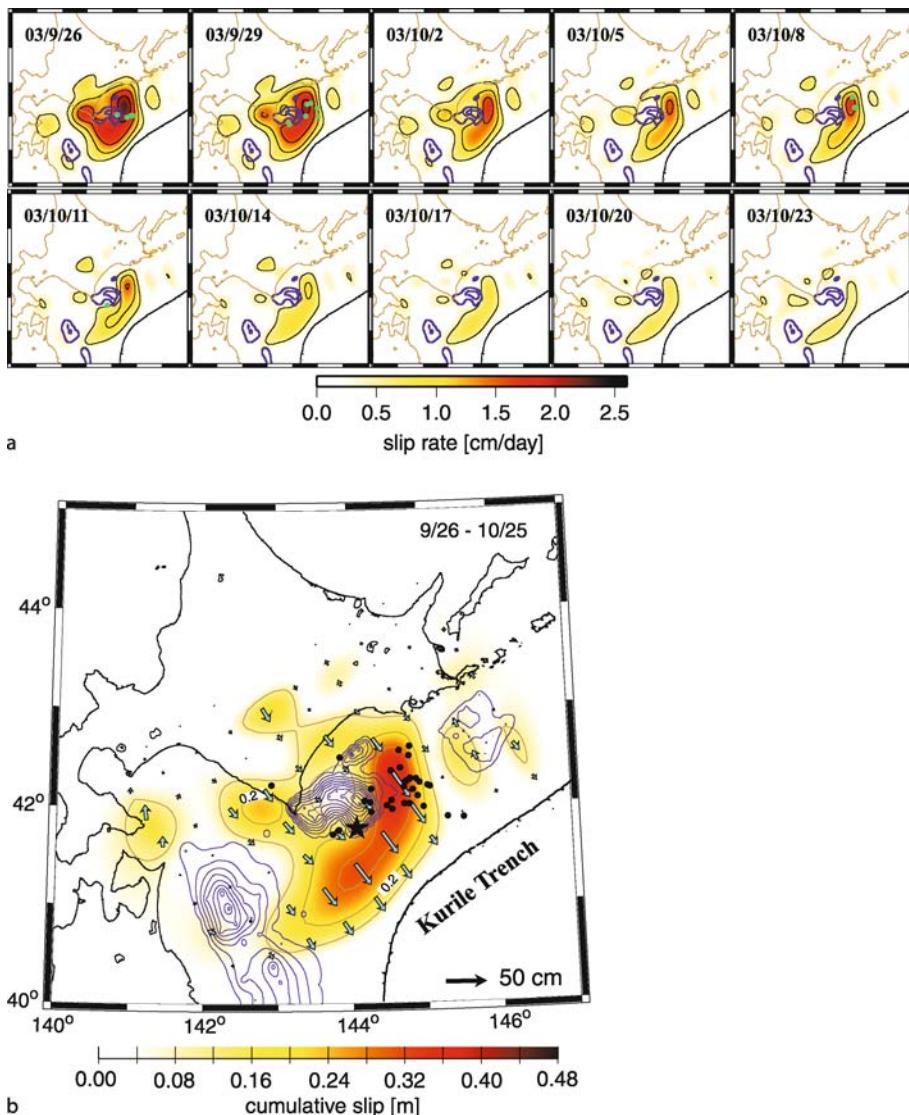
The spatial and temporal evolution of afterslip following the 2003  $M_w$  8.0 Tokachi-Oki earthquake in Japan was estimated from GPS data using a Kalman-filtering technique [97] (Fig. 15a). The rate of afterslip started quite high and died off gradually. These authors noted that the highest afterslip rates tended to localize around the area which slipped in the earthquake. This is not unexpected as the coseismic slip area would have just experienced a reduction in shear stress which would discourage further slip here. Interestingly, the afterslip seems to avoid areas inferred to have slipped in other earthquakes that occurred

in the decades prior to the Tokachi-Oki event as evidenced in the plot of cumulative afterslip in Fig. 15b. This interpretation provides support for the idea that rapid coseismic and gradual aseismic slip may occur on different parts of the fault because of variations in frictional properties. Similar results were found from an analysis of afterslip following the 2005 Nias–Simeulue thrust event [59].

As described earlier in regards to viscoelastic cycle models of interseismic deformation, the stress changes imparted by a moderate or large earthquake to the material below the elastic crust causes time-varying deformation. The geodetically recorded postseismic deformation following several large events has been interpreted to reflect viscoelastic processes (e.g. [156]). The rate at which viscoelastic postseismic deformation decays depends in part on the viscosity of the lower crust and upper mantle. The temporal decay of displacements measured with GPS have been used to infer viscosity values and thus the relative strength of these two layers (e.g. [5,31,45,46,121,122]). Estimates for the viscosity of the lower crust range from  $10^{19}$  to  $10^{21}$  Pa s, and for the upper mantle range from  $10^{17}$  to  $10^{19}$  Pa s.

Stress changes in the crust due to fault slip in an earthquake compress the pore space of rocks in some areas and cause dilation of the pore space elsewhere, depending on the orientation of the fault and sense and distribution of slip. The resulting pore pressure gradients cause fluid to flow from areas of high pressure to those of low pressure. This fluid flow causes further time-dependent strain. Such effects have been observed following several earthquakes including a pair of moderate earthquakes that occurred in 2000 in Iceland [5,69] and the  $M_w$  7.3 Landers event in California [43,118]. If poroelastic effects are confined to the upper few kilometers of the fault zone, the spatial extent of the resulting surface deformation will be localized near the fault. Also, vertical surface displacement is a large component of the poroelastic signal. Because GPS measurements of vertical displacement are noisier than the horizontal data, and because the distribution of GPS stations may be limited near the causative fault, much of the insight into poroelastic deformation following events like the Landers earthquake and the earthquakes in Iceland has come from InSAR data.

It is unlikely that postseismic deformation associated with a given earthquake can be explained by a single process. Given the often limited spatial and temporal data coverage and, in the case of GPS data, the sometimes poor vertical displacement control, it can be difficult to differentiate among different potential sources, e.g., afterslip, viscoelastic, and poroelastic deformation. However, the surface displacements due to each process can exhibit diag-



**GPS: Applications in Crustal Deformation Monitoring, Figure 15**

a Temporal evolution of the rate of afterslip (indicated by colored shading) on the subduction interface following the 2003  $M_w$  8 Tokachi-Oki earthquake. Dates given in upper left of each frame. Blue contours are the areas that slipped in previous earthquakes in this region. The blue contours in the center of the mapped area are those of the Tokachi-Oki earthquake. The green dots are aftershocks. b Cumulative afterslip in the first 30 days following the earthquake estimated from GPS data. Slip magnitude is given by the colored shading; the estimated amount and direction that the upper plate (the area northwest of the Kurile trench) moved relative to the lower plate is shown by the arrows. Afterslip tends to surround the areas inferred to have slipped in the Tokachi-Oki earthquake (epicenter given by the black star) and other events, shown by the blue contours. The black dots are aftershocks. Adapted with permission from [97] (copyright 2006, American Geophysical Union)

nostic patterns in time and space, which, if observed, make it possible to discern distinct causative processes. For example, due to its deeper source viscoelastic deformation should affect a broader geographic region than poroelastic deformation or shallow afterslip. Because viscoelastic deformation involves the response of material with vis-

cosities on the order of  $10^{18}$  to  $10^{19}$  Pa s, while poroelastic processes involve the flow of aqueous fluid through the ground, the signal due to the former is expected to last considerably longer (e.g., several years) than that of the latter (e.g. several months). Several studies have analyzed postseismic deformation, in some cases observed us-

ing multiple data types, and attributed the observed deformation to a combination of two or more effects [5,43,46].

*(c) Volcano Deformation* Volcanic deformation is often characterized by transient signals. Magma or hydrothermal fluids migrate beneath the volcanic edifice, causing inflation or deflation and sometimes culminating in an intrusion or eruption. The steep flanks of volcanoes are often unstable, leading to landslides and in some cases collapse of large sections of the edifice. GPS is very well-suited to monitoring these types of deformation signals, and many of the world's volcanoes have CGPS receivers installed for this purpose. Some of the earliest GPS observations of volcanic deformation come from a submarine volcano near the Izu peninsula in Japan. Here GPS data from two receivers recorded deformation several days before volcanic tremor or visible signs of eruption were apparent [48,145].

Loss of instruments in a volcanic eruption is a real and costly risk. InSAR, which does not require any equipment on the ground, is a widely used method for monitoring volcano deformation. However, InSAR does not provide three-component deformation measurements and does not have the temporal resolution that CGPS does. InSAR also suffers from decorrelation when vegetation, snowfall, or lava flows change the land surface during the time between two image acquisitions, however it still generally provides better spatial coverage than GPS. Dzurisin [40] presents a good overview of InSAR as applied to volcano deformation.

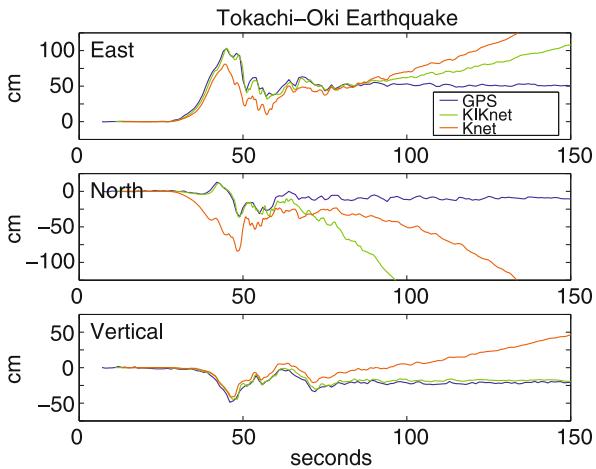
During a recent eruption of Augustine, a stratovolcano in Alaska, GPS data proved to be a valuable complement to other monitoring systems [20]. In the summer of 2005 CGPS stations on the volcano began to record an inflation signal, following an increase in microseismicity below the volcano that had begun in May 2005 or perhaps earlier. Beginning in November 2005 the rate of inflation increased rapidly, but then died off somewhat in early January 2006. The inflation signal has been interpreted as evidence for a dike intrusion into the volcanic edifice which nearly reached the surface [20]. On January 11, 2006 a series of explosive eruptions began at Augustine, destroying two of the six CGPS receivers. By January 17th, though, the eruptions had died down, as had seismicity and gas emissions. Only the GPS data from two of the remaining stations showed continued inflation. The quiescence lasted 10 days before another explosive eruption and effusive lava flows took place. This is an example in which GPS data provided early corroboration in the summer of 2005 that increased seismicity was due to magma movement into the volcanic edifice. Moreover, during the ten days of quiescence when other indicators such as seismic-

ity and gas measurements showed little activity, the GPS data showed that further eruptive activity was likely [20].

**High-Rate GPS** Continuous GPS networks typically record data at 15- or 30-second sampling rates, which is more than adequate for obtaining daily positions. One such CGPS network, the Southern California Integrated GPS Network (SCIGN), recorded data at 30-second intervals for the 1999  $M_w$  7.1 Hector Mine earthquake. Nikolaidis et al. [104] used these data to obtain positions at every observation epoch for the time spanning the earthquake and demonstrated that these observations, although aliased, agreed with those from nearby strong ground motion instruments. However, GPS receivers are capable of recording data at much higher rates, e. g. 1 Hz or greater. In recent years, for scientific as well as surveying and navigation purposes, an increasing number of CGPS sites have been set to record high-rate data and in some cases transmit them in real-time.

Unlike seismic data, GPS receivers provide a direct measure of displacement, and the instrument stays on scale even during the strong shaking of an earthquake. This makes it possible to obtain high-rate displacement time histories without the error introduced by integrating velocity or acceleration records from seismic instruments and without the data loss that occurs when the shaking exceeds the dynamic range of a seismic instrument (Fig. 16). Although the results of Nikolaidis et al. [104] hinted at a potential application of high-rate data, this kind of observation was not yet available for an earthquake. Such an event occurred in 2002 with the  $M_w$  7.9 Denali earthquake in Alaska. Larson et al. [78], obtained displacement-time histories of seismic waves from 1 Hz GPS data recorded during this event and showed they were in good agreement with those obtained by doubly integrating accelerometer data.

Wang et al. [165] demonstrated that GPS receivers operating at 1 Hz provide a faithful recording of signals with periods of 2 seconds or greater. However, GPS receivers will not replace seismometers for recording earthquakes since they cannot capture the higher-frequency components of the seismic waves. Rather, they provide a valuable complement to seismic data. For example, Emore et al. [42] show that high-rate GPS data can be used as a constraint when integrating accelerograms. Moreover, the combined use of seismic and high-rate GPS data extends the observable frequency and amplitude range of seismic waves considerably [78]. Miyazaki et al. [96] demonstrated that GPS seismograms measured by 1 Hz GPS could be used in a similar way to seismic data to infer the spatial and temporal progression of fault slip dur-



**GPS: Applications in Crustal Deformation Monitoring, Figure 16**  
 Comparison of displacement time series from a GPS station and doubly integrated accelerometer data from two nearby stations of the KIKNet and Knet networks in Japan for the 2003  $M_w$  8 Tokachi-Oki earthquake. The GPS data and that from the KIKNet site (blue and green curves) agree well for the first  $\sim$ 50 seconds following the arrival of the seismic signal (at  $\sim$ 30 seconds). The deviation of the accelerometer data after several seconds is likely due to low frequency noise that is amplified by the double integration required to produce displacements. Additionally, the Knet site (red curve) may be affected by tilting during shaking or other problems with sensor orientation, leading to the greater deviation between the red curve and those of the other two instruments. Reprinted with permission from [96] (copyright 2006, American Geophysical Union)

ing the earthquake. High-rate GPS observations can be especially useful in this capacity when used in combination with seismic data (e.g. [63]).

High-rate GPS also vastly improves our ability to track early postseismic deformation. Seismic data do not record any processes, like afterslip, that do not generate seismic waves, so postseismic deformation measurements must be made using other instruments including GPS, strainmeters, and InSAR. Until CGPS networks became common, it was generally not possible to deploy instruments for post-earthquake geodetic measurements until field crews could reach the affected area, days or weeks after the event. As a result, any change in measured position from the last pre-earthquake survey until the first post-earthquake survey contained the full coseismic signal and some portion of postseismic displacement. Even the “coseismic” displacements obtained from differencing daily positions from CGPS data collected the day before and the day after the earthquake can be contaminated by postseismic signals that began immediately after the event. High-rate GPS data, on the other hand, provide a means to record the deformation signal continuously, starting in the seconds af-

ter an earthquake. An example of this was the 2004  $M_w$  6 Parkfield earthquake. Langbein et al. [76] showed that rapid postseismic displacement began immediately after the event, and these authors were able to use the high-rate time series to separate the coseismic and postseismic displacements. For other events, including the 1966  $M_w \sim$ 6 Parkfield earthquake, the moment release estimated from seismic data has often been substantially lower than that estimated using geodetic data (e.g. [138]). However in the case of the 2004 Parkfield event the slip estimated from the coseismic portion of the GPS displacement signal had moment release in good agreement with that estimated from seismic data [76].

The availability and use of high-rate GPS data in real-time for geophysical applications is not yet widespread, but it is growing as more CGPS sites are installed, telemetry of high bandwidth data becomes more feasible, and approaches to processing the data are refined. These data have great potential for aiding real-time deformation monitoring on volcanoes, for early warning of major earthquakes, and for tsunami warning as well.

For example, Mattia et al. [87] highlight the value of high-rate real-time GPS data in a time-critical situation to quickly differentiate localized volcanic processes (in their case a new volcanic vent opening) from more far-reaching dangers (e.g., a feared flank failure that could have caused a local tsunami).

Real-time high-rate GPS data may soon also contribute to rapid earthquake warning and response. Hudnut et al. [61] envision a system consisting of a pair of GPS receivers straddling a fault such as the San Andreas which record data at high rate and transmit these measurements in real time. In the case of a major earthquake which, within seconds, caused surface offset on the fault of more than a few centimeters, the “GPS slip-sensor” would indicate that the event was large significantly sooner than could be expected from seismic data alone. This rapid information could be used to trigger preventive measures for critical infrastructure such as transportation systems (e.g., slowing trains to help prevent derailment) and factories (e.g. halting processes involving hazardous chemicals).

Blewitt et al. [14], using data from the 2004 Sumatra earthquake, demonstrated that GPS data available in real time, even if recorded at the traditional 30-second sampling rate, can be used to reliably estimate displacements larger than about 10 mm for great earthquakes. Such information would significantly improve the robustness of tsunami warning systems, at least for oceanwide tsunamis, by giving a better indication of the true moment magnitude of the earthquake more quickly (e.g. within 15 minutes) than can be achieved with seismic data.

The patterns of strong shaking that occur during an earthquake provide a good indication as to where damage will be the most severe. The U. S. Geological Survey has implemented a system called ShakeMap which rapidly generates a map of shaking intensity based on instrumental recordings. This is a valuable tool for emergency responders and scientists in the aftermath of an earthquake, and is available to the public as well (<http://earthquake.usgs.gov/eqcenter/shakemap/>). However, in many areas seismic instrument coverage is sparse, hindering the generation of accurate and detailed maps. Dreger et al. [36] demonstrate that seismic data can be used to rapidly generate a model of fault slip during an earthquake and that such a model can significantly improve the quality of ShakeMap produced for that event. This is especially true when the direction of earthquake rupture away from the hypocenter shows a preferred orientation along strike, since shaking will be stronger at locations that coincide with this directivity. Current work is focused on incorporation of real-time high-rate GPS data into the fault slip analysis with the ultimate goal of further refining ShakeMaps.

## Future Directions

As evidenced by the applications described in this article, GPS has become an indispensable tool for monitoring crustal deformation hazards and investigating the underlying processes. It provides an affordable means of obtaining surface displacement measurements with millimeter-level precision at any time of day anywhere on earth in all weather conditions with no line of sight requirement. GPS data complement other observations such as seismic and InSAR data by providing a direct measure of displacement in three dimensions that stays on scale. GPS can record deformation over temporal scales of seconds to decades, thus making it possible to track surface waves generated by an earthquake, postseismic deformation, slow slip events, volcanic unrest, fault creep, interseismic strain accumulation, and plate motions.

For the first decade or so, applications of GPS in crustal deformation tended toward periodic measurements in the style of earlier geodetic surveys. However, the size, power consumption, and cost of receivers have decreased steadily, while the data storage capacity has increased. This has fostered the growth of large CGPS networks, accompanied by more centralized and uniform processing and availability of results. This trend is likely to continue, and many future applications of GPS will focus on further exploiting the ability of this tool to provide temporally dense measurements for tracking deformation.

High-rate data show promise for hazard monitoring and response as well as for providing insight into the physical processes underlying time-varying deformation, especially as more sophisticated methods become available for processing these data in real-time and mitigating error sources such as multipath which are particularly problematic for high rate measurements.

In areas with few CGPS sites or where spatially dense coverage is needed to address specific scientific questions, an alternative to traditional campaign GPS measurements has recently been developed [13]. Termed “semi-permanent” or “semi-continuous” GPS, it involves rotating a pool of GPS receivers through several subsets of GPS sites such that each subset is observed for periods of multiple weeks several times a year. Each site to be observed is outfitted with a specially designed antenna mount set in a rock outcrop so that the antenna is attached in the same location and orientation every time the site is occupied, thus eliminating the set-up error inherent in many SGPS surveys. Blewitt et al. [13] have demonstrated that the scatter in the position time series as well as the uncertainties on velocities estimated using data from a semi-permanent network in Nevada are nearly as low as for CGPS networks over a comparable length of time (e.g., 1.5 years). Although this method requires sites that have rock outcrops and that are secure enough for equipment to be left unattended for extended periods of time, its use is likely to grow as it is arguably more cost-effective than either CGPS or SGPS for many applications.

Finally, a series of changes are being made to modernize the GPS signal. For instance, beginning with a generation of satellites called Block IIR-M, first launched in the fall of 2005, a civilian code is now modulated on the L2 carrier in addition to the C/A code that has always been available on the L1 carrier. This makes it easier to obtain the L2 carrier signal and enables elimination of ionospheric delay when positioning with the code measurements alone. A new series of GPS satellites, the Block IIF, set to launch in 2007 will transmit a third carrier frequency called L5 which will aid in ambiguity resolution. The L5 observable, modulated with a civilian-accessible code, will be broadcast at a higher power than the L1 or L2, making it easier to acquire. Future satellites will transmit a second civilian code on the L1 frequency that will be more robust than the C/A code and will be interoperable with a civilian accessible code to be transmitted by the planned Galileo satellites.

## Acknowledgments

John Langbein, Margaret Boettcher, Maurizio Battaglia, Emily Desmarais, Tom Hanks, and Fred Pollitz pro-

vided helpful comments which significantly improved this manuscript.

## Bibliography

### Primary Literature

1. Ammon C, Ji C, Thio HK, Robinson D, Ni S, Hjorleifsdottir V, Kanamori H, Lay T, Das S, Helmberger D, Ichinose G, Polet J, Wald D (2005) Rupture process of the 2004 Sumatra-Andaman earthquake. *Science* 308:1133–1139
2. Argus D, Gordon R (1991) No-net-rotation model of current plate velocities incorporating plate motion model NUVEL-1. *Geophys Res Lett* 18:2039–2042
3. Argus D, Gordon R, Ma C, Eanes R, Heflin M, Owen S, Willis P (2006) GEODEVEL: Plate motions from space geodesy. *Eos Trans AGU* 87(52) Fall Meet Suppl:Abstract G41A-06
4. Argus D, Heflin M (1995) Plate motion and crustal deformation estimated with geodetic data from the Global Positioning System. *Geophys Res Lett* 22:1973–1976
5. Árnadóttir T, Jónsson S, Pollitz F, Jiang W, Feigl K (2005) Post-seismic deformation following the June 2000 earthquake sequence in the south Iceland seismic zone. *J Geophys Res* 110:B12308; [doi:10.1029/2005JB003701](https://doi.org/10.1029/2005JB003701)
6. Banerjee P, Pollitz F, Bürgmann R (2005) The size and duration of the Sumatra-Andaman earthquake from far-field static offsets. *Science* 308:1769–1772
7. Banerjee P, Pollitz F, Nagarajan B, Bürgmann R (2007) Coseismic slip distributions of the 26 December 2004 Sumatra-Andaman and 28 March 2005 Nias earthquakes from GPS static offsets. *Bull Seismol Soc Amer* 97:S86–S102
8. Battaglia M, Segall P, Murray J, Cervelli P, Langbein J (2003) The mechanics of unrest at Long Valley caldera, California: 1. Modeling the geometry of the source using GPS, leveling and two-color EDM data. *J Volc Geotherm Res* 127:195–217
9. Battaglia M, Segall P, Roberts C (2003) The mechanics of unrest at Long Valley caldera, California: 2. Constraining the nature of the source using geodetic and micro-gravity data. *J Volc Geotherm Res* 127:219–245
10. Bennett R, Davis J, Wernicke B (1996) First results from the northern Basin and Range continuous GPS network. *Eos Trans AGU* 77(46) Fall Meet Suppl:150
11. Bilek S, Satake K, Sieh K (2007) Introduction to the special issue on the 2004 Sumatra-Andaman earthquake and the Indian Ocean tsunami. *Bull Seis Soc Amer* 97:S1–S5
12. Bilham R, Engdahl R, Feldl N, Satyabala S (2005) Partial and complete rupture of the Indo-Andaman plate boundary 1847–2004. *Seismol Res Lett* 76:299–311
13. Blewitt G, Hammond WC, Kreemer C (2009) Geodetic observation of contemporary strain in the northern Walker Lane: 1, Semi-permanent GPS strategy. In: Oldow JS, Cashman PH (eds) Late Cenozoic Structure and Evolution of the Great Basin – Sierra Nevada Transition. *Geol Soc Amer* (in press) [doi:10.1130/2009.2447\(1\)](https://doi.org/10.1130/2009.2447(1))
14. Blewitt G, Kreemer C, Hammond W, Plag HP, Stein S, Okal E (2006) Rapid determination of earthquake magnitude using GPS for tsunami warning systems. *Geophys Res Lett* 33:L11309; [doi:10.1029/2006GL026145](https://doi.org/10.1029/2006GL026145)
15. Bock Y, Nikollaids R, de Jonge P, Bevis M (2000) Instantaneous geodetic positioning at medium distances with the Global Positioning System. *J Geophys Res* 105:28223–28253
16. Bos A, Usai S, Spakman W (2004) A joint analysis of GPS motions and InSAR to infer the coseismic surface deformation of the Izmit, Turkey earthquake. *Geophys J Int* 158:849–863
17. Bürgmann R, Kogan M, Steblov G, Hillel G, Levin V, Apel E (2005) Interseismic coupling and asperity distribution along the Kamchatka subduction zone. *J Geophys Res* 110:B07405; [doi:10.1029/2005JB003648](https://doi.org/10.1029/2005JB003648)
18. Bürgmann R, Rosen P, Fielding E (2000) Synthetic aperture radar interferometry to measure Earth's surface topography and its deformation. *Ann Rev Earth Planet Sci* 28:169–209
19. Calais E, Han JY, DeMets C, Nocquet JM (2006) Deformation of the North American plate interior from a decade of continuous GPS measurements. *J Geophys Res* 111:B06402; [doi:10.1029/2005JB004253](https://doi.org/10.1029/2005JB004253)
20. Cervelli P, Fournier T, Freymueller J, Power J (2006) Ground deformation associated with the precursory unrest and early phases of the January 2006 eruption of Augustine Volcano, Alaska. *Geophys Res Lett* 33:L18304; [doi:10.1029/2006GL027219](https://doi.org/10.1029/2006GL027219)
21. Cervelli P, Murray M, Segall P, Aoki Y, Kato T (2001) Estimating source parameters from deformation data, with an application to the March 1997 earthquake swarm off the Izu Peninsula, Japan. *J Geophys Res* 106:11217–11237
22. Cervelli P, Segall P, Johnson K, Lisowski M, Miklius A (2002) Sudden aseismic fault slip on the south flank of Kilauea volcano. *Nature* 415:1014–1018
23. Chen Q, Freymueller J, Wang Q, Yang Z, Xu C, Liu J (2004) A deforming block model for the present-day tectonics of Tibet. *J Geophys Res* 109:B01403; [doi:10.1029/2002JB002151](https://doi.org/10.1029/2002JB002151)
24. Chlieh M, Avouac JP, Hjorleifsdottir V, Song TR, Ji C, Sieh K, Sladen A, Hebert H, Prawirodirdjo L, Bock Y, Galetzka J (2007) Coseismic slip and afterslip of the great  $M_w$  9.15 Sumatra-Andaman earthquake of 2004. *Bull Seis Soc Amer* 97:S152–S173
25. Coe JA, Ellis WL, Godt JW, Savage WZ, Savage JE, Michael JA, Kibler JD, Powers PS, Lidke DJ, Debray S (2003) Seasonal movement of the Slumgullion landslide determined from Global Positioning System surveys and field instrumentation, July 1998 – March 2002. *Eng Geol* 68:67–101
26. d'Alessio M, Johanson I, Bürgmann R, Schmidt D, Murray M (2005) Slicing up the San Francisco Bay Area: Block kinematics and fault slip rates from GPS-derived surface velocities. *J Geophys Res* 110:B06403; [doi:10.1029/2004JB003496](https://doi.org/10.1029/2004JB003496)
27. Delouis B, Giardini D, Lundgren P, Salichon J (2002) Joint inversion of InSAR, GPS, teleseismic, and strong-motion data for the spatial and temporal distribution of earthquake slip; application to the 1999 Izmit mainshock. *Bull Seis Soc Amer* 92:278–299
28. DeMets C, Gordon R, Argus D (2006) Moving beyond NUVEL-1A: The MORVEL estimates of geologically recent global plate motions. *Eos Trans AGU* 87(52) Fall Meet Suppl:Abstract G41A-05
29. DeMets C, Gordon R, Argus D, Stein S (1990) Current plate motions. *Geophys J Int* 101:425–478
30. DeMets C, Gordon R, Argus D, Stein S (1994) Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions. *Geophys Res Lett* 21:2191–2194
31. Deng J, Gurnis M, Kanamori H, Hauksson E (1998) Viscoelastic flow in the lower crust after the 1992 Landers, California, earthquake. *Science* 282:1689–1692

32. Dieterich J, Decker R (1975) Finite element modeling of surface deformation associated with volcanism. *J Geophys Res* 80:4094–4102
33. Dong D, Fang P, Bock Y, Cheng MK, Miyazaki S (2002) Anatomy of apparent seasonal variations from GPS-derived site position time series. *J Geophys Res* 107:2075; doi:[10.1029/2001JB000573](https://doi.org/10.1029/2001JB000573)
34. Douglas A, Beavan J, Wallace L, Townend J (2005) Slow slip on the northern Hikurangi subduction interface, New Zealand. *Geophys Res Lett* 32:L16305; doi:[10.1029/2005GL023607](https://doi.org/10.1029/2005GL023607)
35. Dragert H, Wang K, James TS (2001) A silent slip event on the deeper Cascadia subduction interface. *Science* 292:1525–1528
36. Dreger D, Gee L, Lombard P, Murray M, Romanowicz B (2005) Rapid finite-source analysis and near-fault strong ground motions: Application to the 2003  $M_w$  6.5 San Simeon and 2004  $M_w$  6.0 Parkfield earthquakes. *Seismol Res Lett* 76:40–48
37. Du Y, Segall P, Gao H (1997) Quasi-static dislocations in three dimensional inhomogeneous media. *Geophys Res Lett* 24:2347–2350
38. Dzurisin D (1992) Geodetic leveling as a tool for studying restless volcanoes. In: Ewert J, Swanson D (eds) *Monitoring volcanoes: Techniques and strategies used by the staff of the Cascades Volcano Observatory, 1980–1990*, USGS Bull. 1966, US Geological Survey, Reston, VA, pp 125–134
39. Dzurisin D (2003) A comprehensive approach to monitoring volcano deformation as a window on the eruption cycle. *Rev Geophys* 41; doi:[10.1029/2001RG000107](https://doi.org/10.1029/2001RG000107)
40. Dzurisin D (2007) *Volcano deformation: Geodetic monitoring techniques*. Springer, New York
41. Eberhart-Phillips D, Michael AJ (1993) Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield region, central California. *J Geophys Res* 98:15737–15758
42. Emore G, Haase J, Choi K, Larson K, Yamagiwa A (2007) Recovering seismic displacements through combined use of 1-Hz GPS and strong-motion accelerometers. *Bull Seismol Soc Amer* 97:357–378; doi:[10.1785/0120060153](https://doi.org/10.1785/0120060153)
43. Fialko Y (2004) Evidence of fluid-filled upper crust from observations of postseismic deformation due to the 1992  $M_w$  7.3 Landers earthquake. *J Geophys Res* 109:B08401; doi:[10.1029/2004JB002985](https://doi.org/10.1029/2004JB002985)
44. Fletcher H, Beavan J, Freymueller J, Gilbert L (2001) High interseismic coupling of the Alaska subduction zone SW of Kodiak island inferred from GPS data. *Geophys Res Lett* 28:443–446
45. Freed A, Bürgmann R (2004) Evidence of power-law flow in the Mojave desert mantle. *Nature* 430:548–551
46. Freed A, Bürgmann R, Calais E, Freymueller J, Hreinsdóttir S (2006) Implications of deformation following the 2002 Denali, Alaska, earthquake for postseismic relaxation processes and lithospheric rheology. *J Geophys Res* 111:B01401; doi:[10.1029/2005JB003894](https://doi.org/10.1029/2005JB003894)
47. Freymueller J, Murray M, Segall P, Castillo D (1999) Kinematics of the Pacific-North America plate boundary zone, northern California. *J Geophys Res* 104:7419–7441
48. Fujinawa Y, Shimada S, Ohmi S, Sekiguchi S, Eguchi T, Okada Y (1991) Fixed point GPS observation of crustal movement associated with the 1989 seismic swarm and submarine volcanic activities of Ito, central Japan. *J Phys Earth* 39:141–153
49. Gahalaut VK, Nagarajan B, Catherine JK, Kumar S (2006) Constraints on 2004 Sumatra–Andaman earthquake rupture from GPS measurements in Andaman-Nicobar Islands. *Earth Planet Sci Lett* 242:365–374
50. Gili JA, Corominas J, Rius J (2000) Using Global Positioning System techniques in landslide monitoring. *Engineering Geology* 55:167–192
51. Gladwin M, Gwyther R, Hart R, Breckenridge K (1994) Measurements of the strain field associated with episodic creep events on the San Andreas fault near San Juan Bautista, California. *J Geophys Res* 99:4559–4565
52. Gwyther RL, Gladwin MT, Mee M, Hart RHG (1996) Anomalous shear strain at Parkfield during 1993–94. *Geophys Res Lett* 23:2425–2428
53. Hashimoto M, Hashizume M, Takemoto S, Fukada Y, Fujimori K, Takiguchi H, Satomura M, Otsuka Y, Saito S (2006) Postseismic deformations following the Sumatra–Andaman and Nias earthquakes detected by continuous GPS observation in SE Asia. *Seism Res Lett* 77:289
54. Hernandez B, Cotton F, Campillo M (1999) Contribution of radar interferometry to a two-step inversion of the kinematic process of the 1992 Landers earthquake. *J Geophys Res* 104:13083–13099
55. Hirose H, Obara K (2006) Short-term slow slip and correlated tremor episodes in the Tokai region, central Japan. *Geophys Res Lett* 33:L17311; doi:[10.1029/2006GL026579](https://doi.org/10.1029/2006GL026579)
56. Hofmann-Wellenhof B, Lichtenegger H, Collins J (2001) *Global Positioning System theory and practice*, 5th edn. Springer, New York
57. Hreinsdóttir S, Freymueller JT, Bürgmann R, Mitchell J (2006) Coseismic deformation of the 2002 Denali Fault earthquake: Insights from GPS measurements. *J Geophys Res* 111:B03308; doi:[10.1029/2005JB003676](https://doi.org/10.1029/2005JB003676)
58. Hreinsdóttir S, Freymueller JT, Fletcher HJ, Larsen CF, Bürgmann R (2003) Coseismic slip distribution of the 2002  $M_w$  7.9 Denali Fault earthquake, Alaska, determined from GPS measurements. *Geophys Res Lett* 30:1670; doi:[10.1029/2003GL017447](https://doi.org/10.1029/2003GL017447)
59. Hsu Y, Simons M, Avouac J-P, Galetzka J, Sieh K, Chlieh M, Natawidjaja D, Prawirodirido L, Bock Y (2006) Frictional after-slip following the 2005 Nias-Simeulue earthquake, Sumatra. *Science* 312:1921–1926
60. Hudnut K (1997) The Southern California Integrated GPS Network (SCIGN). Open-File Report, US Geological Survey, Report: OF 97-0467:10-13
61. Hudnut K, Anderson G, Aspiotes A, King N, Moffitt R, Stark K (2002) GPS fault slip sensors. APEC Symposium on Confronting Urban Earthquakes/Seismic Early Warning. Academia Sinica, Taipei, pp 93–96
62. Ihmlé PF, Jordan TH (1994) Teleseismic search for slow precursors to large earthquakes. *Science* 266:1547–1551
63. Ji C, Larson K, Tan Y, Hudnut K, Choi K (2004) Slip history of the 2003 San Simeon earthquake constrained by combining 1-Hz GPS, strong motion, and teleseismic data. *Geophys Res Lett* 31:L17608; doi:[10.1029/2004GL020448](https://doi.org/10.1029/2004GL020448)
64. Johanson I, Bürgmann R (2005) Creep and quakes on the northern transition zone of the San Andreas fault from GPS and InSAR data. *J Geophys Res* 32:L14306; doi:[10.1029/2005GL023150](https://doi.org/10.1029/2005GL023150)
65. Johnson K, Segall P (2004) Imaging the ramp-décollement geometry of the Chelungpu fault using coseismic GPS displacements from the 1999 Chi-Chi, Taiwan earthquake. *Tectonophysics* 378:123–139

66. Johnson K, Segall P (2004) Viscoelastic cycle models of deep stress driven creep along the San Andreas Fault. *J Geophys Res* 109; [doi:10.1029/2004JB003096](https://doi.org/10.1029/2004JB003096)
67. Johnson K, Segall P (2005) A viscoelastic earthquake cycle model for Taiwan. *J Geophys Res* 110:B10404; [doi:10.1029/2004JB003516](https://doi.org/10.1029/2004JB003516)
68. Johnston M, Linde A (2002) Implications of crustal strain during conventional, slow, and silent earthquakes. *Handbook of Earthquake and Engineering Seismology* 81A:589–605
69. Jónsson S, Segall P, Pedersen R, Björnsson G (2003) Post-earthquake ground movements correlated to pore-pressure transients. *Nature* 424:179–183
70. Jónsson S, Zebker H, Segall P, Amelung F (2002) Fault Slip Distribution of the 1999  $M_w$  7.1 Hector Mine, California, Earthquake, estimated from Satellite Radar and GPS Measurements. *Bull Seis Soc Amer* 92:1377–1389
71. Kaverina A, Dreger D, Price E (2002) The combined inversion of seismic and geodetic data for the source process of the 16 October 1999  $M_w$  7.1 Hector Mine, California, earthquake. *Bull Seis Soc Amer* 92:1266–1280
72. King N, Murray M, Prescott W, Clymer R, Romanowicz B (1994) The Bay Area Regional Deformation (BARD) permanent GPS array. *Eos Trans AGU* 75(44) Fall Meet Suppl:470
73. Kreemer C, Blewitt G, Hammond W, Plag HP (2006) Global deformation from the great 2004 Sumatra–Andaman earthquake observed by GPS: Implications for rupture process and global reference frame. *Earth Planets Space* 58:141–148
74. Langbein J (2004) Noise in two-color electronic distance meter measurements revisited. *J Geophys Res* 109:B04406; [doi:10.1029/2003JB002819](https://doi.org/10.1029/2003JB002819)
75. Langbein J, Gwyther RL, Hart RHG, Gladwin MT (1999) Slip-rate increase at Parkfield in 1993 detected by high-precision EDM and borehole tensor strainmeters. *Geophys Res Lett* 26:2529–2532
76. Langbein J, Murray J, Snyder HA (2006) Coseismic and initial postseismic deformation from the 2004 Parkfield, California, earthquake, observed by Global Positioning System, electronic distance meter, creepmeters, and borehole strainmeters. *Bull Seismol Soc Amer* 96:S304–S320; [doi:10.1785/0120050823](https://doi.org/10.1785/0120050823)
77. Larson K (1995) Crustal deformation. *Rev Geophys* 33:371–378; [doi:10.1029/95RG00439](https://doi.org/10.1029/95RG00439)
78. Larson K, Bodin P, Gomberg J (2003) Using 1-Hz GPS data to measure deformations caused by the Denali fault earthquake. *Science* 300:1421–1424; [doi:10.1126/science.1084531](https://doi.org/10.1126/science.1084531)
79. Larson K, Freymueller J, Philipsen S (1997) Global plate velocities from the Global Positioning System. *J Geophys Res* 102:9961–9981
80. Larson K, van Dam T (2000) Measuring postglacial rebound with GPS and absolute gravity. *Geophys Res Lett* 27:3925–3928
81. Lay T, Kanamori H, Ammon C, Nettles M, Ward S, Aster R, Beck S, Bilek S, Brudzinski M, Butler R, DeShon H, Ekstrom G, Satake K, Sipkin S (2005) The great Sumatra–Andaman earthquake of 26 December 2004. *Science* 308:1127–1133
82. Linde A, Gladwin M, Johnston M, Gwyther R, Bilham R (1996) A slow earthquake sequence on the San Andreas fault. *Nature* 383:65–68
83. Lowry A (2006) Resonant slow fault slip in subduction zones forced by climatic load stress. *Nature*, 442:802–805
84. Maerten F, Resor P, Pollard D, Maerten L (2005) Inverting for slip on three-dimensional fault surfaces using angular dislocations. *Bull Seismol Soc Amer* 95:1654–1665
85. Manaker D, Bürgmann R, Prescott W, Langbein J (2003) Distribution of interseismic slip rates and the potential for significant earthquakes on the Calaveras fault, central California. *J Geophys Res* 108:B62287; [doi:10.1029/2002JB001749](https://doi.org/10.1029/2002JB001749)
86. Mao A, Harrison C, Dixon T (1999) Noise in GPS coordinate time series. *J Geophys Res* 104:2797–2816
87. Mattia M, Rossi M, Guglielmino F, Aloisi M, Bock Y (2004) The shallow plumbing system of Stromboli Island as imaged from 1 Hz instantaneous GPS positions. *Geophys Res Lett* 31:L24610; [doi:10.1029/2004GL021281](https://doi.org/10.1029/2004GL021281)
88. Mazzotti S, James TS, Henton J, Adams J (2005) GPS crustal strain, postglacial rebound, and seismic hazard in eastern North America: The Saint Lawrence valley example. *J Geophys Res* 110:B11301; [doi:10.1029/2004JB003590](https://doi.org/10.1029/2004JB003590)
89. McCaffrey R (2005) Block kinematics of the Pacific–North America plate boundary in southwestern United States from inversion of GPS, seismological, and geologic data. *J Geophys Res* 110:B07401; [doi:10.1029/2004JB003307](https://doi.org/10.1029/2004JB003307)
90. Meade B, Hager B (2005) Block models of crustal motion in southern California constrained by GPS measurements. *J Geophys Res* 110:B03403; [doi:10.1029/2004JB003209](https://doi.org/10.1029/2004JB003209)
91. Meltzner A, Sieh K, Abrams M, Agnew D, Hudnut K, Avouac JP, Natawidjaja DH (2006) Uplift and subsidence associated with the great Aceh–Andaman earthquake of 2004. *J Geophys Res* 111:B02407; [doi:10.1029/2005JB003891](https://doi.org/10.1029/2005JB003891)
92. Miller M, Johnson D, Rubin C, Dragert H, Endo E, Humphreys E, Nabelek J, Qamar A (1997) GPS Monitoring of the Cascadia Margin: The Pacific Northwest Geodetic Array (PANGA). *Eos Trans AGU* 78(46) Fall Meet Suppl:167
93. Miller M, Johnson D, Rubin C, Dragert H, Wang K, Qamar A, Goldfinger C (2001) GPS-determination of along-strike variation in Cascadia margin kinematics: Implications for relative plate motion, subduction zone coupling, and permanent deformation. *Tectonics* 20:161–176
94. Miller M, Melbourne T, Johnson D, Sumner W (2002) Periodic slow earthquakes from the Cascadia subduction zone. *Science* 295:2423
95. Milne GA, Davis JL, Mitrovica JX, Scherneck HG, Johansson JM, Vermeer M, Koivula H (2001) Space-geodetic constraints on glacial isostatic adjustment in Fennoscandia. *Science* 291:2381–2385
96. Miyazaki S, Larson K, Choi K, Hikima K, Koketsu K, Bodin P, Haase J, Emore G, Yamagiwa A (2004) Modeling the rupture process of the 2003 September 25 Tokachi-Oki (Hokkaido) earthquake using 1-Hz GPS data. *Geophys Res Lett* 31:L21603; [doi:10.1029/2004GL021457](https://doi.org/10.1029/2004GL021457)
97. Miyazaki S, Segall P, Fukuda J, Kato T (2004) Space time distribution of afterslip following the 2003 Tokachi-oki earthquake: Implications for variations in fault zone frictional properties. *Geophys Res Lett* 31:L06623; [doi:10.1029/2003GL019410](https://doi.org/10.1029/2003GL019410)
98. Miyazaki S, Segall P, McGuire J, Kato T, Hatanaka Y (2006) Spatial and temporal evolution of stress and slip rate during the 2000 Tokai slow earthquake. *J Geophys Res* 111:B03409; [doi:10.1029/2004JB003426](https://doi.org/10.1029/2004JB003426)
99. Mogi K (1958) Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. *Bull Seismol Soc Amer* 36:111–123
100. Mora P, Baldi P, Casula G, Fabris M, Ghirotti M, Mazzini E, Pesci

- A (2003) Global Positioning Systems and digital photogrammetry for the monitoring of mass movements: Application to the Ca' di Malta landslide (northern Appenines, Italy). *Eng Geol* 68:103–121
101. Murray J, Langbein J (2006) Slip on the San Andreas fault at Parkfield, California, over two earthquake cycles, and the implications for seismic hazard. *Bull Seismol Soc Amer* 96:S283–S303
  102. Murray J, Segall P (2005) Spatiotemporal evolution of a slip-rate increase on the San Andreas fault near Parkfield, CA. *J Geophys Res* 110:B09407; [doi:10.1029/2005JB003651](https://doi.org/10.1029/2005JB003651)
  103. Murray J, Segall P, Cervelli P, Prescott W, Svarc J (2001) Inversion of GPS data for spatially variable slip-rate on the San Andreas Fault near Parkfield, CA. *Geophys Res Lett* 28:359–362
  104. Nikolaidis R, Bock Y, de Jonge P, Shearer P, Agnew D, Domseilaar M (2001) Seismic wave observations with the Global Positioning System. *J Geophys Res* 106:21897–21916
  105. Obara K (2002) Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science* 296:1679–1681
  106. Obara K, Hirose H, Yamamizu F, Kasahara K (2004) Episodic slow slip events accompanied by non-volcanic tremors in southwest Japan subduction zone. *Geophys Res Lett* 31; [doi:10.1029/2004GL020848](https://doi.org/10.1029/2004GL020848)
  107. Ohta Y, Freymueller J, Hreinsdóttir S, Suito H (2006) A large slow slip event and the depth of the seismogenic zone in the south central Alaska subduction zone. *Earth Plan Sci Lett* 247:108–116
  108. Ohta Y, Kimata F, Sagiya T (2004) Reexamination of the interplate coupling in the Tokai region, central Japan, based on the GPS data in 1997–2002. *Geophys Res Lett* 31:L24604; [doi:10.1029/2004GL021404](https://doi.org/10.1029/2004GL021404)
  109. Okada Y (1985) Surface deformation due to shear and tensile faults in a half-space. *Bull Seismol Soc Amer* 75:1135–1154
  110. Okada Y (1992) Internal deformation due to shear and tensile faults in a half-space. *Bull Seismol Soc Amer* 82:1018–1040
  111. Owen S, Segall P, Lisowski M, Miklius A, Denlinger R, Sako M (2000) Rapid deformation of Kilauea volcano: GPS measurements between 1990 and 1996. *J Geophys Res* 105:18983–18998
  112. Owen S, Segall P, Lisowski M, Murray M, Bevis M, Foster J (2000) The January 30, 1997 eruptive event on Kilauea Volcano, Hawaii, as monitored by continuous GPS. *Geophys Res Lett* 27:2757–2760
  113. Ozawa S, Murakami M, Fujiwara S, Tobita M (1997) Synthetic aperture radar interferogram of the 1995 Kobe earthquake and its geodetic inversion. *Geophys Res Lett* 24:2327–2330
  114. Ozawa S, Suito H, Imakiire T, Murakmi M (2007) Spatiotemporal evolution of aseismic interplate slip between 1996 and 1998 and between 2002 and 2004, in Bungo channel, southwest Japan. *J Geophys Res* 112:B05409; [doi:10.1029/2006JB004643](https://doi.org/10.1029/2006JB004643)
  115. Park J, Anderson K, Aster R, Butler R, Lay T, Simpson D (2005) Global seismographic network records the great Sumatra–Andaman earthquake. *Eos Trans AGU* 86:57, 60–61
  116. Park J, Song TR, Tromp J, Okal E, Stein S, Roult G, Clevede E, Laske G, Kanamori H, Davis P, Berger J, Braatenberg C, Van Camp M, Lei X, Sun H, Xu H, Rosat S (2005) Earth's free oscillations excited by the 26 December 2004 Sumatra–Andaman earthquake. *Science* 308:1139–1144
  117. Park K, Nerem RS, Davis JL, Schenewerk MS, Milne GA, Mitrovica JX (2002) Investigation of glacial isostatic adjustment in the northeast US using GPS measurements. *Geophys Res Lett* 29:1509; [doi:10.1029/2001GL013782](https://doi.org/10.1029/2001GL013782)
  118. Peltzer G, Rosen P, Rogez F, Hudnut K (1998) Poroelastic rebound along the Landers 1992 earthquake surface rupture. *J Geophys Res* 103:30131–30145
  119. Poland M, Hamburger M, Newman A (2006) The changing shapes of active volcanoes: History, evolution, and future challenges for volcano geodesy. *J Volc Geotherm Res* 150:1–13
  120. Polet J, Kanamori H (2000) Shallow subduction zone earthquakes and their tsunamigenic potential. *Geophys J Int* 142:684–702
  121. Pollitz F (2005) Transient rheology of the upper mantle beneath central Alaska inferred from the crustal velocity field following the 2002 Denali earthquake. *J Geophys Res* 110:B08407; [doi:10.1029/2005JB003672](https://doi.org/10.1029/2005JB003672)
  122. Pollitz F, Wicks C, Thatcher W (2001) Mantle flow beneath a continental strike-slip fault: postseismic deformation after the 1999 Hector Mine earthquake. *Science* 293:1814–1818
  123. Pratt T (2006) Do Episodic Tremor and Slip (ETS) Events Affect Seismicity in the Northern Cascadia Subduction Zone? *Eos Trans AGU* 87(52), Fall Meet Suppl, Abstract T54A-04
  124. Prawirodiridjo L, Bock Y (2004) Instantaneous global plate motion model from 12 years of continuous GPS observations. *J Geophys Res* 109:B08405; [doi:10.1029/2003JB002944](https://doi.org/10.1029/2003JB002944)
  125. Prescott W, Savage J, Svarc J, Manaker D (2001) Deformation across the Pacific–North America plate boundary near San Francisco, California. *J Geophys Res* 106:6673–6682
  126. Pritchard M, Norabuena E, Ji C, Boroschek R, Comte D, Simons M, Dixon T, Rosen P (2007) Geodetic, teleseismic, and strong motion constraints on slip from recent southern Peru subduction zone earthquakes. *J Geophys Res* 112:B03307; [doi:10.1029/2006JB004294](https://doi.org/10.1029/2006JB004294)
  127. Reid HF (1910) The California Earthquake of April 18, 1906. In: Report of the state earthquake investigation commission, vol 2, Carnegie Institute, Washington DC
  128. Reid M, LaHusen R, Schmidt K (2004) Capturing 3-D displacements in active landslides using GPS. Abstracts with Programs. *Geol Soc Amer* 36:331
  129. Rhee J, Dreger D, Bürgmann R, Romanowicz B (2007) Slip of the 2004 Sumatra–Andaman earthquake from joint inversion of long-period global seismic waveforms and GPS static offsets. *Bull Seismol Soc Amer* 97:S115–S127
  130. Rogers G, Dragert H (2003) Episodic tremor and slip on the Cascadia subduction zone: the chatter of silent slip. *Science* 300:1942–1943
  131. Sagiya T (2004) Interplate coupling in the Kanto district, central Japan, and the Boso Peninsula silent earthquake in May 1996. *Pure Appl Geophys* 161:2327–2342; [doi:10.1007/s00024-004-2566-6](https://doi.org/10.1007/s00024-004-2566-6)
  132. Salichon J, Lundgren P, Delouis B, Giardini D (2004) Slip history of the 16 October 1999  $M_w$  7.1 Hector Mine earthquake (California) from the inversion of InSAR, GPS, and teleseismic data. *Bull Seismol Soc Amer* 94:2015–2027
  133. Savage J, Prescott W (1978) Asthenosphere readjustment and the earthquake cycle. *J Geophys Res* 83:3369–3376
  134. Savage J, Svarc J, Yu SB (2005) Postseismic relaxation and transient creep. *J Geophys Res* 110:B11402; [doi:10.1029/2005JB003687](https://doi.org/10.1029/2005JB003687)
  135. Schmalzle G, Dixon T, Malservisi R, Govers R (2006) Strain accumulation across the Carrizo segment of the San Andreas

- fault, California: Impact of laterally varying crustal properties. *J Geophys Res* 111:B05403; doi:[10.1029/2005JB003843](https://doi.org/10.1029/2005JB003843)
136. Schmidt D, Bürgmann R, Nadeau R, d'Alessio M (2005) Distribution of aseismic slip rate on the Hayward fault inferred from seismic and geodetic data. *J Geophys Res* 110:B08406; doi:[10.1029/2004JB003397](https://doi.org/10.1029/2004JB003397)
137. Segall P (2002) Integrating geologic and geodetic estimates of slip rate on the San Andreas Fault system. *Int Geol Rev* 44:62–82
138. Segall P, Harris R (1986) Slip deficit on the San Andreas fault at Parkfield, California, as revealed by inversion of geodetic data. *Science* 233:1409–1413
139. Segall P, Desmarais E, Shelly D, Miklius A, Cervelli P (2006) Earthquakes triggered by silent slip events on Kilauea volcano, Hawaii. *Nature* 442:71–74; doi:[10.1038/nature04938](https://doi.org/10.1038/nature04938)
140. Sella G, Dixon T, Mao A (2002) REVEL: A model for recent plate velocities from space geodesy. *J Geophys Res* 107:B42081; doi:[10.1029/2000JB000033](https://doi.org/10.1029/2000JB000033)
141. Sella G, Stein S, Dixon T, Craymer M, James T, Mazzotti S, Dokka R (2007) Observation of glacial isostatic adjustment in “stable” North America with GPS. *Geophys Res Lett* 34: L02306; doi:[10.1029/2006GL027081](https://doi.org/10.1029/2006GL027081)
142. Shelly D, Beroza G, Ide S (2007) Non-volcanic tremor and low-frequency earthquake swarms. *Nature* 446:305–307
143. Shelly D, Beroza G, Ide S, Nakamura S (2006) Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. *Nature* 442:188–191
144. Shimada S, Bock Y (1992) Crustal deformation measurements in central Japan determined by a Global Positioning System fixed point network. *J Geophys Res* 97:12437–12455
145. Shimada S, Fujinawa Y, Sekiguchi S, Ohmi S, Eguchi T, Okada Y (1990) Detection of a volcanic fracture in Japan using Global Positioning System measurements. *Nature* 343:631–633
146. Sieh K, Jahns RH (1984) Holocene activity of the San Andreas fault at Wallace Creek, California. *Geol Soc Amer Bull* 95:883–896
147. Silver P, Bock Y, Agnew D, Henyey T, Linde A, McEvilly T, Minister JB, Romanowicz B, Sachs I, Smith R, Solomon S, Stein S (1999) A Plate Boundary Observatory. *IRIS Newsletter* XVI:3–9
148. Simons M, Fialko Y, Rivera L (2002) Coseismic deformation from the 1999  $M_w$  7.1 Hector Mine, California, earthquake as inferred from InSAR and GPS observations. *Bull Seismol Soc Amer* 92:1390–1402
149. Sims JD (1990) Geologic map of the San Andreas fault in the Parkfield 7.5-minute quadrangle, Monterey and Fresno counties, California. US Geol Surv Misc Field Studies Map MF-2115
150. Snay R, Cline M, Dillinger W, Foote R, Hilla S, Kass W, Ray J, Rohde J, Sella G, Soler T (2007) Using global positioning system-derived crustal velocities to estimate rates of absolute sea level change from North American tide gauge records. *J Geophys Res* 112:B04409; doi:[10.1029/2006JB004606](https://doi.org/10.1029/2006JB004606)
151. Squarzoni C, Delacourt C, Allemand P (2005) Differential single-frequency GPS monitoring of the La Valette landslide (French Alps). *Eng Geol* 79:215–229
152. Stein S, Okal E (2005) Speed and size of the Sumatra earthquake. *Nature* 434:581–582
153. Subarya C, Chlieh M, Prawirodirdjo L, Avouac JP, Bock Y, Sieh K, Meltzner AJ, Natawidjaja DH, McCaffrey R (2006) Plate-boundary deformation associated with the great Sumatra-Andaman earthquake. *Nature* 440:46–51
154. Suwa Y, Miura S, Hasegawa A, Sato T, Tachibana K (2006) Interplate coupling beneath NE Japan inferred from three-dimensional displacement field. *J Geophys Res* 111:B04402; doi:[10.1029/2004JB003203](https://doi.org/10.1029/2004JB003203)
155. Szeliga W, Melbourne T, Miller M, Santillan V (2004) Southern Cascadia episodic slow earthquakes. *Geophys Res Lett* 31:L16602; doi:[10.1029/2004GL020824](https://doi.org/10.1029/2004GL020824)
156. Thatcher W (1983) Nonlinear strain buildup and the earthquake cycle on the San Andreas fault. *J Geophys Res* 88:5893–5902
157. Thatcher W (1995) Microplate versus continuum descriptions of active tectonic deformation. *J Geophys Res* 100:3885–3894
158. Thatcher W (2007) Microplate model for the present-day deformation of Tibet. *J Geophys Res* 112:B01401; doi:[10.1029/2005JB004244](https://doi.org/10.1029/2005JB004244)
159. Thurber C, Zhang H, Waldhauser F, Hardebeck J, Michael A, Eberhart-Phillips D (2006) Three-dimensional compressional wavespeed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region. *Bull Seismol Soc Amer* 96:S38–S49
160. Vigny C, Simons WJF, Abu S, Bamphenyu R, Satirapod C, Choosakul N, Subarya C, Socquet A, Omar K, Abidin HZ, Ambrosius BAC (2005) Insight into the 2004 Sumatra-Andaman earthquake from GPS measurements in southeast Asia. *Nature* 436:201–206
161. Wahba G (1990) Spline Models for Observational Data. SIAM, Philadelphia PA
162. Wald D, Heaton T (1994) Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bull Seismol Soc Amer* 84:668–691
163. Waldhauser F, Ellsworth W (2000) A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. *Bull Seismol Soc Amer* 90:1353–1368
164. Wallace L, Beavan J, McCaffrey R, Darby D (2004) Subduction zone coupling and tectonic block rotations in the North Island, New Zealand. *J Geophys Res* 109:B12406; doi:[10.1029/2004JB003241](https://doi.org/10.1029/2004JB003241)
165. Wang GQ, Boore D, Tang G, Zhou X (2007) Comparisons of ground motions from collocated and closely spaced one-sample-per-second Global Positioning System and accelerograph recordings of the 2003 M 6.5 San Simeon, California, earthquake in the Parkfield region. *Bull Seismol Soc Amer* 97:76–90
166. Williams S, Bock Y, Fang P, Jamason P, Nikolaidis R, Prawirodirdjo L, Miller M, Johnson D (2004) Error analysis of continuous GPS position time series. *J Geophys Res* 109; doi:[10.1029/2003JB002741](https://doi.org/10.1029/2003JB002741)
167. Working Group on California Earthquake Probabilities (2003) Earthquake probabilities in the San Francisco Bay Region: 2002–2031. US Geol Surv Open File Report 03-214
168. Wright T, Lu Z, Wicks C (2004) Constraining the slip distribution and fault geometry of the  $M_w$  7.9, 3 November 2002, Denali Fault earthquake with interferometric synthetic aperture radar and Global Positioning System data. *Bull Seismol Soc Amer* 94:175–189
169. Yagi Y, Kikuchi M, Nishimura T (2003) Co-seismic slip, post-seismic slip, and largest aftershock associated with the 1994 Sanriku-haruka-oki, Japan, earthquake. *Geophys Res Lett* 30:2177; doi:[10.1029/2003GL018189](https://doi.org/10.1029/2003GL018189)
170. Zhang J, Bock Y, Johnson H, Fang P, Williams S, Genrich J, Wdowinski S, Behr J (1997) Southern California permanent

- GPS geodetic array: Error analysis of daily position estimates and site velocities. *J Geophys Res* 102:18035–18055
171. Zhang PZ, Shen Z, Wang M, Gan W, Bürgmann R, Molnar P, Wang Q, Niu Z, Sun J, Wu J, Hanrong S, Xinzhaoy Y (2004) Continuous deformation of the Tibetan Plateau from global positioning system data. *Geol* 32:809–812

## Books and Reviews

- Bolt B (1999) Earthquakes, 4th edn. W H Freeman and Company, New York
- Menke W (1989) Geophysical data analysis: Discrete inverse theory, rev edn. In: Dmowska R, Holton J (eds) International geophysics series 45. Academic Press, San Diego
- Misra P, Enge P (2001) Global Positioning System: Signals, measurements, and performance. Ganga-Jamuna Press, Lincoln, MA
- Schwartz S, Rokosky J (2007) Slow slip events and seismic tremor at circum-pacific subduction zones. *Rev Geophys* 45:RG3004
- Shearer P (1999) Introduction to seismology. Cambridge University Press, New York
- Strang G, Borre K (1997) Linear algebra, geodesy, and GPS. Wellesley-Cambridge Press, Wellesley, MA

---

## Granular Computing and Data Mining for Ordered Data: The Dominance-Based Rough Set Approach

SALVATORE GRECO<sup>1</sup>, BENEDETTO MATARAZZO<sup>1</sup>,  
ROMAN SŁOWIŃSKI<sup>2,3</sup>

- <sup>1</sup> Faculty of Economics, University of Catania, Catania, Italy  
<sup>2</sup> Poznań University of Technology, Institute of Computing Science, Poznań, Poland  
<sup>3</sup> Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

## Article Outline

- Glossary
- Definition of the Subject
- Introduction: Granular Computing and Ordered Data
- Philosophical Basis of DRSA
- Granular Computing
- Dominance-Based Rough Set Approach
- Fuzzy Set Extensions
  - of the Dominance-Based Rough Set Approach
  - Variable-Consistency Dominance-Based Rough Set Approach (VC-DRSA)
  - Dominance-Based Rough Approximation of a Fuzzy Set
  - Monotonic Rough Approximation
    - of a Fuzzy Set Versus Classical Rough Set
  - Dominance-Based Rough Set Approach
    - to Case-Based Reasoning

## An Algebraic Structure

### for Dominance-Based Rough Set Approach

#### Conclusions

#### Future Directions

#### Bibliography

## Glossary

**Case-based reasoning** Case-based reasoning is a paradigm in machine learning whose idea is that a new problem can be solved by noticing its similarity to a set of problems previously solved. Case-based reasoning regards the inference of some proper conclusions related to a new situation by the analysis of similar cases from a memory of previous cases. Very often similarity between two objects is expressed on a graded scale and this justifies application of fuzzy sets in this context. Fuzzy case-based reasoning is a popular approach in this domain.

**Decision rule** Decision rule is a logical statement of the type “*if... then...*”, where the premise (condition part) specifies values assumed by one or more condition attributes and the conclusion (decision part) specifies an overall judgment.

#### Dominance-based rough set approach (DRSA)

DRSA permits approximation of a set in universe  $U$  based on available ordinal information about objects of  $U$ . Also the decision rules induced within DRSA are based on ordinal properties of the elementary conditions in the premise and in the conclusion, such as “*if property  $f_{i1}$  is present in degree at least  $\alpha_{i1}$  and ... property  $f_{ip}$  is present in degree at least  $\alpha_{ip}$ , then property  $f_{iq}$  is present in degree at least  $\alpha_{iq}$* ”.

**Fuzzy sets** Differently from ordinary sets in which an object belongs or does not belong to a given set, in a fuzzy set an object belongs to a set in some degree. Formally, in universe  $U$  a fuzzy set  $X$  is characterized by its membership function  $\mu_X: U \rightarrow [0, 1]$ , such that for any  $y \in U$ ,  $y$  certainly does not belong to set  $X$  if  $\mu_X(y) = 0$ ,  $y$  certainly belongs to  $X$  if  $\mu_X(y) = 1$ , and  $y$  belongs to  $X$  with a given degree of certainty represented by the value of  $\mu_X(y)$  in all other cases.

**Granular computing** Granular computing is a general computation theory for using granules such as subsets, classes, objects, clusters, and elements of a universe to build an efficient computational model for complex applications with huge amounts of data, information, and knowledge. Granulation of an object  $a$  leads to a collection of granules, with a granule being a clump of points (objects) drawn together by indiscernibility,

similarity, proximity, or functionality. In human reasoning and concept formulation, the granules and the values of their attributes are fuzzy rather than crisp. In this perspective, fuzzy information granulation may be viewed as a mode of generalization, which can be applied to any concept, method, or theory.

**Ordinal properties and monotonicity** Ordinal properties in description of objects are related to graduality of the presence or absence of a property. In this context, it is meaningful to say that a property is more present in one object than in another object. It is important that the ordinal descriptions are handled properly, which means, without introducing any operation, such as sum, averages, or fuzzy operators, like t-norm or t-conorm of Łukasiewicz, taking into account cardinal properties of data not present in the considered descriptions, which would, therefore, give not meaningful results. Monotonicity is strongly related to ordinal properties. It regards relationships between degrees of presence or absence of properties in the objects, like “the more present is property  $f_i$ , the more present is property  $f_j$ ”, or “the more present is property  $f_i$ , the more absent is property  $f_j$ ”. The graded presence or absence of a property can be meaningfully represented using fuzzy sets. More precisely, the degree of presence of property  $f_i$  in object  $y \in U$  is the value given to  $y$  by the membership function of the set of objects having property  $f_i$ .

**Rough set** A rough set in universe  $U$  is an approximation of a set based on available information about objects of  $U$ . The rough approximation is composed of two ordinary sets called *lower and upper approximation*. Lower approximation is a maximal subset of objects which, according to the available information, certainly belong to the approximated set, and upper approximation is a minimal subset of objects which, according to the available information, possibly belong to the approximated set. The difference between upper and lower approximation is called *boundary*.

### Definition of the Subject

This article describes the dominance-based rough set approach (DRSA) to granular computing and data mining. DRSA was first introduced as a generalization of the rough set approach for dealing with multicriteria decision analysis, where preference order is important. The ordering is also important, however, in many other problems of data analysis. Even when the ordering seems absent, the presence or the absence of a property can be represented in ordinal terms, because if two properties are related, the pres-

ence, rather than the absence, of one property should make more (or less) probable the presence of the other property. This is even more apparent when the presence or the absence of a property is graded or fuzzy, because in this case, the more credible the presence of a property, the more (or less) probable the presence of the other property. Since the presence of properties, possibly fuzzy, is the basis of any granulation, DRSA can be seen as a general basis for granular computing.

After presenting the main ideas of DRSA for granular computing and its philosophical basis, the article introduces the basic concepts of DRSA, followed by its extensions in a fuzzy context and in probabilistic terms. This prepares the ground for treating the rough approximation of a fuzzy set, which is the core of the subject. It is also explained why the classical rough set approach is a specific case of DRSA. The article continues with presentation of DRSA for case-based reasoning, where the main ideas of DRSA for granular computing are fruitfully applied. Finally, some basic formal properties of the whole approach are presented in terms of an algebra modeling the logic of DRSA.

### Introduction: Granular Computing and Ordered Data

Granular computing originated in the research of Lin [37, 38, 39, 40, 41, 42] and Zadeh [57, 58, 59, 60] and gained considerable interest in the last decade. The basic components of granular computing are granules, such as subsets, classes, objects, clusters, and elements of a universe. Granulation of an object  $a$  leads to a collection of granules, with a granule being a clump of points (objects) drawn together by indiscernibility, similarity, proximity, or functionality. In human reasoning and concept formulation, the granules and the values of their attributes are fuzzy rather than crisp. In this perspective, fuzzy information granulation may be viewed as a mode of generalization, which can be applied to any concept, method, or theory. Moreover, the theory of fuzzy granulation provides a basis for computing with words, due to the observation that in a natural language, words play the role of labels of fuzzy granules.

Since fuzzy granulation plays a central role in fuzzy logic and in its applications, and rough set theory can be considered as a crisp granulation of set theory, it is interesting to study the relationship between fuzzy sets and rough sets from this point of view. Moreover, noticing that fuzzy granulation that leads to fuzzy logic underlies all applications of granulation, hybridization of fuzzy sets and rough sets can lead to a more general theory of granulation with a potential of application to any domain of human in-

vestigation. This explains the interest in putting together rough sets and fuzzy sets.

Recently, it has been shown that a proper way of handling graduality in rough set theory is to use the DRSA [29]. This implies that DRSA is also a proper way of handling granulation within rough set theory. Let us explain this point in detail.

The rough set approach has been proposed to approximate some relationships existing between concepts. For example, in medical diagnosis the concept of “disease Y” can be represented in terms of such concepts as “low blood pressure” and “high temperature”, or “muscle pain” and “headache”. The classical rough approximation is based on a very coarse representation, that is, for each aspect characterizing a concept (“low blood pressure”, “high temperature”, “muscle pain”, etc.), only its presence or its absence is considered relevant. In this case, the rough approximation involves a very primitive idea of monotonicity related to a scale with only two values: “presence” and “absence”.

Monotonicity gains importance when a finer representation of the concepts is considered. A representation is finer when, for each aspect characterizing a concept, not only its presence or its absence is taken into account, but also the *degree* of its presence or absence is considered relevant. Graduality is typical for fuzzy set philosophy [56] and, therefore, a joint consideration of rough sets and fuzzy sets is worthwhile. In fact, rough sets and fuzzy sets capture the two basic complementary aspects of monotonicity: rough sets deal with relationships between different concepts and fuzzy sets deal with expression of different dimensions in which the concepts are considered. For this reason, many approaches have been proposed to combine fuzzy sets with rough sets (see e.g. [3,5,45,49,51]). Our combination of rough sets and fuzzy sets presents some important advantages with respect to the other approaches, which are discussed below.

The main preoccupation in almost all the studies combining rough sets with fuzzy sets was related to a fuzzy extension of Pawlak’s definition of lower and upper approximations using fuzzy connectives [10,34]. In fact, there is no rule for the choice of the “right” connective, so this choice is always arbitrary to some extent. Another drawback of fuzzy extensions of rough sets involving fuzzy connectives is that they are based on cardinal properties of membership degrees. In consequence, the result of these extensions is sensitive to order preserving transformation of membership degrees. For example, consider the t-conorm of Łukasiewicz as a fuzzy connective; it may be used in the definition of both fuzzy lower approximation (to build fuzzy implication) and fuzzy upper approximation (as a fuzzy counterpart of a union). The t-conorm of

Łukasiewicz is defined as

$$T^*(\alpha, \beta) = \min(\alpha + \beta, 1), \quad \alpha, \beta \in [0, 1].$$

$T^*(\alpha, \beta)$  can be interpreted as follows. Given two fuzzy propositions  $p$  and  $q$ , putting  $v(p) = \alpha$  and  $v(q) = \beta$ ,  $T^*(\alpha, \beta)$  can be interpreted as  $v(p \vee q)$ , the truth value of the proposition  $p \vee q$ . Let us consider the following values of arguments:

$$\alpha = 0.5, \quad \beta = 0.3, \quad \gamma = 0.2, \quad \delta = 0.1,$$

and their order preserving transformation:

$$\alpha' = 0.4, \quad \beta' = 0.3, \quad \gamma' = 0.2, \quad \delta' = 0.05.$$

The values of the t-conorm are in the two cases as follows:

$$\begin{aligned} T^*(\alpha, \delta) &= 0.6, & T^*(\beta, \gamma) &= 0.5, \\ T^*(\alpha', \delta') &= 0.45, & T^*(\beta', \gamma') &= 0.5. \end{aligned}$$

One can see that the order of the results has changed after the order preserving transformation of the arguments. This means that the Łukasiewicz t-conorm takes into account not only the ordinal properties of the truth values, but also their cardinal properties. A natural question arises: is it reasonable to expect from truth values a cardinal content instead of ordinal only? Or, in other words, is it realistic to claim that a human is able to say in a meaningful way not only that

- (a) “proposition  $p$  is more credible than proposition  $q$ ” but even something like
- (b) “proposition  $p$  is two times more credible than proposition  $q$ ”?

It is much safer to consider information of type (a), because information of type (b) is rather meaningless for a human.

Since fuzzy generalization of rough set theory using DRSA takes into account only *ordinal properties* of fuzzy membership degrees, it is the proper way of fuzzy generalization of rough set theory.

Moreover, the classical rough set approach [46,47] can be seen as a specific case of our general model. This is important for several reasons. In particular, this interpretation of DRSA gives an insight into fundamental properties of the classical rough set approach and permits its further generalization.

Rough set theory [46,47] relies on the idea that some knowledge (data, information) is available about objects of a universe of discourse  $U$ . Thus, a subset of  $U$  is defined using the available knowledge about the objects and not on the base of information about membership or non-mem-

bership of the objects to the subset. For example, knowledge about patients suffering from a certain disease may contain information about body temperature, blood pressure, etc. All patients described by the same information are *indiscernible* in view of the available knowledge, and form groups of similar objects. These groups are called *elementary sets*, and can be considered as basic granules of the available knowledge about patients. Elementary sets can be combined into compound concepts. For example, elementary sets of patients can be used to represent a set of patients suffering from a certain disease. Any union of elementary sets is called a crisp set, while other sets are referred to as *rough sets*. Each rough set has boundary line objects, i. e. objects which, in view of the available knowledge, cannot be classified with certainty as members of the set or of its complement. Therefore, in the rough set approach, any set is associated with a pair of crisp sets, called the lower and the upper approximation. Intuitively, in view of the available information, the *lower approximation* consists of all objects which certainly belong to the set and the *upper approximation* contains all objects which possibly belong to the set. The difference between the upper and the lower approximation constitutes the *boundary* region of the rough set. Analogously, for a partition of universe  $U$  into classes, one may consider rough approximation of the partition. It appeared to be particularly useful for analysis of classification problems, being the most common decision problems.

The rough set approach operates on an *information table* composed of a set  $U$  of objects described by a set  $Q$  of attributes. If in the set  $Q$  disjoint sets ( $C$  and  $D$ ) of *condition* and *decision* attributes are distinguished, then the information table is called a *decision table*. It is often assumed, without loss of generality, that set  $D$  is a singleton  $\{d\}$ , and thus decision attribute  $d$  makes a partition of set  $U$  into decision classes corresponding to its values. Data collected in such a decision table correspond to a multiple attribute classification problem. The classical Indiscernibility-Based Rough Set Approach (IRSA) is naturally adapted to analysis of this type of decision problems, because the set of objects can be identified with examples of classification and it is possible to extract all the essential knowledge contained in the decision table using indiscernibility or similarity relations. However, as pointed out by the authors (see e. g. [17,20,26,54]), IRSA cannot extract all the essential knowledge contained in the decision table if a background knowledge about *monotonic relationships* between evaluation of objects on condition attributes and their assignment to decision classes has to be taken into account. Such a background knowledge is typical for data describing various phenomena, as well as for data describing

multiple criteria decision problems (see e. g. [9]), e. g., “the larger the mass and the smaller the distance, the larger the gravity”, “the more a tomato is red, the more it is ripe” or “the better the school marks of a pupil, the better his overall classification”. The monotonic relationships, typical for multiple criteria decision problems, follow from preferential ordering of value sets of attributes (scales of criteria), as well as preferential ordering of decision classes.

In order to take into account the *ordinal* properties of the considered attributes and the monotonic relationships between condition and decision attributes, a number of methodological changes to the original rough set theory were necessary. The main change was the replacement of the indiscernibility relation with a *dominance relation*, which permits approximation of ordered sets. The dominance relation is a very natural and rational concept within multiple criteria decision analysis. The dominance-based rough set approach (DRSA) has been proposed and characterized by the authors (see e. g. [17,20,24,25,26,54]).

Let us mention that ordered value sets of attributes and a kind of order dependency among attributes has also been considered by specialists of relational databases (see, e. g., [13]). There is, however, a striking difference between consideration of orders in database queries and consideration of orders in knowledge discovery. More precisely, assuming ordered domains of attributes, knowledge discovery tends to discover monotonic relationships between ordered attributes, e. g., if a student is at least medium in networks, and at least good in databases, then his overall evaluation is at least medium. On the other hand, assuming an order of attribute value sets and order dependency in databases, one can exploit this given information for a more efficient answer to a query, e. g., when the dates of bank checks and their numbers are ordered, there is an order dependency between these two attributes, because in day  $x$  a check cannot hold a number smaller than checks from day  $x - 1$ , which permits us to prune the search tree and make the search more efficient.

Looking at DRSA from a granular computing perspective, one can observe that DRSA permits us to deal with *ordered data* by considering a specific type of information granules defined by means of *dominance-based constraints* having a syntax of the type: “ $x$  is at least  $R$ ” or “ $x$  is at most  $R$ ”, where  $R$  is a qualifier from a properly ordered scale. In evaluation space, such granules are *dominance cones*. In this sense, the contribution of DRSA consists of:

- Extending the paradigm of granular computing to problems involving ordered data,
- Specifying a proper syntax and modality of information granules (the dominance-based constraints which

should be adjoined to other modalities of information constraints, such as possibilistic, veristic, and probabilistic [60]),

- Defining a methodology dealing properly with this type of information granules, and resulting in a theory of computing with words and reasoning about data in the case of ordered data.

Let us observe that other modalities of information constraints, such as veristic, possibilistic, and probabilistic, have also to deal with ordered values (with qualifiers relative to grades of truth, possibility, and probability). Therefore, granular computing with ordered data and DRSA as a proper way of reasoning about ordered data, are very important in the future development of the whole domain of granular computing.

Indeed the DRSA approach proposed in [15,16] avoids arbitrary choice of fuzzy connectives and not meaningful operations on membership degrees. It exploits only ordinal character of the membership degrees and proposes a methodology of fuzzy rough approximation that infers the most cautious conclusion from available imprecise information. In particular, any approximation of knowledge about concept  $Y$  using knowledge about concept  $X$  is based on positive or negative relationships between premises and conclusions, i. e.:

- "The more  $x$  is  $X$ , the more it is  $Y$ " (positive relationship),
- "The more  $x$  is  $X$ , the less it is  $Y$ " (negative relationship).

The following simple relationships illustrate i) and ii):

- "The larger the market share of a company, the greater its profit" (positive relationship), and
- "The greater the debt of a company, the smaller its profit" (negative relationship).

These relationships have the form of *gradual decision rules* [4]. Examples of these decision rules are: "if a car is speedy with credibility at least 0.8 and it has high fuel consumption with credibility at most 0.7, then it is a good car with a credibility at least 0.9", and "if a car is speedy with credibility at most 0.5 and it has high fuel consumption with credibility at least 0.8, then it is a good car with a credibility at most 0.6". It is worth noting that the syntax of gradual decision rules is based on monotonic relationships between degrees of credibility, that can also be found in dominance-based decision rules induced from preference-ordered data. This explains why one can build a fuzzy rough approximation using DRSA.

Finally, the fuzzy rough approximation taking into account monotonic relationships can be applied to case-based reasoning [28]. In this perspective, it is interesting to consider monotonicity of the type "the more similar is  $y$  to  $x$ , the more credible is that  $y$  belongs to the same class as  $x$ ". Application of DRSA in this context leads to decision rules similar to the gradual decision rules:

"the more object  $z$  is similar to a referent object  $x$  w.r.t. condition attribute  $s$ , the more  $z$  is similar to a referent object  $x$  w.r.t. decision attribute  $t$ ",

or, equivalently, but more technically,

$$s(z, x) \geq \alpha \Rightarrow t(z, x) \geq \alpha ,$$

where functions  $s$  and  $t$  measure the credibility of similarity with respect to condition attribute and decision attribute, respectively. When there are multiple condition and decision attributes, functions  $s$  and  $t$  aggregate similarity with respect to these attributes.

The decision rules induced on the basis of DRSA do not need the aggregation of the similarity with respect to different attributes into one comprehensive similarity. This is important because it permits us to avoid using aggregation operators (weighted average, min, etc.) which are always arbitrary to some extent. Moreover, the DRSA decision rules permit us to consider different thresholds for degrees of credibility in the premise and in the conclusion.

The article is organized as follows. Next Section presents the philosophical basis of DRSA granular computing. Section "[Dominance-Based Rough Set Approach](#)" recalls the main steps of DRSA for multiple criteria classification, or, in general, for ordinal classification. Section "[Fuzzy Set Extensions of the Dominance-Based Rough Set Approach](#)" presents, a review of fuzzy set extensions of DRSA based on fuzzy connectives. In Sect. "[Variable-Consistency Dominance-Based Rough Set Approach \(VC-DRSA\)](#)", a "probabilistic" version of DRSA, the variable-consistency dominance-based rough set approach is presented. Dominance-based rough approximation of a fuzzy set is presented in Sect "[Dominance-Based Rough Approximation of a Fuzzy Set](#)". The explanation that IRSA is a particular case of DRSA is presented in Sect. "[Monotonic Rough Approximation of a Fuzzy Set versus Classical Rough Set](#)". Section "[Dominance-Based Rough Set Approach to Case-Based Reasoning](#)" is devoted to DRSA for case-based reasoning. In Sect. "[An Algebraic Structure for Dominance-Based Rough Set Approach](#)" an algebra modeling logic of DRSA is presented. Section "[Conclusions](#)" contains conclusions. In Sect. "[Future Directions](#)" some issues for further developments are presented.

## Philosophical Basis of DRSA Granular Computing

It is interesting to analyze the relationships between DRSA and granular computing from the point of view of the philosophical basis of rough set theory proposed by Pawlak. Since according to Pawlak [48], rough set theory refers to some ideas of Gottlob Frege (vague concepts), Gottfried Leibniz (indiscernibility), George Boole (reasoning methods), Jan Łukasiewicz (multi-valued logic), and Thomas Bayes (inductive reasoning), it is meaningful to give an account for DRSA generalization of rough sets, justifying it in reference to some of these main ideas recalled by Pawlak.

The *identity of indiscernibles* is a principle of analytic ontology first explicitly formulated by Gottfried Leibniz in his Discourse on Metaphysics, Sect. 9 [43]. Two objects  $x$  and  $y$  are defined indiscernible, if  $x$  and  $y$  have the same properties. The principle of identity of indiscernibles states that

$$\text{if } x \text{ and } y \text{ are indiscernible, then } x = y. \quad (\text{II1})$$

This can be expressed also as *if  $x \neq y$ , then  $x$  and  $y$  are discernible*, i. e. there is at least one property that  $x$  has and  $y$  does not, or vice versa. The converse of the principle of identity of indiscernibles is called *indiscernibility of identicals* and states that *if  $x = y$ , then  $x$  and  $y$  are indiscernible*, i. e. they have the same properties. This is equivalent to saying that if there is at least one property that  $x$  has and  $y$  does not, or vice versa, then  $x \neq y$ . The conjunction of both principles is often referred to as “Leibniz’s law”.

Rough set theory is based on a weaker interpretation of Leibniz’s law, having as objective the ability to classify objects falling under the same concept. This reinterpretation of Leibniz’s law is based on a reformulation of the principle of identity of indiscernibles as follows:

$$\begin{aligned} &\text{if } x \text{ and } y \text{ are indiscernible,} \\ &\text{then } x \text{ and } y \text{ belong to the same class.} \end{aligned} \quad (\text{II2})$$

Let us observe that the word “class” in the previous sentence can be considered as synonymous with “granule”. Thus, from the point of view of granular computing, (II2) can be rewritten as

$$\begin{aligned} &\text{if } x \text{ and } y \text{ are indiscernible,} \\ &\text{then } x \text{ and } y \text{ belong to the same granule of classification.} \end{aligned} \quad (\text{II2}')$$

Notice also that the principle of indiscernibility of identicals cannot be reformulated in analogous terms. In fact, such an analogous reformulation would amount to stating that *if  $x$  and  $y$  belong to the same class, then  $x$*

*and  $y$  are indiscernible*. This principle is too strict, however, because there can be two discernible objects  $x$  and  $y$  belonging to the same class. Thus, within rough set theory, the principle of indiscernibility of identicals should continue to hold in its original formulation (i. e. *if  $x = y$ , then  $x$  and  $y$  are indiscernible*). It is worthwhile to observe that the relaxation in the consequence of the implication from (II1) to (II2), implies an implicit relaxation also in the antecedent. In fact, one could say that two objects are identicals if they have the same properties, if one would be able to take into account all conceivable properties. For human limitations this is not the case, therefore, one can imagine that (II2) can be properly reformulated as

$$\begin{aligned} &\text{if } x \text{ and } y \text{ are indiscernible taking into account} \\ &\text{a given set of properties,} \\ &\text{then } x \text{ and } y \text{ belong to the same class.} \end{aligned} \quad (\text{II2}'')$$

This weakening in the antecedent of the implication means also that the objects indiscernible with respect to a given set of properties can be seen as a granule, such that, finally, the (II2) could be rewritten in terms of granulation as

$$\begin{aligned} &\text{if } x \text{ and } y \text{ belong to the same granule with respect} \\ &\text{to a given set of properties,} \\ &\text{then } x \text{ and } y \text{ belong to the same classification granule.} \end{aligned} \quad (\text{II2}''')$$

For this reason, rough set theory needs a still weaker form of the principle of identity of indiscernibles. Such a principle can be formulated using the idea of vagueness due to Gottlob Frege. According to Frege “the concept must have a sharp boundary – to the concept without a sharp boundary there would correspond an area that had not a sharp boundary-line all around”. Therefore, following this intuition, the principle of identity of indiscernibles can be further reformulated as

$$\begin{aligned} &\text{if } x \text{ and } y \text{ are indiscernible,} \\ &\text{then } x \text{ and } y \text{ should belong to the same class.} \end{aligned} \quad (\text{II3})$$

In terms of granular computing, (II3) can be rewritten as

$$\begin{aligned} &\text{if } x \text{ and } y \text{ belong to the same granule with respect} \\ &\text{to a given set of properties,} \\ &\text{then } x \text{ and } y \text{ should belong to the same classification} \\ &\text{granule.} \end{aligned} \quad (\text{II3}')$$

This reformulation of the principle of identity of indiscernibles implies that there is an inconsistency in the statement that  $x$  and  $y$  are indiscernible, and  $x$  and  $y$  be-

long to different classes. Thus, Leibniz's principle of identity of indiscernibles and Frege's intuition about vagueness found the basic idea of the rough set concept proposed by Pawlak.

The above reconstruction of the basic idea of Pawlak's rough set should be completed, however, by referring to another basic idea. This is the idea of Georg Boole that concerns a property which is satisfied or not satisfied. It is quite natural to weaken this principle admitting that a property can be satisfied to some degree. This idea of graduality can be attributed to Jan Łukasiewicz and his proposal of many-valued logic where, in addition to well-known truth values "true" and "false", other truth values representing partial degrees of truth were present. Łukasiewicz's idea of graduality has been reconsidered, generalized and fully exploited by Zadeh [56] within fuzzy set theory, where graduality concerns membership to a set.

In this sense, any proposal of putting rough sets and fuzzy sets together can be seen as a reconstruction of the rough set concept, where Boole's idea of binary logic is abandoned in favor of Łukasiewicz's idea of many-valued logic, such that Leibniz's principle of identity of indiscernibles and Frege's intuition about vagueness are combined with the idea that a property is satisfied to some degree.

Putting aside, for the moment, Frege's intuition about vagueness, but taking into account the concept of graduality, the principle of identity of indiscernibles can reformulated as follows:

*if the grade of each property for  $x$  is greater than  
or equal to the grade for  $y$ ,  
then  $x$  belongs to the considered class in a grade  
at least as high as  $y$ .* (II4)

Taking into account the paradigm of granular computing, (II4) can be rewritten as

*if  $x$  belongs to the granules defined by considered  
properties more than  $y$ , because the grade of each  
property for  $x$  is greater than or equal to the  
grade for  $y$ ,  
then  $x$  belongs to the considered classification granule  
in a grade at least as high as  $y$ .* (II4')

Considering the concept of graduality together with Frege's intuition about vagueness, one can reformulate the principle of identity of indiscernibles as follows:

*if the grade of each property for  $x$  is greater than  
or equal to the grade for  $y$ ,  
then  $x$  should belong to the considered class  
in a grade at least as high as  $y$ .* (II5)

In terms of granular computing, (II5) can be rewritten as

*if  $x$  belongs to the granules defined by considered  
properties more than  $y$ , because the grade of each  
property for  $x$  is greater than or  
equal to the grade for  $y$ ,  
then  $x$  should belong to the considered classification  
granule in a grade at least as high as  $y$ .* (II5')

The formulation (II5') of the principle of identity of indiscernibles is perfectly concordant with the rough set concept defined within the dominance-based rough set approach [20].

DRSA has been proposed by the authors to deal with ordinal properties of data related to preferences in decision problems [26,54]. The fundamental feature of DRSA is that it handles monotonicity of comprehensive evaluation of objects with respect to preferences relative to evaluation of these objects on particular attributes. For example, the more preferred is a car with respect to such attributes as maximum speed, acceleration, fuel consumption, and price, the better is its comprehensive evaluation. The type of monotonicity considered within DRSA is also meaningful for problems where relationships between different aspects of a phenomenon described by data are to be taken into account, even if preferences are not considered. Indeed, monotonicity concerns, in general, mutual trends existing between different variables, like distance and gravity in physics, or inflation rate and interest rate in economics. Whenever a relationship between different aspects of a phenomenon is discovered, this relationship can be represented by a monotonicity with respect to some specific measures of the considered aspects. Formulation (II5) of the principle of identity of indiscernibles refers to this type of monotonic relationships. So, in general, the monotonicity permits us to translate into a formal language a primitive intuition of relationships between different concepts of our knowledge corresponding to the principle of identity of indiscernibles formulated as (II5').

### Dominance-Based Rough Set Approach

This section presents the main concepts of the dominance-based rough set approach (for a more complete presentation see, for example, [17,20,26,54]).

Information about objects is represented in the form of an information table. The rows of the table are labeled by objects, whereas columns are labeled by attributes and entries of the table are attribute-values. Formally, an information system (table) is the 4-tuple  $S = \langle U, Q, V, \phi \rangle$ , where  $U$  is a finite set of objects,  $Q$  is a finite set of attributes,  $V = \bigcup_{q \in Q} V_q$  and  $V_q$  is the set of values of the

attribute  $q$ , and  $\phi: U \times Q \rightarrow V_q$  is a total function such that  $\phi(x, q) \in V_q$  for every  $q \in Q$ ,  $x \in U$ , called an information function [47]. The set  $Q$  is, in general, divided into set  $C$  of condition attributes and set  $D$  of decision attributes.

Condition attributes with value sets ordered according to decreasing or increasing preference are called *criteria*. For criterion  $q \in Q$ ,  $\succeq_q$  is a *weak preference relation* on  $U$  such that  $x \succeq_q y$  means “ $x$  is at least as good as  $y$  with respect to criterion  $q$ ”. It is supposed that  $\succeq_q$  is a complete preorder, i. e. a strongly complete and transitive binary relation, defined on  $U$  on the basis of evaluations  $\phi(\cdot, q)$ . Without loss of generality, the preference is supposed to increase with the value of  $\phi(\cdot, q)$  for every criterion  $q \in C$ , such that for all  $x, y \in U$ ,  $x \succeq_q y$  if and only if  $\phi(x, q) \geq \phi(y, q)$ .

Furthermore, it is supposed that the set of decision attributes  $D$  is a singleton  $d$ . Values of decision attribute  $d$  makes a partition of  $U$  into a finite number of decision classes,  $Cl = \{Cl_t, t = 1, \dots, n\}$ , such that each  $x \in U$  belongs to one and only one class  $Cl_t \in Cl$ . It is supposed that the classes are preference-ordered, i. e. for all  $r, s \in \{1, \dots, n\}$ , such that  $r > s$ , the objects from  $Cl_r$  are preferred to the objects from  $Cl_s$ . More formally, if  $\succeq$  is a *comprehensive weak preference relation* on  $U$ , i. e. if for all  $x, y \in U$ ,  $x \succeq y$  means “ $x$  is at least as good as  $y$ ”, it is supposed:  $[x \in Cl_r, y \in Cl_s, r > s] \Rightarrow [x \succeq y \text{ and } \text{not } y \succeq x]$ . The above assumptions are typical for consideration of *ordinal classification problems* (also called *multiple criteria sorting problems*).

The sets to be approximated are called *upward union* and *downward union* of classes, respectively:

$$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s, \quad Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s, \quad t = 1, \dots, n.$$

The statement  $x \in Cl_t^{\geq}$  means “ $x$  belongs to at least class  $Cl_t$ ”, while  $x \in Cl_t^{\leq}$  means “ $x$  belongs to at most class  $Cl_t$ ”. Let us remark that  $Cl_1^{\geq} = Cl_n^{\leq} = U$ ,  $Cl_n^{\geq} = Cl_n$  and  $Cl_1^{\leq} = Cl_1$ . Furthermore, for  $t = 2, \dots, n$ ,

$$Cl_{t-1}^{\leq} = U - Cl_t^{\geq} \quad \text{and} \quad Cl_t^{\geq} = U - Cl_{t-1}^{\leq}.$$

The key idea of the rough set approach is representation (approximation) of knowledge generated by decision attributes, by “*granules of knowledge*” generated by condition attributes.

In DRSA, where condition attributes are criteria and decision classes are preference ordered, the represented knowledge is a collection of *upward* and *downward unions of classes* and the “*granules of knowledge*” are sets of objects defined using a dominance relation.

$x$  dominates  $y$  with respect to  $P \subseteq C$  (shortly,  $x$  *P-dominates*  $y$ ), denoted by  $xD_Py$ , if for every criterion  $q \in P$ ,  $\phi(x, q) \geq \phi(y, q)$ . The relation of *P-dominance* is reflexive and transitive, that is it is a partial preorder.

Given a set of criteria  $P \subseteq C$  and  $x \in U$ , the “*granules of knowledge*” used for approximation in DRSA are:

- A set of objects dominating  $x$ , called *P-dominating set*,  $D_P^+(x) = \{y \in U : yD_Px\}$ ,
- A set of objects dominated by  $x$ , called *P-dominated set*,  $D_P^-(x) = \{y \in U : xD_Py\}$ .

Note that the “*granules of knowledge*” defined above have the form of upward (positive) and downward (negative) *dominance cones* in the evaluation space.

Let us recall that the *dominance principle* (or Pareto principle) requires that an object  $x$  dominating object  $y$  on all considered criteria (i. e.  $x$  having evaluations at least as good as  $y$  on all considered criteria) should also dominate  $y$  on the decision (i. e.  $x$  should be assigned to at least as good a decision class as  $y$ ). This principle is the only objective principle that is widely agreed upon in the multiple criteria comparisons of objects.

Given  $P \subseteq C$ , the inclusion of an object  $x \in U$  to the upward union of classes  $Cl_t^{\geq}$ ,  $t = 2, \dots, n$ , is *inconsistent with the dominance principle* if one of the following conditions holds:

- $x$  belongs to class  $Cl_t$  or better but it is *P-dominated* by an object  $y$  belonging to a class worse than  $Cl_t$ , i. e.  $x \in Cl_t^{\geq}$  but  $D_P^+(x) \cap Cl_{t-1}^{\leq} \neq \emptyset$ ,
- $x$  belongs to a worse class than  $Cl_t$  but it *P-dominates* an object  $y$  belonging to class  $Cl_t$  or better, i. e.  $x \notin Cl_t^{\geq}$  but  $D_P^-(x) \cap Cl_t^{\geq} \neq \emptyset$ .

If, given a set of criteria  $P \subseteq C$ , the inclusion of  $x \in U$  to  $Cl_t^{\geq}$ , where  $t = 2, \dots, n$ , is inconsistent with the dominance principle, then  $x$  belongs to  $Cl_t^{\geq}$  with some ambiguity. Thus,  $x$  belongs to  $Cl_t^{\geq}$  without any ambiguity with respect to  $P \subseteq C$ , if  $x \in Cl_t^{\geq}$  and there is no inconsistency with the dominance principle. This means that all objects *P-dominating*  $x$  belong to  $Cl_t^{\geq}$ , i. e.  $D_P^+(x) \subseteq Cl_t^{\geq}$ .

Furthermore,  $x$  possibly belongs to  $Cl_t^{\geq}$  with respect to  $P \subseteq C$  if one of the following conditions holds:

- According to decision attribute  $d$ ,  $x$  belongs to  $Cl_t^{\geq}$ ,
- According to decision attribute  $d$ ,  $x$  does not belong to  $Cl_t^{\geq}$ , but it is inconsistent in the sense of the dominance principle with an object  $y$  belonging to  $Cl_t^{\geq}$ .

In terms of ambiguity,  $x$  possibly belongs to  $Cl_t^{\geq}$  with respect to  $P \subseteq C$ , if  $x$  belongs to  $Cl_t^{\geq}$  with or without any ambiguity. Because of the reflexivity of the dominance relation  $D_P$ , the above conditions can be summarized as

follows:  $x$  possibly belongs to class  $Cl_t$  or better, with respect to  $P \subseteq C$ , if among the objects  $P$ -dominated by  $x$  there is an object  $y$  belonging to class  $Cl_t$  or better, i.e.  $D_P^-(x) \cap Cl_t^{\geq} \neq \emptyset$ .

The  $P$ -lower approximation of  $Cl_t^{\geq}$ , denoted by  $\underline{P}Cl_t^{\geq}$ , and the  $P$ -upper approximation of  $Cl_t^{\geq}$ , denoted by  $\overline{P}(Cl_t^{\geq})$ , are defined as follows ( $t = 1, \dots, n$ ):

$$\begin{aligned}\underline{P}(Cl_t^{\geq}) &= \{x \in U : D_P^+(x) \subseteq Cl_t^{\geq}\}, \\ \overline{P}(Cl_t^{\geq}) &= \{x \in U : D_P^-(x) \cap Cl_t^{\geq} \neq \emptyset\}.\end{aligned}$$

Analogously, one can define the  $P$ -lower approximation and the  $P$ -upper approximation of  $Cl_t^{\leq}$  as follows ( $t = 1, \dots, n$ ):

$$\begin{aligned}\underline{P}(Cl_t^{\leq}) &= \{x \in U : D_P^-(x) \subseteq Cl_t^{\leq}\}, \\ \overline{P}(Cl_t^{\leq}) &= \{x \in U : D_P^+(x) \cap Cl_t^{\leq} \neq \emptyset\}.\end{aligned}$$

The  $P$ -lower and  $P$ -upper approximations so defined satisfy the following *inclusion properties* for each  $t \in \{1, \dots, n\}$  and for all  $P \subseteq C$ :

$$\begin{aligned}\underline{P}(Cl_t^{\geq}) &\subseteq Cl_t^{\geq} \subseteq \overline{P}(Cl_t^{\geq}), \\ \underline{P}(Cl_t^{\leq}) &\subseteq Cl_t^{\leq} \subseteq \overline{P}(Cl_t^{\leq}).\end{aligned}$$

The  $P$ -lower and  $P$ -upper approximations of  $Cl_t^{\geq}$  and  $Cl_t^{\leq}$  have an important *complementarity property*, according to which,

$$\begin{aligned}\underline{P}(Cl_t^{\geq}) &= U - \overline{P}(Cl_{t-1}^{\leq}) \quad \text{and} \\ \overline{P}(Cl_t^{\geq}) &= U - \underline{P}(Cl_{t-1}^{\leq}), \quad t = 2, \dots, n, \\ \underline{P}(Cl_t^{\leq}) &= U - \overline{P}(Cl_{t+1}^{\geq}) \quad \text{and} \\ \overline{P}(Cl_t^{\leq}) &= U - \underline{P}(Cl_{t+1}^{\geq}), \quad t = 1, \dots, n-1.\end{aligned}$$

The  $P$ -boundary of  $Cl_t^{\geq}$  and  $Cl_t^{\leq}$ , denoted by  $Bn_P(Cl_t^{\geq})$  and  $Bn_P(Cl_t^{\leq})$  respectively, are defined as follows ( $t = 1, \dots, n$ ):

$$\begin{aligned}Bn_P(Cl_t^{\geq}) &= \overline{P}(Cl_t^{\geq}) - \underline{P}(Cl_t^{\geq}), \\ Bn_P(Cl_t^{\leq}) &= \overline{P}(Cl_t^{\leq}) - \underline{P}(Cl_t^{\leq}).\end{aligned}$$

Because of the complementarity property,  $Bn_P(Cl_t^{\geq}) = Bn_P(Cl_{t-1}^{\leq})$ , for  $t = 2, \dots, n$ .

The dominance-based rough approximations of upward and downward unions of classes can serve to induce “if... then...” decision rules. It is meaningful to consider the following five types of decision rules:

- 1) Certain  $D_{\geq}$ -decision rules: if  $x_{q1} \succeq_{q1} r_{q1}$  and  $x_{q2} \succeq_{q2} r_{q2}$  and ...  $x_{qp} \succeq_{qp} r_{qp}$ , then certainly  $x$  belongs to  $Cl_t^{\geq}$ , where, for each  $w_q, z_q \in X_q$ , “ $w_q \succeq_q z_q$ ” means “ $w_q$  is at least as good as  $z_q$ ”.

- 2) Possible  $D_{\geq}$ -decision rules: if  $x_{q1} \succeq_{q1} r_{q1}$  and  $x_{q2} \succeq_{q2} r_{q2}$  and ...  $x_{qp} \succeq_{qp} r_{qp}$ , then  $x$  possibly belongs to  $Cl_t^{\geq}$ .
- 3) Certain  $D_{\leq}$ -decision rules: if  $x_{q1} \preceq_{q1} r_{q1}$  and  $x_{q2} \preceq_{q2} r_{q2}$  and ...  $x_{qp} \preceq_{qp} r_{qp}$ , then certainly  $x$  belongs to  $Cl_t^{\leq}$ , where, for each  $w_q, z_q \in X_q$ , “ $w_q \preceq_q z_q$ ” means “ $w_q$  is at most as good as  $z_q$ ”.
- 4) Possible  $D_{\leq}$ -decision rules: if  $x_{q1} \preceq_{q1} r_{q1}$  and  $x_{q2} \preceq_{q2} r_{q2}$  and ...  $x_{qp} \preceq_{qp} r_{qp}$ , then  $x$  possibly belongs to  $Cl_t^{\leq}$ .
- 5) Approximate  $D_{\geq\leq}$ -decision rules: if  $x_{q1} \succeq_{q1} r_{q1}$  and ...  $x_{qk} \succeq_{qk} r_{qk}$  and  $x_{q(k+1)} \preceq_{q(k+1)} r_{q(k+1)}$  and ...  $x_{qp} \preceq_{qp} r_{qp}$ , then  $x \in Cl_s^{\geq} \cap Cl_t^{\leq}$ , where  $s < t$ .

The rules of type 1) and 3) represent certain knowledge extracted from the decision table, while the rules of type 2) and 4) represent possible knowledge. Rules of type 5) represent doubtful knowledge.

## Fuzzy Set Extensions of the Dominance-Based Rough Set Approach

The concept of dominance can be refined by introducing gradedness through the use of fuzzy sets. Here are basic definitions of fuzzy connectives [10,34]. For each proposition  $p$ , one can consider its truth value  $v(p)$  ranging from  $v(p) = 0$  ( $p$  is definitely false) to  $v(p) = 1$  ( $p$  is definitely true); and for all intermediate values, the greater  $v(p)$ , the more credible is the truth of  $p$ . A negation is a non-increasing function  $N: [0, 1] \rightarrow [0, 1]$  such that  $N(0) = 1$  and  $N(1) = 0$ . Given proposition  $p$ ,  $N(v(p))$  states the credibility of the negation of  $p$ . A t-norm  $T$  and a t-conorm  $T^*$  are two functions  $T: [0, 1] \times [0, 1] \rightarrow [0, 1]$  and  $T^*: [0, 1] \times [0, 1] \rightarrow [0, 1]$ , such that given two propositions,  $p$  and  $q$ ,  $T(v(p), v(q))$  represents the credibility of the conjunction of  $p$  and  $q$ , and  $T^*(v(p), v(q))$  represents the credibility of the disjunction of  $p$  and  $q$ . t-norm  $T$  and t-conorm  $T^*$  must satisfy the following properties:

$$T(\alpha, \beta) = T(\beta, \alpha) \quad \text{and} \quad T^*(\alpha, \beta) = T^*(\beta, \alpha), \\ \text{for all } \alpha, \beta \in [0, 1],$$

$$T(\alpha, \beta) \leq T(\gamma, \delta) \quad \text{and} \quad T^*(\alpha, \beta) \leq T^*(\gamma, \delta), \\ \text{for all } \alpha, \beta, \gamma, \delta \in [0, 1], \\ \text{such that } \alpha \leq \gamma \quad \text{and} \quad \beta \leq \delta,$$

$$\begin{aligned}T(\alpha, T(\beta, \gamma)) &= T(T(\alpha, \beta), \gamma) \quad \text{and} \\ T^*(\alpha, T^*(\beta, \gamma)) &= T^*(T^*(\alpha, \beta), \gamma), \\ \text{for all } \alpha, \beta, \gamma \in [0, 1], \\ T(1, \alpha) &= \alpha \quad \text{and} \quad T^*(0, \alpha) = \alpha, \\ \text{for all } \alpha \in [0, 1].\end{aligned}$$

A negation is strict iff it is strictly decreasing and continuous. A negation  $N$  is involutive iff, for all  $\alpha \in [0, 1]$ ,  $N(N(\alpha)) = \alpha$ . A strong negation is an involutive strict negation. If  $N$  is a strong negation, then  $(T, T^*, N)$  is a de Morgan triplet iff  $N(T^*(\alpha, \beta)) = T(N(\alpha), N(\beta))$ . A fuzzy implication is a function  $I: [0, 1] \times [0, 1] \rightarrow [0, 1]$  such that, given two propositions  $p$  and  $q$ ,  $I(v(p), v(q))$  represents the credibility of the implication of  $q$  by  $p$ . A fuzzy implication must satisfy the following properties (see [10]):

$$I(\alpha, \beta) \geq I(\gamma, \beta) \quad \text{for all } \alpha, \beta, \gamma \in [0, 1],$$

such that  $\alpha \leq \gamma$ ,

$$I(\alpha, \beta) \geq I(\alpha, \gamma) \quad \text{for all } \alpha, \beta, \gamma \in [0, 1],$$

such that  $\beta \geq \gamma$ ,

$$I(0, \alpha) = 1, I(\alpha, 1) = 1 \quad \text{for all } \alpha \in [0, 1],$$

$$I(1, 0) = 0.$$

An implication  $I_{N, T^*}^\rightarrow$  is a  $T^*$ -implication if there is a t-conorm  $T^*$  and a strong negation  $N$  such that  $I_{N, T^*}^\rightarrow(\alpha, \beta) = T^*(N(\alpha), \beta)$ . A fuzzy similarity relation on the universe  $U$  is a fuzzy binary relation (i.e. function  $R: U \times U \rightarrow [0, 1]$ ) reflexive ( $R(x, x) = 1$  for all  $x \in U$ ), symmetric ( $R(x, y) = R(y, x)$  for all  $x, y \in U$ ) and transitive (given t-norm  $T$ ,  $T(R(x, y), R(y, z)) \leq R(x, z)$  for all  $x, y, z \in U$ ).

Let  $\succeq_q$  be a fuzzy weak preference relation on  $U$  with respect to criterion  $q \in C$ , i.e.  $\succeq_q: U \times U \rightarrow [0, 1]$ , such that, for all  $x, y \in U$ ,  $\succeq_q(x, y)$  represents the credibility of the proposition “ $x$  is at least as good as  $y$  with respect to criterion  $q$ ”. Suppose that  $\succeq_q$  is a fuzzy partial  $T$ -preorder, i.e. that it is reflexive ( $\succeq_q(x, x) = 1$  for each  $x \in U$ ) and  $T$ -transitive ( $T(\succeq_q(x, y), \succeq_q(y, z)) \leq \succeq_q(x, z)$ , for each  $x, y, z \in U$ ) (see [10]). Using the fuzzy weak preference relations  $\succeq_q$ ,  $q \in C$ , a fuzzy dominance relation on  $U$  (denotation  $D_P(x, y)$ ) can be defined, for all  $P \subseteq C$ , as follows:

$$D_P(x, y) = T_{q \in P}(\succeq_q(x, y)).$$

Given  $(x, y) \in U \times U$ ,  $D_P(x, y)$  represents the credibility of the proposition “ $x$  is at least as good as  $y$  with respect to each criterion  $q$  from  $P$ ”. Since the fuzzy weak preference relations  $\succeq_q$  are supposed to be partial  $T$ -preorders, then also the fuzzy dominance relation  $D_P$  is a partial  $T$ -preorder. Furthermore, let  $\mathbf{Cl} = \{Cl_t, t = 1, \dots, n\}$  be a set of fuzzy classes in  $U$  such that, for each  $x \in U$ ,  $Cl_t(x)$  represents the membership function of  $x$  to  $Cl_t$ . It is supposed, as before, that the classes of  $\mathbf{Cl}$  are increasingly ordered, i.e. that for all  $r, s \in \{1, \dots, n\}$  such that  $r > s$ , the objects from  $Cl_r$  have a better comprehensive evaluation than the objects from

$Cl_s$ . On the basis of the membership functions of the fuzzy class  $Cl_t$ , fuzzy membership functions of two other sets can be defined as follows:

- 1) The upward union fuzzy set  $Cl_t^\geq$ , whose membership function  $Cl_t^\geq(x)$  represents the credibility of the proposition “ $x$  is at least as good as the objects in  $Cl_t$ ”:

$$Cl_t^\geq(x) = \left\{ \begin{array}{ll} 1 & \text{if } \exists s \in \{1, \dots, n\}: Cl_s(x) > 0 \\ Cl_t(x) & \text{otherwise} \end{array} \right\};$$

- 2) the downward union fuzzy set  $Cl_t^\leq$ , whose membership function  $Cl_t^\leq(x)$  represents the credibility of the proposition “ $x$  is at most as good as the objects in  $Cl_t$ ”:

$$Cl_t^\leq(x) = \left\{ \begin{array}{ll} 1 & \text{if } \exists s \in \{1, \dots, n\}: Cl_s(x) > 0 \\ Cl_t(x) & \text{otherwise} \end{array} \right\}.$$

The  $P$ -lower and the  $P$ -upper approximations of  $Cl_t^\geq$  with respect to  $P \subseteq C$  are fuzzy sets in  $U$ , whose membership functions, denoted by  $\underline{P}[Cl_t^\geq(x)]$  and  $\overline{P}[Cl_t^\geq(x)]$  respectively, are defined as:

$$\underline{P}[Cl_t^\geq(x)] = T_{y \in U}(T^*(N(D_P(y, x)), Cl_t^\geq(y))),$$

$$\overline{P}[Cl_t^\geq(x)] = T_{y \in U}^*(T(D_P(x, y), Cl_t^\geq(y))).$$

$\underline{P}[Cl_t^\geq(x)]$  represents the credibility of the proposition “for all  $y \in U$ ,  $y$  does not dominate  $x$  with respect to criteria from  $P$  or  $y$  belongs to  $Cl_t^\geq$ ”, while  $\overline{P}[Cl_t^\geq(x)]$  represents the credibility of the proposition “there is at least one  $y \in U$  dominated by  $x$  with respect to criteria from  $P$  which belongs to  $Cl_t^\geq$ ”.

The  $P$ -lower and  $P$ -upper approximations of  $Cl_t^\leq$  with respect to  $P \subseteq C$ , denoted by  $\underline{P}[Cl_t^\leq(x)]$  and  $\overline{P}[Cl_t^\leq(x)]$  respectively, can be defined, analogously, as:

$$\underline{P}[Cl_t^\leq(x)] = T_{y \in U}(T^*(N(D_P(x, y)), Cl_t^\leq(y))),$$

$$\overline{P}[Cl_t^\leq(x)] = T_{y \in U}^*(T(D_P(y, x), Cl_t^\leq(y))).$$

$\underline{P}[Cl_t^\leq(x)]$  represents the credibility of the proposition “for all  $y \in U$ ,  $x$  does not dominate  $y$  with respect to criteria from  $P$  or  $y$  belongs to  $Cl_t^\leq$ ”, while  $\overline{P}[Cl_t^\leq(x)]$  represents the credibility of the proposition “there is at least one  $y \in U$  dominating  $x$  with respect to criteria from  $P$  which belongs to  $Cl_t^\leq$ ”.

Let us remark that, using the definition of the  $T^*$ -implication, it is possible to rewrite the definitions of  $\underline{P}[Cl_t^\geq(x)]$ ,  $\overline{P}[Cl_t^\geq(x)]$ ,  $\underline{P}[Cl_t^\leq(x)]$  and  $\overline{P}[Cl_t^\leq(x)]$ , in the

following way:

$$\begin{aligned}\underline{P}[Cl_t^{\geq}(x)] &= T_{y \in U} \left( I_{T^*, N}^{\rightarrow} (D_P(y, x), Cl_t^{\geq}(y)) \right) , \\ \overline{P}[Cl_t^{\geq}(x)] &= \\ T_{y \in U}^* \left( N \left( I_{T^*, N}^{\rightarrow} (D_P(x, y), N(Cl_t^{\geq}(y))) \right) \right) , \\ \underline{P}[Cl_t^{\leq}(x)] &= T_{y \in U} \left( I_{T^*, N}^{\rightarrow} (D_P(x, y), Cl_t^{\leq}(y)) \right) , \\ \overline{P}[Cl_t^{\leq}(x)] &= \\ T_{y \in U}^* \left( N \left( I_{T^*, N}^{\rightarrow} (D_P(y, x), N(Cl_t^{\leq}(y))) \right) \right) .\end{aligned}$$

The following results can be proved:

- 1) for each  $x \in U$  and for each  $t \in \{1, \dots, n\}$ ,

$$\begin{aligned}\underline{P}[Cl_t^{\geq}(x)] &\leq Cl_t^{\geq}(x) \leq \overline{P}[Cl_t^{\geq}(x)] , \\ \underline{P}[Cl_t^{\leq}(x)] &\leq Cl_t^{\leq}(x) \leq \overline{P}[Cl_t^{\leq}(x)] ;\end{aligned}$$

- 2) if  $(T, T^*, N)$  constitute a de Morgan triplet and if  $N[Cl_t^{\geq}(x)] = Cl_{t-1}^{\leq}(x)$  for each  $x \in U$  and  $t = 2, \dots, n$ , then

$$\begin{aligned}\underline{P}[Cl_t^{\geq}(x)] &= N(\overline{P}[Cl_{t-1}^{\leq}(x)]) , \\ \overline{P}[Cl_t^{\geq}(x)] &= N(\underline{P}[Cl_{t-1}^{\leq}(x)]) , \quad t = 2, \dots, n , \\ \underline{P}[Cl_t^{\leq}(x)] &= N(\overline{P}[Cl_{t+1}^{\geq}(x)]) , \\ \overline{P}[Cl_t^{\leq}(x)] &= N(\underline{P}[Cl_{t+1}^{\geq}(x)]) , \quad t = 1, \dots, n-1 ;\end{aligned}$$

- 3) for all  $P \subseteq R \subseteq C$ , for all  $x \in U$  and for each  $t \in \{1, \dots, n\}$ ,

$$\begin{aligned}\underline{P}[Cl_t^{\geq}(x)] &\leq R[Cl_t^{\geq}(x)] , \quad \overline{P}[Cl_t^{\geq}(x)] \geq \overline{R}[Cl_t^{\geq}(x)] , \\ \underline{P}[Cl_t^{\leq}(x)] &\leq R[Cl_t^{\leq}(x)] , \quad \overline{P}[Cl_t^{\leq}(x)] \geq \overline{R}[Cl_t^{\leq}(x)] .\end{aligned}$$

Results 1) to 3) can be read as fuzzy counterparts of the following results well-known within the classical rough set approach:

- 1) (*inclusion property*) says that  $Cl_t^{\geq}$  and  $Cl_t^{\leq}$  include their  $P$ -lower approximations and are included in their  $P$ -upper approximations;
- 2) (*complementarity property*) says that the  $P$ -lower ( $P$ -upper) approximation of  $Cl_t^{\geq}$  is the complement of the  $P$ -upper ( $P$ -lower) approximation of its complementary set  $Cl_{t-1}^{\leq}$ , (analogous property holds for  $Cl_t^{\leq}$  and  $Cl_{t+1}^{\geq}$ );
- 3) (*monotonicity with respect to sets of attributes*) says that enlarging the set of criteria, the membership to the lower approximation does not decrease and the membership to the upper approximation does not increase.

Greco, Inuiguchi, and Słowiński [14] proposed, moreover, the following fuzzy rough approximations based on dominance, which go in line with the fuzzy rough approximation by Dubois and Prade [3,5], concerning classical rough sets:

$$\begin{aligned}\underline{P}[Cl_t^{\geq}(x)] &= \inf_{y \in U} (I(D_P(y, x), Cl_t^{\geq}(y))) , \\ \overline{P}[Cl_t^{\geq}(x)] &= \sup_{y \in U} (T(D_P(x, y), Cl_t^{\geq}(y))) , \\ \underline{P}[Cl_t^{\leq}(x)] &= \inf_{y \in U} (I(D_P(x, y), Cl_t^{\leq}(y))) , \\ \overline{P}[Cl_t^{\leq}(x)] &= \sup_{y \in U} (T(D_P(y, x), Cl_t^{\leq}(y))) .\end{aligned}$$

Using fuzzy rough approximations based on DRSA, one can induce decision rules having the same syntax as the decision rules obtained from crisp DRSA. In this case, however, each decision rule has a fuzzy credibility.

### Variable-Consistency Dominance-Based Rough Set Approach (VC-DRSA)

The definitions of rough approximations introduced in Sect. “Dominance-Based Rough Set Approach” are based on a strict application of the dominance principle. However, when defining non-ambiguous objects, it is reasonable to accept a limited proportion of negative examples, particularly for large data tables. Such an extended version of DRSA is called the variable-consistency DRSA model (VC-DRSA) [31].

For any  $P \subseteq C$ ,  $x \in U$  belongs to  $Cl_t^{\geq}$  without any ambiguity at consistency level  $l \in (0, 1]$ , if  $x \in Cl_t^{\geq}$  and at least  $l * 100\%$  of all objects  $y \in U$  dominating  $x$  with respect to  $P$  also belong to  $Cl_t^{\geq}$ , i.e., for  $t = 2, \dots, n$ ,

$$\frac{|D_P^+(x) \cap Cl_t^{\geq}|}{|D_P^+(x)|} \geq l .$$

The level  $l$  is called *consistency level* because it controls the degree of consistency with respect to objects qualified as belonging to  $Cl_t^{\geq}$  without any ambiguity. In other words, if  $l < 1$ , then at most  $(1 - l) * 100\%$  of all objects  $y \in U$  dominating  $x$  with respect to  $P$  do not belong to  $Cl_t^{\geq}$  and thus contradict the inclusion of  $x$  in  $Cl_t^{\geq}$ .

Analogously, for any  $P \subseteq C$ ,  $x \in U$  belongs to  $Cl_t^{\leq}$  without any ambiguity at consistency level  $l \in (0, 1]$ , if  $x \in Cl_t^{\leq}$  and at least  $l * 100\%$  of all the objects  $y \in U$  dominated by  $x$  with respect to  $P$  also belong to  $Cl_t^{\leq}$ , i.e., for  $t = 1, \dots, n-1$ ,

$$\frac{|D_P^-(x) \cap Cl_t^{\leq}|}{|D_P^-(x)|} \geq l .$$

The concept of non-ambiguous objects at some consistency level  $l$  leads naturally to the corresponding definition of  $P$ -lower approximations of the unions of classes

$Cl_t^{\geq}$  and  $Cl_t^{\leq}$ , respectively:

$$\underline{P}^l(Cl_t^{\geq}) = \left\{ x \in Cl_t^{\geq} : \frac{|D_P^+(x) \cap Cl_t^{\geq}|}{|D_P^+(x)|} \geq l \right\}, \\ t = 2, \dots, n,$$

$$\underline{P}^l(Cl_t^{\leq}) = \left\{ x \in Cl_t^{\leq} : \frac{|D_P^-(x) \cap Cl_t^{\leq}|}{|D_P^-(x)|} \geq l \right\}, \\ t = 1, \dots, n-1.$$

Given  $P \subseteq C$  and consistency level  $l$ , the corresponding  $P$ -upper approximations of  $Cl_t^{\geq}$  and  $Cl_t^{\leq}$ , denoted by  $\overline{P}^l(Cl_t^{\geq})$  and  $\overline{P}^l(Cl_t^{\leq})$ , respectively, can be defined as a complement of  $\underline{P}^l(Cl_{t-1}^{\leq})$  and  $\underline{P}^l(Cl_{t+1}^{\geq})$  with respect to  $U$ :

$$\overline{P}^l(Cl_t^{\geq}) = U - \underline{P}^l(Cl_{t-1}^{\leq}), \quad t = 2, \dots, n,$$

$$\overline{P}^l(Cl_t^{\leq}) = U - \underline{P}^l(Cl_{t+1}^{\geq}), \quad t = 1, \dots, n-1.$$

$\overline{P}^l(Cl_t^{\geq})$  can be interpreted as a set of all the objects belonging to  $Cl_t^{\geq}$ , possibly ambiguous at consistency level  $l$ . Analogously,  $\overline{P}^l(Cl_t^{\leq})$  can be interpreted as a set of all the objects belonging to  $Cl_t^{\leq}$ , possibly ambiguous at consistency level  $l$ . The  $P$ -boundaries ( $P$ -doubtful regions) of  $Cl_t^{\geq}$  and  $Cl_t^{\leq}$  at consistency level  $l$  are defined as:

$$Bn_P^l(Cl_t^{\geq}) = \overline{P}^l(Cl_t^{\geq}) - \underline{P}^l(Cl_t^{\geq}), \quad t = 2, \dots, n$$

$$Bn_P^l(Cl_t^{\leq}) = \overline{P}^l(Cl_t^{\leq}) - \underline{P}^l(Cl_t^{\leq}), \quad t = 1, \dots, n-1.$$

The variable consistency model of the dominance-based rough set approach provides some degree of flexibility in assigning objects to lower and upper approximations of the unions of decision classes. The following properties can be easily proved: for  $0 < l' < l \leq 1$ ,

$$\underline{P}^l(Cl_t^{\geq}) \subseteq \underline{P}^{l'}(Cl_t^{\geq}) \quad \text{and} \quad \overline{P}^l(Cl_t^{\geq}) \supseteq \overline{P}^{l'}(Cl_t^{\geq}), \\ t = 2, \dots, n,$$

$$\underline{P}^l(Cl_t^{\leq}) \subseteq \underline{P}^{l'}(Cl_t^{\leq}) \quad \text{and} \quad \overline{P}^l(Cl_t^{\leq}) \supseteq \overline{P}^{l'}(Cl_t^{\leq}), \\ t = 1, \dots, n-1.$$

The following two basic types of variable-consistency decision rules can be considered:

1.  $D_{\geq}$ -decision rules with the following syntax: "if  $\phi(x, q_1) \geq r_{q_1}$  and  $\phi(x, q_2) \geq r_{q_2}$  and ...  $\phi(x, q_p) \geq r_{q_p}$ , then  $x \in Cl_t^{\geq}$ " with confidence  $\alpha$  (i.e. in fraction  $\alpha$  of considered cases), where  $P = \{q_1, \dots, q_p\} \subseteq C$ ,  $(r_{q_1}, \dots, r_{q_p}) \in V_{q_1} \times V_{q_2} \times \dots \times V_{q_p}$  and  $t = 2, \dots, n$ ;

2.  $D_{\leq}$ -decision rules with the following syntax: "if  $\phi(x, q_1) \leq r_{q_1}$  and  $\phi(x, q_2) \leq r_{q_2}$  and ...  $\phi(x, q_p) \leq r_{q_p}$ , then  $x \in Cl_t^{\leq}$ " with confidence  $\alpha$ , where  $P = \{q_1, \dots, q_p\} \subseteq C$ ,  $(r_{q_1}, \dots, r_{q_p}) \in V_{q_1} \times V_{q_2} \times \dots \times V_{q_p}$  and  $t = 1, \dots, n-1$ .

The *variable consistency* model is inspired by the *variable precision* model proposed by Ziarko [61,62] within the classical indiscernibility-based rough set approach.

### Dominance-Based Rough Approximation of a Fuzzy Set

This section shows how the dominance-based rough set approach can be used for rough approximation of fuzzy sets.

A *fuzzy information base* is the 3-tuple  $B = \langle U, F, \varphi \rangle$ , where  $U$  is a finite set of *objects* (universe),  $F = \{f_1, f_2, \dots, f_m\}$  is a finite set of *properties*, and  $\varphi: U \times F \rightarrow [0, 1]$  is a function such that  $\varphi(x, f_h) \in [0, 1]$  expresses the credibility that object  $x$  has property  $f_h$ . Each object  $x$  from  $U$  is described by a vector

$$\text{Des}_F(x) = [\varphi(x, f_1), \dots, \varphi(x, f_m)],$$

called *description* of  $x$  in terms of the degrees to which it has properties from  $F$ ; it represents the available information about  $x$ . Obviously,  $x \in U$  can be described in terms of any non-empty subset  $E \subseteq F$  and in this case

$$\text{Des}_E(x) = [\varphi(x, f_h), f_h \in E].$$

For any  $E \subseteq F$ , the dominance relation  $D_E$  can be defined as follows: for all  $x, y \in U$ ,  $x$  dominates  $y$  with respect to  $E$  (denotation  $x D_E y$ ) if, for any  $f_h \in E$ ,

$$\varphi(x, f_h) \geq \varphi(y, f_h).$$

Given  $E \subseteq F$  and  $x \in U$ , let

$$D_E^+(x) = \{y \in U : y D_E x\}, \quad D_E^-(x) = \{y \in U : x D_E y\}.$$

Let us consider a fuzzy set  $X$  in  $U$ , with its membership function  $\mu_X: U \rightarrow [0, 1]$ . For each cutting level  $\alpha \in [0, 1]$  and for  $* \in \{\geq, >\}$ , the  $E$ -lower and the  $E$ -upper approximation of  $X^{*\alpha} = \{y \in U : \mu_X(y) * \alpha\}$  with respect to  $E \subseteq F$  (denotation  $\underline{E}(X^{*\alpha})$  and  $\overline{E}(X^{*\alpha})$ , respectively), can be defined as:

$$\underline{E}(X^{*\alpha}) = \{x \in U : D_E^+(x) \subseteq X^{*\alpha}\},$$

$$\overline{E}(X^{*\alpha}) = \{x \in U : D_E^-(x) \cap X^{*\alpha} \neq \emptyset\}.$$

Rough approximations  $\underline{E}(X^{*\alpha})$  and  $\overline{E}(X^{*\alpha})$  can be expressed in terms of unions of granules  $D_E^+(x)$  as follows:

$$\begin{aligned}\underline{E}(X^{*\alpha}) &= \bigcup_{x \in U} \{D_E^+(x) : D_E^+(x) \subseteq X^{*\alpha}\}, \\ \overline{E}(X^{*\alpha}) &= \bigcup_{x \in U} \{D_E^+(x) : D_E^-(x) \cap X^{*\alpha} \neq \emptyset\}.\end{aligned}$$

Analogously, for each cutting level  $\alpha \in [0, 1]$  and for  $\diamond \in \{\leq, <\}$ , the  $E$ -lower and the  $E$ -upper approximation of  $X^{\diamond\alpha} = \{y \in U : \mu_X(y) \diamond \alpha\}$ , with respect to  $E \subseteq F$  (denotation  $\underline{E}(X^{\diamond\alpha})$  and  $\overline{E}(X^{\diamond\alpha})$ , respectively), can be defined as:

$$\begin{aligned}\underline{E}(X^{\diamond\alpha}) &= \{x \in U : D_E^-(x) \subseteq X^{\diamond\alpha}\}, \\ \overline{E}(X^{\diamond\alpha}) &= \{x \in U : D_E^+(x) \cap X^{\diamond\alpha} \neq \emptyset\}.\end{aligned}$$

The rough approximations  $\underline{E}(X^{\diamond\alpha})$  and  $\overline{E}(X^{\diamond\alpha})$  can be expressed in terms of unions of granules  $D_E^-(x)$  as follows:

$$\begin{aligned}\underline{E}(X^{\diamond\alpha}) &= \bigcup_{x \in U} \{D_E^-(x) : D_E^-(x) \subseteq X^{\diamond\alpha}\}, \\ \overline{E}(X^{\diamond\alpha}) &= \bigcup_{x \in U} \{D_E^-(x) : D_E^+(x) \cap X^{\diamond\alpha} \neq \emptyset\}.\end{aligned}$$

Let us remark that the rough approximations  $\underline{E}(X^{\geq\alpha})$ ,  $\overline{E}(X^{\geq\alpha})$ ,  $\underline{E}(X^{\leq\alpha})$  and  $\overline{E}(X^{\leq\alpha})$  can be rewritten as follows:

$$\begin{aligned}\underline{E}(X^{\geq\alpha}) &= \{x \in U : \forall w \in U, w D_E x \Rightarrow w \in X^{\geq\alpha}\}, \\ \overline{E}(X^{\geq\alpha}) &= \{x \in U : \exists w \in U \text{ such that } x D_E w \text{ and } w \in X^{\geq\alpha}\}, \\ \underline{E}(X^{\leq\alpha}) &= \{x \in U : \forall w \in U, x D_E w \Rightarrow w \in X^{\leq\alpha}\}, \\ \overline{E}(X^{\leq\alpha}) &= \{x \in U : \exists w \in U \text{ such that } w D_E x \text{ and } w \in X^{\leq\alpha}\}.\end{aligned}$$

Rough approximations  $\underline{E}(X^{>\alpha})$ ,  $\overline{E}(X^{>\alpha})$ ,  $\underline{E}(X^{<\alpha})$  and  $\overline{E}(X^{<\alpha})$  can be rewritten analogously by a simple replacement of “ $\geq$ ” with “ $>$ ”, and “ $\leq$ ” with “ $<$ ”.

This reformulation of the rough approximations is concordant with the syntax of decision rules obtained in DRSA. For example,  $\underline{E}(X^{\geq\alpha})$  is concordant with decision rules of the type

*“if object  $y$  has property  $f_{i1}$  to degree at least  $h_{i1}$ , and has property  $f_{i2}$  to degree at least  $h_{i2}$ , ..., and has property  $f_{ip}$  to degree at least  $h_{ip}$ , then object  $y$  belongs to set  $X$  to degree at least  $\alpha$ ”,*

where  $\{i1, \dots, ip\} = E$  and  $h_{i1} = \varphi(x, f_{i1}), \dots, h_{ip} = \varphi(x, f_{ip})$ .

Let us remark that in the above approximations, even if  $X^{\geq\alpha} = Y^{\leq\alpha}$ , their approximations are, in general, different

due to the different directions of cutting the membership functions of  $X$  and  $Y$ . Of course, a similar remark holds also for  $X^{>\alpha}$  and  $Y^{<\alpha}$ . Considerations of the directions in the cuts  $X^{\geq\alpha}$ ,  $X^{>\alpha}$  and  $X^{\leq\alpha}$ ,  $X^{<\alpha}$  are important in the definition of the rough approximations of unions and intersections of cuts.

The rough approximations  $\underline{E}(X^{\geq\alpha})$ ,  $\overline{E}(X^{\geq\alpha})$ ,  $\underline{E}(X^{\leq\alpha})$ ,  $\overline{E}(X^{\leq\alpha})$  and  $\underline{E}(X^{>\alpha})$ ,  $\overline{E}(X^{>\alpha})$ ,  $\underline{E}(X^{<\alpha})$ ,  $\overline{E}(X^{<\alpha})$  satisfy the following inclusion properties: for any  $0 \leq \alpha \leq 1$ ,

$$\begin{aligned}\underline{E}(X^{\geq\alpha}) &\subseteq X^{\geq\alpha} \subseteq \overline{E}(X^{\geq\alpha}), \\ \underline{E}(X^{\leq\alpha}) &\subseteq X^{\leq\alpha} \subseteq \overline{E}(X^{\leq\alpha}), \\ \underline{E}(X^{>\alpha}) &\subseteq X^{>\alpha} \subseteq \overline{E}(X^{>\alpha}), \\ \underline{E}(X^{<\alpha}) &\subseteq X^{<\alpha} \subseteq \overline{E}(X^{<\alpha}).\end{aligned}$$

Furthermore, the following complementary properties hold: for any  $0 \leq \alpha \leq 1$ ,

$$\begin{aligned}\underline{E}(X^{\geq\alpha}) &= U - \overline{E}(X^{\leq\alpha}), & \underline{E}(X^{\leq\alpha}) &= U - \overline{E}(X^{>\alpha}), \\ \underline{E}(X^{>\alpha}) &= U - \overline{E}(X^{\leq\alpha}), & \underline{E}(X^{<\alpha}) &= U - \overline{E}(X^{\geq\alpha}).\end{aligned}$$

The following properties of monotonicity with respect to sets of properties also hold: for any  $E_1 \subseteq E_2 \subseteq F$  and for any  $0 \leq \alpha \leq 1$ ,

$$\begin{aligned}\underline{E}_1(X^{\geq\alpha}) &\subseteq \underline{E}_2(X^{\geq\alpha}), & \underline{E}_1(X^{>\alpha}) &\subseteq \underline{E}_2(X^{>\alpha}), \\ \underline{E}_1(X^{\leq\alpha}) &\subseteq \underline{E}_2(X^{\leq\alpha}), & \underline{E}_1(X^{<\alpha}) &\subseteq \underline{E}_2(X^{<\alpha}), \\ \overline{E}_1(X^{\geq\alpha}) &\supseteq \overline{E}_2(X^{\geq\alpha}), & \overline{E}_1(X^{>\alpha}) &\supseteq \overline{E}_2(X^{>\alpha}), \\ \overline{E}_1(X^{\leq\alpha}) &\supseteq \overline{E}_2(X^{\leq\alpha}), & \overline{E}_1(X^{<\alpha}) &\supseteq \overline{E}_2(X^{<\alpha}).\end{aligned}$$

One can consider also fuzzy rough approximations  $\underline{X}_E^{\uparrow}$ ,  $\underline{X}_E^{\downarrow}$ ,  $\overline{X}_E^{\uparrow}$ ,  $\overline{X}_E^{\downarrow}$ , which are fuzzy sets with membership functions defined, respectively, as follows: for any  $y \in U$ ,

$$\begin{aligned}\mu_{\underline{X}_E^{\uparrow}}(y) &= \max \{\alpha \in [0, 1] : y \in \underline{E}(X^{\geq\alpha})\}, \\ \mu_{\underline{X}_E^{\downarrow}}(y) &= \min \{\alpha \in [0, 1] : y \in \underline{E}(X^{\leq\alpha})\}, \\ \mu_{\overline{X}_E^{\uparrow}}(y) &= \max \{\alpha \in [0, 1] : y \in \overline{E}(X^{\geq\alpha})\}, \\ \mu_{\overline{X}_E^{\downarrow}}(y) &= \min \{\alpha \in [0, 1] : y \in \overline{E}(X^{\leq\alpha})\}.\end{aligned}$$

The membership function  $\mu_{\underline{X}_E^{\uparrow}}(y)$  is defined as the upward lower fuzzy rough approximation of  $X$  with respect to  $E$  and can be interpreted in the following way. For any  $\alpha, \beta \in [0, 1]$ ,  $\alpha < \beta$  implies  $X^{\geq\alpha} \supseteq X^{\geq\beta}$ . Therefore, the greater the cutting level  $\alpha$ , the smaller  $X^{\geq\alpha}$  and, consequently, the smaller also its lower approximation  $\underline{E}(X^{\geq\alpha})$ . Thus, for each  $y \in U$  and for each fuzzy set  $X$ , there is a threshold  $k(y)$ ,  $0 \leq k(y) \leq \mu_X(y)$ , such that  $y \in \underline{E}(X^{\geq\alpha})$  if  $\alpha \leq k(y)$ , and  $y \notin \underline{E}(X^{\geq\alpha})$  if  $\alpha > k(y)$ .

Since  $k(y) = \mu_{\underline{X}_E}^{\uparrow}(y)$ , this explains the interest of  $\mu_{\underline{X}_E}^{\uparrow}(y)$ . Analogous interpretation holds for  $\mu_{\overline{X}_E}^{\uparrow}(y)$ , defined as the membership function of the upward upper fuzzy rough approximation of  $X$  with respect to  $E$ .

The membership function  $\mu_{\overline{X}_E}^{\downarrow}(y)$  is defined as the downward lower fuzzy rough approximation of  $X$  with respect to  $E$  and can be interpreted as follows. For any  $\alpha, \beta \in [0, 1]$ ,  $\alpha < \beta$  implies  $X^{\leq\alpha} \subseteq X^{\leq\beta}$ . Therefore, the greater the cutting level  $\alpha$ , the greater  $X^{\leq\alpha}$  and, consequently, its lower approximation  $\underline{E}(X^{\geq\alpha})$ . Thus, for each  $y \in U$  and for each fuzzy set  $X$ , there is a threshold  $h(y)$ ,  $\mu_X(y) \leq h(y) \leq 1$ , such that  $y \in \underline{E}(X^{\leq\alpha})$  if  $\alpha \geq h(y)$ , and  $y \notin \underline{E}(X^{\leq\alpha})$  if  $\alpha < h(y)$ . Observe that  $h(y) = \mu_{\overline{X}_E}^{\downarrow}(y)$ . Analogous interpretation holds for  $\mu_{\underline{X}_E}^{\downarrow}(y)$ , defined as the membership function of the downward upper fuzzy rough approximation of  $X$  with respect to  $E$ .

The membership functions of the upward and downward lower and upper fuzzy rough approximations can also be rewritten in the following equivalent formulations, which has been proposed and investigated by Greco, Inuiguchi, and Słowiński [16]:

$$\begin{aligned}\mu_{\underline{X}_E}^{\uparrow}(y) &= \min \{\mu_X(z) : z \in D_E^+(y)\}, \\ \mu_{\overline{X}_E}^{\uparrow}(y) &= \max \{\mu_X(z) : z \in D_E^-(y)\}, \\ \mu_{\underline{X}_E}^{\downarrow}(y) &= \max \{\mu_X(z) : z \in D_E^-(y)\}, \\ \mu_{\overline{X}_E}^{\downarrow}(y) &= \min \{\mu_X(z) : z \in D_E^+(y)\}.\end{aligned}$$

The membership functions of the fuzzy rough approximations  $\mu_{\underline{X}_E}^{\uparrow}(y)$ ,  $\mu_{\overline{X}_E}^{\uparrow}(y)$ ,  $\mu_{\underline{X}_E}^{\downarrow}(y)$  and  $\mu_{\overline{X}_E}^{\downarrow}(y)$ , satisfy the following inclusion properties: for any  $y \in U$ ,

$$\begin{aligned}\mu_{\underline{X}_E}^{\uparrow}(y) &\leq \mu_X(y) \leq \mu_{\overline{X}_E}^{\uparrow}(y), \\ \mu_{\underline{X}_E}^{\downarrow}(y) &\leq \mu_X(y) \leq \mu_{\overline{X}_E}^{\downarrow}(y).\end{aligned}$$

Furthermore, the following complementary property holds: for any  $y \in U$ ,

$$\mu_{\underline{X}_E}^{\uparrow}(y) = \mu_{\overline{X}_E}^{\downarrow}(y), \quad \mu_{\underline{X}_E}^{\downarrow}(y) = \mu_{\overline{X}_E}^{\uparrow}(y).$$

The following properties of monotonicity with respect to sets of properties also hold: for any  $E_1 \subseteq E_2 \subseteq F$ ,

$$\begin{aligned}\mu_{\underline{X}_{E_1}}^{\uparrow}(y) &\leq \mu_{\underline{X}_{E_2}}^{\uparrow}(y), \quad \mu_{\underline{X}_{E_1}}^{\downarrow}(y) \geq \mu_{\underline{X}_{E_2}}^{\downarrow}(y), \\ \mu_{\overline{X}_{E_1}}^{\uparrow}(y) &\geq \mu_{\overline{X}_{E_2}}^{\uparrow}(y), \quad \mu_{\overline{X}_{E_1}}^{\downarrow}(y) \leq \mu_{\overline{X}_{E_2}}^{\downarrow}(y).\end{aligned}$$

### Monotonic Rough Approximation of a Fuzzy Set Versus Classical Rough Set

What is the relationship between the classical rough set and the DRSA approximation of a fuzzy set? Greco,

Matarazzo, and Słowiński [27,29] proved that the former is a particular case of the latter, as shown below.

Let us remember that in the classical rough set approach [46,47], the original information is expressed by means of an *information system*, that is the 4-tuple  $S = \langle U, Q, V, \phi \rangle$ , where  $U$  is a finite set of *objects* (universe),  $Q = \{q_1, q_2, \dots, q_m\}$  is a finite set of *attributes*,  $V_q$  is the set of values of the attribute  $q$ ,  $V = \bigcup_{q \in Q} V_q$  and  $\phi: U \times Q \rightarrow V$  is a total function such that  $\phi(x, q) \in V_q$  for each  $q \in Q$ ,  $x \in U$ , called *information function*.

Therefore, each object  $x$  from  $U$  is described by a vector

$$\text{Des}_Q(x) = [\phi(x, q_1), \phi(x, q_2), \dots, \phi(x, q_m)],$$

called *description* of  $x$  in terms of the evaluations of the attributes from  $Q$ ; it represents the available information about  $x$ . Obviously,  $x \in U$  can be described in terms of any non-empty subset  $P \subseteq Q$ .

With every (non-empty) subset of attributes  $P$  there is associated an *indiscernibility relation* on  $U$ , denoted by  $I_P$ :

$$I_P = \{(x, y) \in U \times U : \phi(x, q) = \phi(y, q), \forall q \in P\}.$$

If  $(x, y) \in I_P$ , it is said that the objects  $x$  and  $y$  are  $P$ -indiscernible. Clearly, the indiscernibility relation thus defined is an equivalence relation (reflexive, symmetric, and transitive). The family of all the equivalence classes of the relation  $I_P$  is denoted by  $U|I_P$ , and the equivalence class containing an element  $x \in U$  by  $I_P(x)$ , i.e.

$$I_P(x) = \{y \in U : \phi(y, q) = \phi(x, q), \forall q \in P\}.$$

The equivalence classes of the relation  $I_P$  are called  *$P$ -elementary sets*.

Let  $S$  be an information system,  $X$  a non-empty subset of  $U$  and  $\emptyset \neq P \subseteq Q$ . The  *$P$ -lower approximation* and the  *$P$ -upper approximation* of  $X$  in  $S$  are defined, respectively, as:

$$\underline{P}(X) = \{x \in U : I_P(x) \subseteq X\},$$

$$\overline{P}(X) = \{x \in U : I_P(x) \cap X \neq \emptyset\}.$$

The elements of  $\underline{P}(X)$  are all and only those objects  $x \in U$  which belong to the equivalence classes generated by the indiscernibility relation  $I_P$ , contained in  $X$ ; the elements of  $\overline{P}(X)$  are all and only those objects  $x \in U$  which belong to the equivalence classes generated by the indiscernibility relation  $I_P$ , containing at least one object  $x$  belonging to  $X$ . In other words,  $\underline{P}(X)$  is the largest union of the  $P$ -elementary sets included in  $X$ , while  $\overline{P}(X)$  is the smallest union of the  $P$ -elementary sets containing  $X$ .

**Granular Computing and Data Mining for Ordered Data: The Dominance-Based Rough Set Approach, Table 1**  
Information base  $B$

	$f_1$	$f_2$
$x_1$	0	1
$x_2$	1	0
$x_3$	1	1

Any information system can be expressed in terms of a specific type of an information base. An *information base* is called *Boolean* if  $\varphi: U \times F \rightarrow \{0, 1\}$ . A partition  $F = \{F_1, \dots, F_r\}$  of the set of properties  $F$ , with  $\text{card}(F_k) \geq 2$  for all  $k = 1, \dots, r$ , is called *canonical* if, for each  $x \in U$  and for each  $F_k \in F$ ,  $k = 1, \dots, r$ , there exists only one  $f_j \in F_k$  for which  $\varphi(x, f_j) = 1$  (and, therefore, for each  $f_i \in F_k - \{f_j\}$ ,  $\varphi(x, f_i) = 0$ ). The condition  $\text{card}(F_k) \geq 2$  for all  $k = 1, \dots, r$ , is necessary because, otherwise, there would be at least one class  $F_k = \{f'\}$  of the partition such that  $\varphi(x, f') = 1$  for all  $x \in U$ , and this would mean that property  $f'$  gives no information and can be removed. Observe now that any *information system*  $S = \langle U, Q, V, \phi \rangle$  can be transformed to a Boolean information base  $B = \langle U, F, \varphi \rangle$  assigning to each  $v \in V_q$ ,  $q \in Q$ , one property  $f_{qv} \in F$  such that  $\varphi(x, f_{qv}) = 1$  if  $\phi(x, q) = v$ , and  $\varphi(x, f_{qv}) = 0$  otherwise. Let us remark that  $F = \{F_1, \dots, F_r\}$ , with  $F_q = \{f_{qv}, v \in V_q\}$ ,  $q \in Q$ , is a canonical partition of  $F$ . The opposite transformation, from a Boolean information base to an information system, is not always possible, i. e. there may exist Boolean information bases which cannot be transformed into information systems, because their sets of properties do not admit any canonical partition, as shown by the following example.

**Example 1** Let us consider a Boolean information base  $B$ , such that  $U = \{x_1, x_2, x_3\}$ ,  $F = \{f_1, f_2\}$  and function  $\varphi$  is defined by Table 1. One can see that  $F = \{\{f_1, f_2\}\}$  is not a canonical partition because  $\varphi(x_3, f_1) = \varphi(x_3, f_2) = 1$ , while the definition of canonical partition  $F$  does not allow that for an object  $x \in U$ ,  $\varphi(x, f_1) = \varphi(x, f_2) = 1$ . Therefore, this Boolean information base has no equivalent information system. Let us remark that also the Boolean information base  $B'$  presented in Table 2, where  $U = \{x_1, x_2, x_4\}$  and  $F = \{f_1, f_2\}$ , cannot be transformed to an information system, because partition  $F = \{\{f_1, f_2\}\}$  is not canonical. Indeed,  $\varphi(x_4, f_1) = \varphi(x_4, f_2) = 0$ , while definition of canonical partition  $F$  does not allow that for an object  $x \in U$ ,  $\varphi(x, f_1) = \varphi(x, f_2) = 0$ .

The above says that consideration of rough approximation in the context of a Boolean information base is

**Granular Computing and Data Mining for Ordered Data: The Dominance-Based Rough Set Approach, Table 2**  
Information base  $B'$

	$f_1$	$f_2$
$x_1$	0	1
$x_2$	1	0
$x_4$	0	0

**Granular Computing and Data Mining for Ordered Data: The Dominance-Based Rough Set Approach, Table 3**  
Information system  $S$

	$q_1$	$q_2$
$x_1$	0	1
$x_2$	1	0
$x_3$	1	1

more general than the same consideration in the context of an information system. This means, of course, that the rough approximation considered in the context of a fuzzy information base is yet more general.

It is worth stressing that the Boolean information bases  $B$  and  $B'$  are not Boolean information systems. In fact, on one hand, a Boolean information base provides information about absence ( $\varphi(x, f) = 0$ ) or presence ( $\varphi(x, f) = 1$ ) of properties  $f \in F$  in objects  $x \in U$ . On the other hand, a Boolean information system provides information about values assigned by attributes  $q \in Q$ , whose sets of values are  $V_q = \{0, 1\}$ , to objects  $x \in U$ , such that  $\phi(x, q) = 1$  or  $\phi(x, q) = 0$  for all  $x \in U$  and  $q \in Q$ . Observe, therefore, that to transform a Boolean information system  $S$  into a Boolean information base  $B$ , each attribute  $q$  of  $S$  corresponds to two properties  $f_{q0}$  and  $f_{q1}$  of  $B$ , such that for all  $x \in U$

- $\varphi(x, f_{q0}) = 1$  and  $\varphi(x, f_{q1}) = 0$  if  $\phi(x, q) = 0$ ,
- $\varphi(x, f_{q0}) = 0$  and  $\varphi(x, f_{q1}) = 1$  if  $\phi(x, q) = 1$ .

Thus, the Boolean information base  $B$  in Table 1 and the Boolean information system  $S$  in Table 3 are different, though they could seem identical. In fact, the Boolean information system  $S$  in Table 3 can be transformed into the Boolean information base  $B''$  in Table 4, which is clearly different from  $B$ .

The equivalence between rough approximations in the context of a fuzzy information base and the classical definition of rough approximations in the context of an information system can be stated as follows [27,29]. Let us consider an information system and the corresponding Boolean information base; for each  $P \subseteq Q$ , let  $E^P$  be the

**Granular Computing and Data Mining for Ordered Data: The Dominance-Based Rough Set Approach, Table 4**  
**Information base  $B''$**

	$f_{q10}$	$f_{q11}$	$f_{q20}$	$f_{q21}$
$x_1$	1	0	0	1
$x_2$	0	1	1	0
$x_3$	0	1	0	1

set of all the properties corresponding to values  $v$  of attributes in  $P$ . Let  $X$  be a non-fuzzy (classical) set in  $U$  (i.e.  $\mu_X: U \rightarrow \{0, 1\}$  and, therefore, for any  $y \in U$ ,  $\mu_X(y) = 1$  or  $\mu_X(y) = 0$ ); then it is:

$$\underline{E}^P(X^{\geq 1}) = \underline{P}(X^{\geq 1}), \quad \overline{E}^P(X^{\geq 1}) = \overline{P}(X^{\geq 1}), \\ \underline{E}^P(X^{\leq 0}) = \underline{P}(U - X^{\geq 1}), \quad \overline{E}^P(X^{\leq 0}) = \overline{P}(U - X^{\geq 1}).$$

This result proves that the rough approximation of a non-fuzzy set  $X$  in a Boolean information base admitting a canonical partition is equivalent to the classical rough approximation of set  $X$  in the corresponding information system. Therefore, the classical rough approximation is a particular case of the dominance-based rough approximation in a fuzzy information base.

### Dominance-Based Rough Set Approach to Case-Based Reasoning

This section presents rough approximation of a fuzzy set using a similarity relation in the context of case-based reasoning [28].

Case-based reasoning (for a general introduction to case-based reasoning see e.g. [35]; for a fuzzy set approach to case-based reasoning see [6]) is a paradigm in machine learning whose idea is that a new problem can be solved by noticing its similarity to a set of problems previously solved. Case-based reasoning regards the inference of some proper conclusions related to a new situation by the analysis of similar cases from a memory of previous cases. It is based on two principles [36]:

- a) Similar problems have similar solutions;
- b) Types of encountered problems tend to recur.

Gilboa and Schmeidler [12] observed that the basic idea of case-based reasoning can be found in the following sentence of Hume [33]: “From causes which appear *similar* we expect similar effects. This is the sum of all our experimental conclusions.” Rephrasing Hume, one can say that “the more similar are the causes, the more similar one expects the effects.” Therefore, measuring similarity is the essential point of all case-based reasoning and, particularly, of the

fuzzy set approach to case-based reasoning [6]. This explains the many problems that measuring similarity generates within case-based reasoning. Problems of modeling similarity are relative to two levels:

- At the level of similarity with respect to single features: how to define a meaningful similarity measure with respect to a single feature?
- At the level of similarity with respect to all features: how to properly aggregate the similarity measures with respect to single features in order to obtain a comprehensive similarity measure?

For the above reasons, [28] proposes a DRSA approach to case-based reasoning, which tries to be possibly “neutral” and “objective” with respect to similarity relation. At the level of similarity concerning single features, the DRSA approach to case-based reasoning considers only ordinal properties of similarity, and at the level of aggregation, it does not impose any particular functional aggregation based on some very specific axioms (see, for example, [12]), but it considers a set of decision rules based on the general monotonicity property of comprehensive similarity with respect to similarity of single features. Therefore, the DRSA approach to case-based reasoning is only very little “invasive”, compared to the many other existing approaches.

Let us consider a *pairwise fuzzy information base* being the 3-tuple

$$\mathcal{B} = \langle U, F, \sigma \rangle,$$

where  $U$  is a finite set of *objects* (universe),  $F = \{f_1, f_2, \dots, f_m\}$  is a finite set of *features*, and  $\sigma: U \times U \times F \rightarrow [0, 1]$  is a function such that  $\sigma(x, y, f_h) \in [0, 1]$  expresses the credibility that object  $x$  is similar to object  $y$  w.r.t. feature  $f_h$ . The minimal requirement function  $\sigma$  must satisfy is that, for all  $x \in U$  and for all  $f_h \in F$ ,  $\sigma(x, x, f_h) = 1$ . Therefore, each pair of objects  $(x, y) \in U \times U$  is described by a vector

$$\text{Des}_F(x, y) = [\sigma(x, y, f_1), \dots, \sigma(x, y, f_m)],$$

called *description* of  $(x, y)$  in terms of the credibilities of similarity with respect to features from  $F$ ; it represents the available information about similarity between  $x$  and  $y$ . Obviously, similarity between  $x$  and  $y$ ,  $x, y \in U$ , can be described in terms of any non-empty subset  $E \subseteq F$  as follows:

$$\text{Des}_E(x, y) = [\sigma(x, y, f_h), f_h \in E].$$

With respect to any  $E \subseteq F$ , the dominance relation  $D_E$  can be defined on  $U \times U$  as follows: for any  $x, y, w, z \in U$ ,

$(x, y)$  dominates  $(w, z)$  with respect to  $E$  (denotation  $(x, y)D_E(w, z)$ ) if, for any  $f_h \in E$ ,

$$\sigma(x, y, f_h) \geq \sigma(w, z, f_h).$$

Given  $E \subseteq F$  and  $x, y \in U$ , let

$$D_E^+(y, x) = \{w \in U : (w, x)D_E(y, x)\},$$

$$D_E^-(y, x) = \{w \in U : (y, x)D_E(w, x)\}.$$

In the pair  $(y, x)$ ,  $x$  is considered to be a *reference object*, while  $y$  can be called a *limit object*, because it is conditioning the membership of  $w$  in  $D_E^+(y, x)$  and in  $D_E^-(y, x)$ .

For each  $x \in U$  and  $\alpha \in [0, 1]$  and  $* \in \{\geq, >\}$ , the lower approximation of  $X^{*\alpha}$ ,  $\underline{E}_\sigma(X^{*\alpha})$ , and the upper approximation of  $X^{*\alpha}$ ,  $\overline{E}_\sigma(X^{*\alpha})$ , based on similarity  $\sigma$  with respect to  $E \subseteq F$  and  $x$ , respectively, can be defined as:

$$\underline{E}(x)_\sigma(X^{*\alpha}) = \{y \in U : D_E^+(y, x) \subseteq X^{*\alpha}\},$$

$$\overline{E}(x)_\sigma(X^{*\alpha}) = \{y \in U : D_E^-(y, x) \cap X^{*\alpha} \neq \emptyset\}.$$

For the sake of simplicity, in the following only  $\underline{E}(x)_\sigma(X^{\geq\alpha})$  and  $\overline{E}(x)_\sigma(X^{\geq\alpha})$  with  $x \in X^{\geq\alpha}$  are considered. Of course, analogous considerations hold for  $\underline{E}(x)_\sigma(X^{>\alpha})$  and  $\overline{E}(x)_\sigma(X^{>\alpha})$ . Observe that the lower approximation of  $X^{\geq\alpha}$  with respect to  $x$  contains all the objects  $y \in U$  such that any object  $w$ , being similar to  $x$  at least as much as  $y$  is similar to  $x$  w.r.t. all the considered features  $E \subseteq F$ , also belongs to  $X^{\geq\alpha}$ . Thus, the data from the fuzzy pairwise information base  $B$  confirm that if  $w$  is similar to  $x$  not less than  $y \in \underline{E}(x)_\sigma(X^{\geq\alpha})$  is similar to  $x$  w.r.t. all the considered features  $E \subseteq F$ , then  $w$  belongs to  $X^{\geq\alpha}$ . In other words,  $x$  is a reference object and  $y \in \underline{E}(x)_\sigma(X^{\geq\alpha})$  is a limit object which belongs “certainly” to set  $X$  with credibility at most  $\alpha$ ; the limit is understood such that all objects  $w$  that are similar to  $x$  w.r.t. considered features at least as much as  $y$  is similar to  $x$ , also belong to  $X$  with credibility at most  $\alpha$ .

Analogously, the upper approximation of  $X^{\geq\alpha}$  with respect to  $x$  contains all objects  $y \in U$  such that there is at least one object  $w$ , being similar to  $x$  at most as much as  $y$  is similar to  $x$  w.r.t. all the considered features  $E \subseteq F$ , which belongs to  $X^{\geq\alpha}$ . Thus, the data from the fuzzy pairwise information base  $B$  confirm that if  $w$  is similar to  $x$  not less than  $y \in \overline{E}(x)_\sigma(X^{\geq\alpha})$  is similar to  $x$  w.r.t. all the considered features  $E \subseteq F$ , then it is possible that  $w$  belongs to  $X^{\geq\alpha}$ . In other words,  $x$  is a reference object and  $y \in \overline{E}(x)_\sigma(X^{\geq\alpha})$  is a limit object which belongs “possibly” to set  $X$  with credibility at least  $\alpha$ ; the limit is understood such that all objects  $z \in U$  similar to  $x$  not less than  $y$  w.r.t. considered features, possibly belong to  $X^{\geq\alpha}$ .

For each  $x \in U$  and  $\alpha \in [0, 1]$  and  $\diamond \in \{\leq, <\}$ , the lower approximation of  $X^{\diamond\alpha}$ ,  $\underline{E}(x)_\sigma(X^{\diamond\alpha})$ , and the upper approximation of  $X^{\diamond\alpha}$ ,  $\overline{E}(x)_\sigma(X^{\diamond\alpha})$ , based on similarity  $\sigma$  with respect to  $E \subseteq F$  and  $x$ , respectively, can be defined as:

$$\underline{E}(x)_\sigma(X^{\diamond\alpha}) = \{y \in U : D_E^-(y, x) \subseteq X^{\diamond\alpha}\},$$

$$\overline{E}(x)_\sigma(X^{\diamond\alpha}) = \{y \in U : D_E^+(y, x) \cap X^{\diamond\alpha} \neq \emptyset\}.$$

For the sake of simplicity, in the following only  $\underline{E}(x)_\sigma(X^{\leq\alpha})$  and  $\overline{E}(x)_\sigma(X^{\leq\alpha})$  with  $x \in X^{\leq\alpha}$  are considered. Of course, analogous considerations hold for  $\underline{E}(x)_\sigma(X^{<\alpha})$  and  $\overline{E}(x)_\sigma(X^{<\alpha})$ . Observe that the lower approximation of  $X^{\leq\alpha}$  with respect to  $x$  contains all the objects  $y \in U$  such that any object  $w$ , being similar to  $x$  at most as much as  $y$  is similar to  $x$  w.r.t. all the considered features  $E \subseteq F$ , also belongs to  $X^{\leq\alpha}$ . Thus, the data from the fuzzy pairwise information base  $B$  confirm that if  $w$  is similar to  $x$  not more than  $y \in \underline{E}(x)_\sigma(X^{\leq\alpha})$  is similar to  $x$  w.r.t. all the considered features  $E \subseteq F$ , then  $w$  belongs to  $X^{\leq\alpha}$ . In other words,  $x$  is a reference object and  $y \in \underline{E}(x)_\sigma(X^{\leq\alpha})$  is a limit object which belongs “certainly” to set  $X$  with credibility at most  $\alpha$ ; the limit is understood such that all objects  $w$  that are similar to  $x$  w.r.t. considered features at most as much as  $y$  is similar to  $x$ , also belong to  $X$  with credibility at most  $\alpha$ .

Analogously, the upper approximation of  $X^{\leq\alpha}$  with respect to  $x$  contains all the objects  $y \in U$  such that there is at least one object  $w$ , being similar to  $x$  at least as much as  $y$  is similar to  $x$  with respect to all the considered features  $E \subseteq F$ , which belongs to  $X^{\leq\alpha}$ . Thus, the data from the fuzzy pairwise information base  $B$  confirm that if  $w$  is similar to  $x$  not more than  $y \in \overline{E}(x)_\sigma(X^{\leq\alpha})$  is similar to  $x$  w.r.t. all the considered features  $E \subseteq F$ , then it is possible that  $w$  belongs to  $X^{\leq\alpha}$ . In other words,  $x$  is a reference object and  $y \in \overline{E}(x)_\sigma(X^{\leq\alpha})$  is a limit object which belongs “possibly” to set  $X$  with credibility at most  $\alpha$ ; the limit is understood such that all objects  $z \in U$  similar to  $x$  not more than  $y$  w.r.t. considered features, possibly belong to  $X^{\leq\alpha}$ .

Observe that the rough approximations  $\underline{E}(x)_\sigma(X^{\geq\alpha})$ ,  $\overline{E}(x)_\sigma(X^{\geq\alpha})$ ,  $\underline{E}(x)_\sigma(X^{\leq\alpha})$  and  $\overline{E}(x)_\sigma(X^{\leq\alpha})$  can be rewritten as

$$\begin{aligned} \underline{E}(x)_\sigma(X^{\geq\alpha}) &= \{y \in U : \forall w \in U, (w, x)D_E(y, x) \\ &\Rightarrow w \in X^{\geq\alpha}\}, \end{aligned}$$

$$\begin{aligned} \overline{E}(x)_\sigma(X^{\geq\alpha}) &= \{y \in U : \exists w \in U \text{ such that} \\ &(y, x)D_E(w, x) \text{ and } w \in X^{\geq\alpha}\}, \end{aligned}$$

$$\begin{aligned} \underline{E}(x)_\sigma(X^{\leq\alpha}) &= \{y \in U : \forall w \in U, (y, x)D_E(w, x) \\ &\Rightarrow w \in X^{\leq\alpha}\}, \end{aligned}$$

$$\begin{aligned} \overline{E}(x)_\sigma(X^{\leq\alpha}) &= \{y \in U : \exists w \in U \text{ such that} \\ &(w, x)D_E(y, x) \text{ and } w \in X^{\leq\alpha}\}. \end{aligned}$$

This formulation of the rough approximation is concordant with the syntax of the decision rules induced by means of DRSA from a fuzzy pairwise information base. More precisely,

- $\underline{E}(x)_\sigma(X^{\geq\alpha})$  is concordant with decision rules of the type: “if object  $w$  is similar to object  $x$  w.r.t. feature  $f_{i1}$  to degree at least  $h_{i1}$  and w.r.t. feature  $f_{i2}$  to degree at least  $h_{i2}$  and … and w.r.t. feature  $f_{ip}$  to degree at least  $h_{ip}$ , then object  $w$  belongs to set  $X$  to degree at least  $\alpha$ ”,
- $\overline{E}(x)_\sigma(X^{\geq\alpha})$  is concordant with decision rules of the type: “if object  $w$  is similar to object  $x$  w.r.t. feature  $f_{i1}$  to degree at least  $h_{i1}$  and w.r.t. feature  $f_{i2}$  to degree at least  $h_{i2}$  and … and w.r.t. feature  $f_{ip}$  to degree at least  $h_{ip}$ , then object  $w$  could belong to set  $X$  to degree at least  $\alpha$ ”,
- $\underline{E}(x)_\sigma(X^{\leq\alpha})$  is concordant with decision rules of the type: “if object  $w$  is similar to object  $x$  w.r.t. feature  $f_{i1}$  to degree at most  $h_{i1}$  and w.r.t. feature  $f_{i2}$  to degree at most  $h_{i2}$  and … and w.r.t. feature  $f_{ip}$  to degree at most  $h_{ip}$ , then object  $w$  belongs to set  $X$  to degree at most  $\alpha$ ”,
- $\overline{E}(x)_\sigma(X^{\leq\alpha})$  is concordant with decision rules of the type: “if object  $w$  is similar to object  $x$  w.r.t. feature  $f_{i1}$  to degree at most  $h_{i1}$  and w.r.t. feature  $f_{i2}$  to degree at most  $h_{i2}$  and … and w.r.t. feature  $f_{ip}$  to degree at most  $h_{ip}$ , then object  $w$  could belong to set  $X$  to degree at most  $\alpha$ ”, where  $\{i1, \dots, ip\} = E$  and  $h_{i1}, \dots, h_{ip} \in [0, 1]$ .

The above definitions of rough approximations and the syntax of decision rules are based on ordinal properties of similarity relations only. In fact, no algebraic operation, such as sum or product, involving cardinal properties of function  $\sigma$  measuring credibility of similarity relations is considered. This is an important characteristic of our approach in comparison with alternative approaches to case-based reasoning.

Let us remark that, similarly to DRSA approximation in an information base, in the case of DRSA approximation in a fuzzy pairwise information base, even if for two fuzzy sets  $X$  and  $Y$  it is  $X^{\geq\alpha} = Y^{\leq\alpha}$ , their approximations may be different due to the different directions of cutting the membership functions of sets  $X$  and  $Y$ .

Rough approximations in a fuzzy pairwise information base satisfy the following interesting properties.

**Theorem 1** Given a fuzzy pairwise information base  $B = \langle U, F, \sigma \rangle$  and a fuzzy set  $X$  in  $U$  with membership function  $\mu_X(\cdot)$ , the following properties hold for any  $E \subseteq F$ :

1. For any  $0 \leq \alpha \leq 1$ ,

$$\begin{aligned}\underline{E}(x)_\sigma(X^{\leq\alpha}) &\subseteq X^{\leq\alpha} \subseteq \overline{E}(x)_\sigma(X^{\leq\alpha}), \\ \underline{E}(x)_\sigma(X^{\geq\alpha}) &\subseteq X^{\geq\alpha} \subseteq \overline{E}(x)_\sigma(X^{\geq\alpha}), \\ \underline{E}(x)_\sigma(X^{<\alpha}) &\subseteq X^{<\alpha} \subseteq \overline{E}(x)_\sigma(X^{<\alpha}), \\ \underline{E}(x)_\sigma(X^{>\alpha}) &\subseteq X^{>\alpha} \subseteq \overline{E}(x)_\sigma(X^{>\alpha}).\end{aligned}$$

2. For any  $0 \leq \alpha \leq 1$ ,

$$\begin{aligned}\underline{E}(x)_\sigma(X^{\leq\alpha}) &= U - \overline{E}(x)_\sigma(X^{>\alpha}), \\ \underline{E}(x)_\sigma(X^{\geq\alpha}) &= U - \overline{E}(x)_\sigma(X^{<\alpha}).\end{aligned}$$

3. For any  $0 \leq \alpha \leq \beta \leq 1$ ,

$$\begin{aligned}\underline{E}(x)_\sigma(X^{\leq\alpha}) &\subseteq \underline{E}(x)_\sigma(X^{\leq\beta}), \\ \underline{E}(x)_\sigma(X^{<\alpha}) &\subseteq \underline{E}(x)_\sigma(X^{<\beta}), \\ \underline{E}(x)_\sigma(X^{\geq\alpha}) &\supseteq \underline{E}(x)_\sigma(X^{\geq\beta}), \\ \underline{E}(x)_\sigma(X^{>\alpha}) &\supseteq \underline{E}(x)_\sigma(X^{>\beta}), \\ \overline{E}(x)_\sigma(X^{\leq\alpha}) &\subseteq \overline{E}(x)_\sigma(X^{\leq\beta}), \\ \overline{E}(x)_\sigma(X^{<\alpha}) &\subseteq \overline{E}(x)_\sigma(X^{<\beta}), \\ \overline{E}(x)_\sigma(X^{\geq\alpha}) &\supseteq \overline{E}(x)_\sigma(X^{\geq\beta}), \\ \overline{E}(x)_\sigma(X^{>\alpha}) &\supseteq \overline{E}(x)_\sigma(X^{>\beta}).\end{aligned}$$

4. For any  $x, y, w, z \in U$  and for any  $0 \leq \alpha \leq 1$ ,

$$\begin{aligned}&[(y, x)D_E(w, x) \text{ and } w \in \underline{E}(x)_\sigma(X^{\geq\alpha})] \\ &\quad \Rightarrow y \in \underline{E}(x)_\sigma(X^{\geq\alpha}), \\ &[(y, x)D_E(w, x) \text{ and } w \in \underline{E}(x)_\sigma(X^{>\alpha})] \\ &\quad \Rightarrow y \in \underline{E}(x)_\sigma(X^{>\alpha}), \\ &[(y, x)D_E(w, x) \text{ and } w \in \overline{E}(x)_\sigma(X^{\geq\alpha})] \\ &\quad \Rightarrow y \in \overline{E}(x)_\sigma(X^{\geq\alpha}), \\ &[(y, x)D_E(w, x) \text{ and } w \in \overline{E}(x)_\sigma(X^{>\alpha})] \\ &\quad \Rightarrow y \in \overline{E}(x)_\sigma(X^{>\alpha}), \\ &[(w, x)D_E(y, x) \text{ and } w \in \underline{E}(x)_\sigma(X^{\leq\alpha})] \\ &\quad \Rightarrow y \in \underline{E}(x)_\sigma(X^{\leq\alpha}), \\ &[(w, x)D_E(y, x) \text{ and } w \in \underline{E}(x)_\sigma(X^{<\alpha})] \\ &\quad \Rightarrow y \in \underline{E}(x)_\sigma(X^{<\alpha}), \\ &[(w, x)D_E(y, x) \text{ and } w \in \overline{E}(x)_\sigma(X^{\leq\alpha})] \\ &\quad \Rightarrow y \in \overline{E}(x)_\sigma(X^{\leq\alpha}), \\ &[(w, x)D_E(y, x) \text{ and } w \in \overline{E}(x)_\sigma(X^{<\alpha})] \\ &\quad \Rightarrow y \in \overline{E}(x)_\sigma(X^{<\alpha}).\end{aligned}$$

5. For any  $E_1 \subseteq E_2 \subseteq F$  and for any  $0 \leq \alpha \leq 1$ ,

$$\underline{E}_1(x)_\sigma(X^{\leq\alpha}) \subseteq \underline{E}_2(x)_\sigma(X^{\leq\alpha}),$$

$$\underline{E}_1(x)_\sigma X^{<\alpha} \subseteq \underline{E}_2(x)_\sigma(X^{<\alpha}),$$

$$\underline{E}_1(x)_\sigma(X^{\geq\alpha}) \subseteq \underline{E}_2(x)_\sigma(X^{\geq\alpha}),$$

$$\underline{E}_1(x)_\sigma(X^{>\alpha}) \subseteq \underline{E}_2(x)_\sigma(X^{>\alpha}),$$

$$\overline{E}_1(x)_\sigma(X^{\leq\alpha}) \supseteq \overline{E}_2(x)_\sigma(X^{\leq\alpha}),$$

$$\overline{E}_1(x)_\sigma(X^{<\alpha}) \supseteq \overline{E}_2(x)_\sigma(X^{<\alpha}),$$

$$\overline{E}_1(x)_\sigma(X^{\geq\alpha}) \supseteq \overline{E}_2(x)_\sigma(X^{\geq\alpha}),$$

$$\overline{E}_1(x)_\sigma(X^{>\alpha}) \supseteq \overline{E}_2(x)_\sigma(X^{>\alpha}).$$

### An Algebraic Structure for Dominance-Based Rough Set Approach

This section presents an algebraic characterization of DRSA in terms of the bipolar complemented de Morgan Brouwer–Zadeh distributive lattice [30], being a generalization of the de Morgan Brouwer–Zadeh distributive lattice [2], already proposed to characterize the classical rough set approach [1].

A system  $\langle \Sigma, \Sigma^+, \Sigma^-, \wedge, \vee, '^+, '^-, \sim^+, \sim^-, 0, 1 \rangle$  is a *bipolar complemented quasi-Brouwer–Zadeh distributive lattice* if the following properties (1b)–(4b) hold:

- (1b)  $\Sigma$  is a distributive lattice with respect to the join and the meet operations  $\vee$  and  $\wedge$
- (1b')  $\Sigma^+, \Sigma^- \subseteq \Sigma$  are distributive lattices with respect to the join and the meet operations  $\vee$  and  $\wedge$ .  $\Sigma$  is bounded by the least element 0 and the greatest element 1, which implies that also  $\Sigma^+$  and  $\Sigma^-$  are bounded.
- (2b) The unary operations  $'^+ : \Sigma^+ \rightarrow \Sigma^-$  and  $'^- : \Sigma^- \rightarrow \Sigma^+$  are Kleene (also Zadeh or fuzzy) bipolar complementations, that is for arbitrary  $a, b \in \Sigma^+$  and  $c, d \in \Sigma^-$

$$(K1b) \quad a'^+{}' = a, \quad c'^-{}' = c,$$

$$(K2b) \quad (a \vee b)'^+ = a'^+ \wedge b'^+,$$

$$(c \vee d)'^- = c'^- \wedge d'^-,$$

$$(K3b) \quad a \wedge a'^+ \leq b \vee b'^+, \quad c \wedge c'^- \leq d \vee d'^-.$$

- (3b) The unary operations  $\sim^+ : \Sigma^+ \rightarrow \Sigma^-$  and  $\sim^- : \Sigma^- \rightarrow \Sigma^+$  are Brouwer (or intuitionistic) bipolar complementations, that is for arbitrary  $a, b \in \Sigma^+$  and  $c, d \in \Sigma^-$

$$(B1b) \quad a \wedge a^{\sim^+ \sim^-} = a, \quad c \wedge c^{\sim^- \sim^+} = c$$

$$(B2b) \quad (a \vee b)^{\sim^+} = a^{\sim^+} \wedge b^{\sim^+},$$

$$(c \vee d)^{\sim^-} = c^{\sim^-} \wedge d^{\sim^-},$$

$$(B3b) \quad a \wedge a^{\sim^+} = 0, \quad c \wedge c^{\sim^-} = 0.$$

- (4b) Complementation  $'^+$  and complementation  $\sim^+$  in one hand and complementation  $'^-$  and complementation  $\sim^-$  in the other hand are linked by the interconnection rule, that is, for arbitrary  $a \in \Sigma^+$  and arbitrary  $b \in \Sigma^-$ :  $(in-b) a^{\sim^+} \leq a'^+, b^{\sim^-} \leq b'^-$ .

A structure  $\langle \Sigma, \Sigma^+, \Sigma^-, \wedge, \vee, '^+, '^-, \sim^+, \sim^-, 0, 1 \rangle$  is a *bipolar complemented Brouwer–Zadeh distributive lattice* if it is a quasi-Brouwer–Zadeh distributive lattice satisfying the stronger interconnection rule, that is, for arbitrary  $a \in \Sigma^+$  and arbitrary  $b \in \Sigma^-$ :  $(s-in-b) a^{\sim^+ \sim^-} = a'^+ \wedge b'^-, b^{\sim^- \sim^+} = b'^- \wedge a'^+$ .

A bipolar complemented Brouwer–Zadeh distributive lattice is a *bipolar complemented de Morgan Brouwer–Zadeh distributive lattice*, if it satisfies also the  $\vee$  de Morgan property that is, for arbitrary  $a, b \in \Sigma^+$  and  $c, d \in \Sigma^-$ :

$$(B2a-b) \quad (a \wedge b)^{\sim^+} = a^{\sim^+} \vee b^{\sim^+}$$

$$(c \wedge d)^{\sim^-} = c^{\sim^-} \vee d^{\sim^-}.$$

The bipolar complemented de Morgan Brouwer–Zadeh distributive lattice is an algebraic structure which can be given to the collection of all rough approximations within the dominance-based rough set approach as follows. Let us consider a fuzzy information base  $B = \langle U, F, \varphi \rangle$ . Given  $G \subseteq F$ , for each  $X \subseteq U$ , its *upward lower approximation*  $\underline{G}^{(>)}(X)$  and its *upward upper approximation*  $\overline{G}^{(>)}(X)$  can be defined as:

$$\underline{G}^{(>)}(X) = \{x \in U : D_G^+(x) \subseteq X\},$$

$$\overline{G}^{(>)}(X) = \{x \in U : D_G^-(x) \cap X \neq \emptyset\}.$$

Analogously, given  $G \subseteq F$ , for each  $X \subseteq U$ , its *downward lower approximation*  $\underline{G}^{(<)}(X)$  and its *downward upper approximation*  $\overline{G}^{(<)}(X)$  can be defined as:

$$\underline{G}^{(<)}(X) = \{x \in U : D_G^-(x) \subseteq X\},$$

$$\overline{G}^{(<)}(X) = \{x \in U : D_G^+(x) \cap X \neq \emptyset\}.$$

Fixed  $G \subseteq F$ , for any  $X \subseteq U$  let us consider the pairs  $\langle \underline{G}^{(<)}(X), U - \overline{G}^{(<)}(X) \rangle$  and  $\langle \underline{G}^{(>)}(X), U - \overline{G}^{(>)}(X) \rangle$  and the sets

$$B = \{(I, E) : I, E \subseteq U \text{ such that } I \cap E = \emptyset\},$$

$$B^- = \{(I, E) : \exists X \subseteq U \text{ for which } I = \underline{G}^{(<)}(X) \text{ and } E = U - \overline{G}^{(<)}(X)\},$$

$$B^+ = \{(I, E) : \exists X \subseteq U \text{ for which } I = \underline{G}^{(>)}(X) \text{ and } E = U - \overline{G}^{(>)}(X)\}.$$

The following result [30] holds.

**Theorem 2** *The structure  $\langle B, B^+, B^-, \sqcap, \sqcup, \neg\neg, \neg^+, \approx^-, \approx^+, \langle \emptyset, U \rangle, \langle U, \emptyset \rangle \rangle$  where, for any  $\langle I_1, E_1 \rangle, \langle I_2, E_2 \rangle \in B$ ,  $\langle I_3, E_3 \rangle \in B^-$ ,  $\langle I_4, E_4 \rangle \in B^+$ ,*

$$\langle I_1, E_1 \rangle \sqcap \langle I_2, E_2 \rangle = \langle I_1 \cap I_2, E_1 \cup E_2 \rangle$$

$$\langle I_1, E_1 \rangle \sqcup \langle I_2, E_2 \rangle = \langle I_1 \cup I_2, E_1 \cap E_2 \rangle$$

$$\langle I_3, E_3 \rangle \neg\neg = \langle E_3, I_3 \rangle, \langle I_4, E_4 \rangle \neg^+ = \langle E_4, I_4 \rangle$$

$$\langle I_3, E_3 \rangle \approx^- = \langle E_3, U - E_3 \rangle, \langle I_4, E_4 \rangle \approx^+ = \langle E_4, U - E_4 \rangle$$

is a bipolar complemented de Morgan Brouwer-Zadeh distributive lattice.

Let us observe that within the bipolar complemented de Morgan Brouwer-Zadeh distributive lattice,  $\Sigma^+$  has to be interpreted as the set of “positive” concepts, while  $\Sigma^-$  as the set of “negative” concepts. For example, in a problem of evaluation of students, the concept of “good students” is “positive”, while the concept of “bad students” is “negative”. Within DRSA, each concept is represented by the pair  $(I, E)$ , where  $I$  (the interior) is the lower approximation of a set  $X \subseteq U$  and  $E$  (the exterior) is the complement in  $U$  of the upper approximation of  $X$ . Intuitively, each concept is represented by  $I$ , being the set of objects that surely belong to the concept, and  $E$  being the set of objects that surely does not belong to the concept. Therefore, the positive concept “good students” is represented by the pair  $(I_G, E_G)$ , where  $I_G$  represents the set of students “surely good” and  $E_G$  represents the set of students “surely not good”. Analogously, the negative concept “bad students” is represented by the pair  $(I_B, E_B)$ , where  $I_B$  represents the set of students “surely bad” and  $E_B$  represents the set of students “surely not bad”. To illustrate the concept of join  $\sqcup$  and the meet  $\sqcap$ , let us consider the concepts of “good students in Mathematics”, represented by the pair  $(I_{GM}, E_{GM})$ , and “good students in Literature”, represented by the pair  $(I_{GL}, E_{GL})$ . In this context, the concept of “good students in Mathematics or Literature” is represented by

$$(I_{GM}, E_{GM}) \sqcup (I_{GL}, E_{GL}) = (I_{GM} \cup I_{GL}, E_{GM} \cap E_{GL}).$$

This means that to the concept of “good students in Mathematics or Literature” surely belongs the set of students “surely good in Mathematics” or “surely good in Literature”, i.e.  $I_{GM} \cup I_{GL}$ , while surely does not belong the set of students “surely not good in Mathematics” and “surely

not good in Literature”, i.e.  $E_{GM} \cap E_{GL}$ . Analogously, the concept of “good students in Mathematics and Literature” is represented by

$$(I_{GM}, E_{GM}) \sqcap (I_{GL}, E_{GL}) = (I_{GM} \cap I_{GL}, E_{GM} \cup E_{GL}).$$

This means that to the concept of “good students in Mathematics and Literature” surely belongs the set of students “surely good in Mathematics” and “surely good in Literature”, i.e.  $I_{GM} \cap I_{GL}$ , while surely does not belong the set of students “surely not good in Mathematics” or “surely not good in Literature”, i.e.  $E_{GM} \cup E_{GL}$ .

To the concept of “good students”  $(I_G, E_G)$ , the Kleene bipolar complementation  $\neg^+$  associates the set  $(E_G, I_G)$ , that is to the negation of the concept of “good students” surely belongs the set of student “surely not good”  $E_G$ , and surely does not belong the set of all the students “surely good”  $I_G$ . Considering the Brouwer bipolar complementation  $\approx^+$ , the negation of the concept of “good students”  $(I_G, E_G)$  is given by  $(E_G, U - E_G)$ , that is to the negation of the concept of “good students” surely belongs the set of students “surely not good”  $E_G$ , as in the case of the Kleene bipolar complementation, and surely does not belong the set of all the other students  $U - E_G$ , differently from the case of the Kleene bipolar complementation. Similar considerations hold for the Kleene bipolar complementation  $\neg\neg$  and the Brouwer bipolar complementation  $\approx^-$  of negative concepts.

## Conclusions

This article provides arguments for the claim that the dominance-based rough set approach is a proper way of handling monotonically ordered data in granular computing. Referring to some ideas of Leibniz, Frege, Boole, and Łukasiewicz, DRSA represents fundamental concepts of rough set theory in terms of a generalization that takes into account ordinal properties of data, permitting us to deal with the graduality of fuzzy sets. DRSA in the context of ordinal classification, its fuzzy extension, and a rough probabilistic model of DRSA (variable consistency-dominance-based rough set approach) have been presented. Moreover, dominance-based rough approximation of a fuzzy set has been discussed, which infers the most cautious conclusions from available imprecise information; different to almost all known fuzzy rough set approaches, the dominance-based rough approximation of a fuzzy set does not require any fuzzy connective which is always arbitrary to some extent.

Knowledge induced from dominance-based rough approximations of fuzzy sets, or more generally, from ordi-

nal data, is represented in terms of gradual decision rules. The dominance-based rough approximations of fuzzy sets generalize the classical rough approximations of crisp sets, as proved by showing that the classical rough set approach is one of its particular cases. Because of considering only ordinal character of the graduality of fuzzy sets, and due to eliminating all fuzzy connectives, the dominance-based rough approximations of fuzzy sets give a new insight into both rough sets and fuzzy sets, and enable further generalizations of both of them. The recently proposed DRSA for fuzzy case-based reasoning is an example of this capacity. This article also exhibits some important merits of DRSA within granular computing, which can be summarized as follows:

- DRSA extends the paradigm of granular computing to problems involving ordered data,
- It specifies a syntax and modality of information granules, defined by means of dominance-based constraints, which are appropriate for dealing with ordered data,
- It provides a methodology for dealing with this type of information granules, which results in a theory of computing with words and reasoning about ordered data,
- It is supported by a robust and meaningful algebraic model, the bipolar complemented de Morgan Brouwer–Zadeh distributive lattice, and this ensures the solidity of the obtained results.

## Future Directions

Granular computing with ordered data is a very general problem, because also other modalities of information constraints, such as veristic, possibilistic, and probabilistic modalities, have to deal with ordered value sets (with qualifiers relative to grades of truth, possibility, and probability). For this reason, granular computing with ordered data based on DRSA is a very promising research field. There is a great potential in both theoretical investigations, such as extension of DRSA to other algebraic models than the Brouwer–Zadeh lattice (for a general survey of algebraic structures for rough set theory see Chap. 12 in [50]), and practical applications, such as customization of DRSA to case-based reasoning for specific domains, like medical diagnosis or computer security.

## Bibliography

1. Cattaneo G, Ciucci D (2004) Algebraic structures for rough sets. In: Transaction on rough sets II. LNCS, vol 3135. Springer, Berlin, pp 208–252
2. Cattaneo G, Giuntini R, Pilla R (1999)  $BZMV^{dM}$  algebras and stonian MV-algebras, (applications to fuzzy sets and rough approximations). *Fuzzy Sets Syst* 108:201–222
3. Dubois D, Prade H (1990) Rough fuzzy sets and fuzzy rough sets. *Int J General Syst* 17:191–209
4. Dubois D, Prade H (1992) Gradual inference rules in approximate reasoning. *Inf Sci* 61:103–122
5. Dubois D, Prade H (1992) Putting rough sets and fuzzy sets together. In: Słowiński R (ed) Intelligent decision support – Handbook of applications and advances of the rough sets theory. Kluwer, Dordrecht pp 203–232
6. Dubois D, Prade H, Esteva F, Garcia P, Godo L, Lopez de Mantara R (1998) Fuzzy set modelling in case-based reasoning. *Int J Intell Syst* 13:345–373
7. Dubois D, Grzymala-Busse J, Inuiguchi M, Polkowski L (eds) (2004) Transactions on rough sets II: Rough sets and fuzzy sets. LNCS, vol 3135. Springer, Berlin
8. Dyer J (2005) MAUT – Multiattribute utility theory. In: Figueira J, Greco S, Ehrgott M (eds) Multiple criteria decision analysis: State of the art surveys. Springer, Berlin, pp 266–294
9. Figueira J, Greco S, Ehrgott M (eds) (2005) Multiple criteria decision analysis: State of the art surveys. Springer, Berlin
10. Fodor J, Roubens M (1994) Fuzzy preference modelling and multicriteria decision support. Kluwer, Dordrecht
11. Fortemps P, Greco S, Słowiński R (2008) Multicriteria decision support using rules that represent rough-graded preference relations. *Eur J Operational Res* 188:206–223
12. Gilboa I, Schmeidler D (2001) A theory of case-based decisions. Cambridge University Press, Cambridge
13. Ginsburg S, Hull R (1983) Order dependency in the relational model. *Theor Comput Sci* 26:149–195
14. Greco S, Inuiguchi M, Słowiński R (2002) Dominance-based rough set approach using possibility and necessity measures. In: Alpigini JJ, Peters JF, Skowron A, Zhong N (eds) Rough sets and current trends in computing. LNAI, vol 2475. Springer, Berlin, pp 85–92
15. Greco S, Inuiguchi M, Słowiński R (2004) A new proposal for rough fuzzy approximations and decision rule representation. In: Dubois D, Grzymala-Busse J, Inuiguchi M, Polkowski L (eds) Transactions on rough sets II: Rough sets and fuzzy sets. LNCS, vol 3135. Springer, Berlin, pp 156–164
16. Greco S, Inuiguchi M, Słowiński R (2006) Fuzzy rough sets and multiple-premise gradual decision rules. *Int J Approx Reason* 41:179–211
17. Greco S, Matarazzo B, Słowiński R (1999) The use of rough sets and fuzzy sets in MCDM. In: Gal T, Stewart T, Hanne T (eds) Advances in multiple criteria decision making. Kluwer, Boston, pp 14.1–14.59
18. Greco S, Matarazzo B, Słowiński R (2000) Rough set processing of vague information using fuzzy similarity relations. In: Calude C, Paun G (eds) From finite to infinite. Springer, Berlin, pp 149–173
19. Greco S, Matarazzo B, Słowiński R (2000) A fuzzy extension of the rough set approach to multicriteria and multiattribute sorting. In: Fodor J, De Baets B, Perny P (eds) Preferences and decisions under incomplete information. Physica, Heidelberg, pp 131–154
20. Greco S, Matarazzo B, Słowiński R (2001) Rough sets theory for multicriteria decision analysis. *Eur J Operational Res* 129:1–47

21. Greco S, Matarazzo B, Słowiński R (2001) Rough set approach to decisions under risk. In: Ziarko W, Yao Y (eds) *Rough sets and current trends in computing*. LNAI, vol 2005. Springer, Berlin, pp 160–169
22. Greco S, Matarazzo B, Słowiński R (2002) Preference representation by means of conjoint measurement and decision rule model. In: Bouyssou D, Jacquet-Lagréze E, Perny P, Słowiński R, Vanderpooten D, Vincke P (eds) *Aiding decisions with multiple criteria – Essays in Honor of Bernard Roy*. Kluwer, Dordrecht, pp 263–313
23. Greco S, Matarazzo B, Słowiński R (2004) Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *Eur J Operational Res* 158:271–292
24. Greco S, Matarazzo B, Słowiński R (2004) Dominance-based rough set approach to knowledge discovery (I) – General perspective. In: Zhong N, Liu J (eds) *Intelligent technologies for information analysis*. Springer, Berlin, pp 513–552
25. Greco S, Matarazzo B, Słowiński R (2004) Dominance-based rough set approach to knowledge discovery (II) – Extensions and applications. In: Zhong N, Liu J (eds) *Intelligent technologies for information analysis*. Springer, Berlin, pp 553–612
26. Greco S, Matarazzo B, Słowiński R (2005) Decision rule approach. In: Figueira J, Greco S, Ehrgott M (eds) *Multiple criteria decision analysis: State of the art surveys*. Springer, Berlin, pp 507–563
27. Greco S, Matarazzo B, Słowiński R (2005) Generalizing rough set theory through dominance-based rough set approach. In: Slezak D, Yao J, Peters J, Ziarko W, Hu X (eds) *Rough sets, fuzzy sets, data mining, and granular computing*. LNAI, vol 3642. Springer, Berlin, pp 1–11
28. Greco S, Matarazzo B, Słowiński R (2006) Dominance-based rough set approach to case-based reasoning. In: Torra V, Narukawa Y, Valls A, Domingo-Ferrer J (eds) *Modelling decisions for artificial intelligence*. LNAI, vol 3885. Springer, Berlin, pp 7–18
29. Greco S, Matarazzo B, Słowiński R (2007) Dominance-based rough set approach as a proper way of handling graduality in rough set theory. In: *Transactions on rough sets VII*. LNAI, vol 4400. Springer, Berlin, pp 36–52
30. Greco S, Matarazzo B, Słowiński R (2008) An algebraic structure for dominance-based rough set approach. In: Proc. 3rd Int Conference on rough sets and knowledge technology (RSKT 2008), LNAI. Springer, Berlin, pp 252–259
31. Greco S, Matarazzo B, Słowiński R, Stefanowski J (2001) Variable consistency model of dominance-based rough set approach. In: Ziarko W, Yao Y (eds) *Rough sets and current trends in computing*. LNAI, vol 2005. Springer, Berlin, pp 170–181
32. Greco S, Predki B, Słowiński R (2002) Searching for an equivalence between decision rules and concordance-discordance preference model in multicriteria choice problems. *Control Cybern* 31:921–935
33. Hume D (1748) *An enquiry concerning human understanding*. Oxford, Clarendon Press
34. Klement EP, Mesiar R, Pap E (2000) *Triangular norms*. Kluwer, Dordrecht
35. Kolodner J (1993) Case-based reasoning. Morgan Kaufmann, San Mateo
36. Leake DB (1996) CBR in context: the present and future. In: Leake D (ed) *Case-based reasoning: Experiences, lessons, and future directions*. AAAI Press/MIT Press, Menlo Park, pp 1–30
37. Lin TY (1988) Neighborhood systems and relational databases. In: *Proceedings of the ACM Conference on Computer Science*, Atlanta, p 725
38. Lin TY (1989) Neighborhood systems and approximation in database and knowledge base systems. In: *Proceedings of the Fourth International Symposium on Methodologies of Intelligent Systems, Poster Session, October 12–15*, pp 75–86
39. Lin TY (1992) Topological and fuzzy rough sets. In: Słowiński R (ed) *Intelligent decision support – Handbook of application and advances of the rough sets theory*. Kluwer, Dordrecht, pp 287–304
40. Lin TY (1997) Granular computing. *Announcement of the BISC Special Interest Group on Granular Computing*
41. Lin TY (1998) Granular computing on binary relations I: Data mining and neighborhood systems. In: Skowron A, Polkowski L (eds) *Rough sets in knowledge discovery*. Physica, Heidelberg, pp 107–121
42. Lin TY (1998) Granular computing on binary relations II: Rough set representations and belief functions. In: Skowron A, Polkowski L (eds) *Rough sets in knowledge discovery*. Physica, Heidelberg, pp 121–140
43. Loemker L (ed and trans), Leibniz GW (1969) *Philosophical papers and letters*, 2nd edn. Reidel, Dordrecht
44. Nakamura A, Gao JM (1991) A logic for fuzzy data analysis. *Fuzzy Sets Syst* 39:127–132
45. Pal SK, Skowron A (eds) (1999) *Rough-fuzzy hybridization: A new trends in decision making*. Springer, Singapore
46. Pawlak Z (1982) Rough sets. *Int J Comput Inf Sci* 11:341–356
47. Pawlak Z (1991) Rough sets. Kluwer, Dordrecht
48. Pawlak Z (2001) Rough set theory. *Künstliche Intelligenz* 3: 38–39
49. Peters JF, Skowron A, Dubois D, Grzymala-Busse J, Inuiguchi M, Polkowski L (eds) (2005) *Rough sets and fuzzy sets, transaction on rough sets II*. Springer, Berlin
50. Polkowski L (2002) Rough set: mathematical foundations. *Physica*, Heidelberg
51. Radzikowska AM, Kerre EE (2002) A comparative study of fuzzy rough sets. *Fuzzy Sets Syst* 126:137–155
52. Słowiński R, Greco S, Matarazzo B (2002) Axiomatization of utility, outranking and decision-rule preference models for multiple-criteria classification problems under partial inconsistency with the dominance principle. *Control Cybern* 31:1005–1035
53. Słowiński R, Greco S, Matarazzo B (2002) Mining decision-rule preference model from rough approximation of preference relation. In: Proc. 26th IEEE Annual Int. Conference on Computer Software & Applications (COMPSAC 2002), Oxford, pp 1129–1134
54. Słowiński R, Greco S, Matarazzo B (2005) Rough set based decision support. In: Burke EK, Kendall G (eds) *Search methodologies: Introductory tutorials in optimization and decision support techniques*. Springer, New York, pp 475–527
55. Stewart T (2005) Dealing with uncertainties in MCDA. In: Figueira J, Greco S, Ehrgott M (eds) *Multiple criteria decision analysis: State of the art surveys*. Springer, Berlin, pp 445–470
56. Zadeh LA (1965) Fuzzy sets. *Inf Control* 8:338–353
57. Zadeh LA (1979) Fuzzy sets and information granularity. In: Gupta M, Ragade RK, Yager RR (eds) *Advances in fuzzy set theory and applications*. North-Holland, Amsterdam, pp 3–18

58. Zadeh LA (1996) Key roles of information granulation and fuzzy logic in human reasoning, concept formulation and computing with words. In: Proceedings of the 5th IEEE International Conference on Fuzzy Systems, New Orleans, p 1
59. Zadeh LA (1997) Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets Syst* 90:111–127
60. Zadeh LA (1999) From computing with numbers to computing with words – from manipulation of measurements to manipulation of perception. *IEEE Trans Circuits Syst – I: Fundament Theor Appl* 45:105–119
61. Ziarko W (1993) Variable precision rough sets model. *J Comput Syst Sci* 46:39–59
62. Ziarko W (1998) Rough sets as a methodology for data mining. In: Polkowski L, Skowron A (eds) *Rough Sets in Knowledge Discovery*, vol 1. Physica, Heidelberg, pp 554–576

## Granular Computing, Information Models for

STEVEN A. DEMURJIAN

Department of Computer Science & Engineering,  
The University of Connecticut, Storrs, USA

### Article Outline

- Glossary
- Definition of the Subject
- Introduction
- Candidate Information Models
- Suitability of Information Models for Granular Computing
- Future Directions
- Bibliography

### Glossary

**Attributed-based data model (ABDM)** ABDM is a data (information) model that organizes data (information) in attribute-value pairs, providing a means for both definition and access.

**Data (information) model** A data (information) model is used in database design to capture the structure or schema of the data. All data (information) that is entered into a database must conform to the definitions of the data (information) model.

**Data (information) table** A set of similar granules used in granular computing are collected into a data (information) table in order to allow them to be conceptualized and reasoned on in a formal manner.

**Extensible markup language (XML)** XML is a data (information) model that is a standard for exchanging

and sharing data (information) across the internet, among databases, and so on.

**Functional data model (FDM)** FDM is a semantic data (information) model to represent data (information) as it appears in the “real-world” with a functional/logic basis.

**Granular computing** A discipline of information theory/computer science that uses a formal theory to reason about and analyze data (information) in granules.

**Granule** A piece of data (information) of varied size and complexity that is used to represent data (information) as it occurs in some “real-world” context.

**Relational data model (RDM)** RDM is a dominant data (information) model in commercial and open source database management systems with a basis in set theory.

### Definition of the Subject

Granular computing (GrC) is an emerging discipline of information theory that strives to allow reasoning and analysis based on varying levels of information granularity (from fine to coarse). In GrC, the entire information universe can be organized based on many different criteria, allowing the information to be abstracted, aggregated, classified, generalized, and so on, based on various characteristics (e.g., data similarity, operational usage, etc.). As a result, GrC relies on information models to describe the universe, the elements of the universe, and the composition of each element. In this chapter, alternative candidate information models for GrC are explored, including: the attribute-based data model, the Relational data model, the functional data model, and the extensible markup-language (XML). This includes both a description of these models and an analysis of the suitability in support of GrC.

### Introduction

In information theory, one approach to support the reasoning and analysis of information based on varied levels of conceptualization is the emerging discipline of *Granular computing* (GrC) [10,13,14,26], a term jointly coined by Lin and Zadeh and having a basis in research conducted by Lin on neighborhoods (e.g., granules) in databases [7] and computer security [8]. Information granulation is an approach that partitions the entire information universe into *granules* [11,12,24,25]. This partitioning can be based on many different criteria: aggregating objects that demonstrate similarity of features or usage by classification (e.g., defining a relation such as a Course table in a university application); abstracting away details via generalization (e.g., creating an object-oriented class Person that would

be the root of a hierarchy); distinguishing among different characteristics of objects by specialization (e.g., Students or Faculty sub-classes of Person); summarizing the characteristics of similar objects into a coarser conceptualization (e.g., collecting multiple Computer Science Faculty instances into a single object); and so on. In all of these situations, the information universe is described in its entirety, and from that initial characterization, it is possible to define elements and components of elements.

In support of GrC, one of the key initial considerations is to define the information universe via an appropriate data representation. For GrC, this data representation has most frequently been based on data tables as formally defined for rough sets [15], where a *data table* has many different equivalent nomenclatures, including: a knowledge representation system, an attribute-value system, and an information table. For a data table, columns are labeled as attributes (of the objects) and rows are labeled as instances (of the objects). Formally, the data table is characterized as a pair that contains a finite non-empty universe  $U$  and a finite non-empty set of primitive attributed  $A$ . Additionally, for each  $a$  in  $A$ , there is a non-empty set of values, and a function that maps the universe  $U$  to these values. Collectively, these data tables based on rough sets have been used as the basis to define foundational models for granular computing [13,14,18] and to achieve privacy protection for medical data by partitioning attributes into identifying, easily known, and sensitive categories [20].

The main objective of this chapter is to explore alternative candidate information models for GrC to support the data table (and dependencies that may exist both within a table and across multiple data tables). Specifically, a variety of models that are old and new are explored, including: the attribute-based data model [6], the relational data model [1], the functional data model [17], and the extensible markup-language [22]. The *attribute-based data model (ABDM)*, proposed in the early 1970, was touted for its ability to formally define “real-world” records of information, to support their retrieval, and has been shown capable of capturing data from relational, hierarchical, network, and functional data models [3,4]. The *relational data model (RDM)* proposed in mid-1970 along with the current SQL 1992 (SQL2) has evolved into the dominant database model for commercial and open source data base systems, with extensions to support object-oriented and other features underway (SQL3). The *functional data model (FDM)* and its associated Daplex programming language, proposed in 1981, offers unique programming-like features and its functional formal underpinnings can provide a means to reason about data and constraints, and moreover, to capture semantics. Fi-

nally, the *extensible markup-language (XML)* has emerged as an information-representation standard format, allowing information to be modeled and more easily exchanged among programs, web-sites, databases, etc. All of these models offer different capabilities in support of data tables and GrC.

The remainder of this chapter contains two sections and a conclusion. In Sect. “[Candidate Information Models](#)”, the four information models, ADBM, RDM, FDM, and XML, are examined. Using this as a basis, in Sect. “[Suitability of Information Models for Granular Computing](#)”, the suitability of the four information models in support of GrC is investigated by considering them against three different criteria. Finally, Sect. “[Future Directions](#)” offers concluding remarks.

## Candidate Information Models

In this section, background on the four information models to be considered for their suitability for GrC are reviewed, namely: the attribute-based data model (ABDM) [6], the relational data model (RDM) [1], the functional data model (FDM) [17], and the extensible markup-language (XML) [22]. To serve as a common context for the discussion, a university database is utilized as a example. In the university database, faculty are tracked by name, identifier, office phone, and department (e.g., computer science, mathematics, English, etc.), and students are tracked by name, identifier, grade point average, and campus address (e.g., dormitory). For each course in the catalog, there is a unique course number (e.g., a combination of department and number such as CS123, MATH233, etc.), the course title and description, and the list of prerequisites for the course. The courses offered are also tracked, by course number/section number pairs (e.g., courses may be offered in multiple sections), and the FacultyID teaching each offering (by faculty identifier), and the term (e.g., semester, quarter, etc.) of the offering (e.g., Fall2007, Spring2008, Summer2008, etc.). Likewise, the students enrolled by the course number/section number pair for each term are tracked.

## The Relational Data Model (RDM)

Relational data is organized into tuples of *relations*. A database is a collection of relations. The attributes of a relation are distinct. The tuples of a relation have the property that no two tuples are identical. In addition, one or more attributes of the relation may be defined as the *primary key* of the relation; a relation can have multiple candidate keys, one of which is chosen as the primary key. The primary key of the relation is used to uniquely identify

the tuples of the relation. To establish dependencies across two or more relations, a *foreign key* can be defined consisting of one or more attributes of one relation that reference the primary key of another relation. Foreign keys are employed to establish referential integrity constraints across two relations where the foreign key of one relation references the primary key of another relation (with self-reference allowed). Non-primary key attributes are allowed to have null values; this includes a foreign key. A definition of the University database in tabular (or relational) form is as below:

```

Student (Name, StudentID*, GPA, CampusAddress)
Faculty (Name, FacultyID*, Ophone, Department)
Courses (Course#*, Title, Description,
          PCourse#*)
OfferedCourses (Course#*, Section#*, 
                  FacultyID*, Term*)
EnrolledStudents (Course#*, Section#*, 
                  StudentID*, Term*)

```

Briefly, let us describe the data relationships of this database. First, each student of the Student relation is uniquely identified by a StudentID (primary key), has a Name, a grade point average (GPA), and a CampusAddress. Next, each faculty member of the Faculty relation also has a unique FacultyID (primary key), a Name, an office phone number, and a Department affiliation. Third, each course of the Courses relation (catalog) is uniquely identified by a Course# (primary key), has a Title, a Description and may have zero or more prerequisite courses (PCourse#-foreign key). Fourth, the OfferedCourses relation tracks the courses offered by sections taught by faculty each term, with these four attributes forming a compound primary key. Fifth, the EnrolledStudents relation tracks the students enrolled in sections of courses for each term, again by a compound primary key. In terms of referential integrity, the foreign key PCourse# self references the relation Course; note that a course without a prerequisite will have a null foreign key.

### The Attribute-Based Data Model (ABDM)

In ABDM, the initial step in defining a database is the specification of a collection of the attributes that are in the database. The *database records* in ADBM consist of sets of attribute-value pairs. An *attribute-value pair* is a member of the Cartesian product of the attribute name and the value domain of the attribute. As an example, <Department, Computer Science> is an attribute-value pair having Computer Science as the value for the Depart-

ment attribute. Database records have the property that for a given set of attribute-value pairs, no two pairs in the set have the same attribute name. In ADBM, data is considered in the following constructs: database, file, record, attribute-value pair (keyword) attribute-value range, record body, directory, directory keyword and non-directory keyword. Informally, a *database* consists of a collection of files. Each *file* contains a group of records which are characterized by a unique set of keywords. A *record* is composed of two parts. The first part is a collection of *attribute-value pairs (keywords)*. The second part of the record, which is optional, is for unformatted textual information, and is referred to as the *record body*. A record contains at most one attribute-value pair for each attribute defined in the database. An example of a record equivalent to the Faculty relation is:

```

(<File, Faculty>, <Record Number, 123>,
<Name, John Smith>, <FacultyID, 12121212>,
<Ophone, 5551111>, <Department, Computer
Science>,
{This is an assistant professor up for
tenure in 2012})

```

The angle brackets, <, >, enclose an attribute-value pair, i.e., keyword. The curly brackets, {, }, include the record body. The record is enclosed in parentheses. The first attribute-value pair of all records of a file, by convention, is the same. In particular, the attribute is File and the value is the file name (akin to a relation name in RDM). Similar records for Student, Courses, CourseOfferings, and EnrolledStudents can also be defined; referential integrity is achieved by using the unique record numbers that are given to every instance.

Finally, in ABDM, the indexing criteria for a given database is also a critical part of the definition of the database. In particular, for indexing purposes, there are two different types of attribute-value pairs in a record (or a file), with all of the indexing data maintained in the *directory* of the database. Certain attribute-value pairs of a record (or a file) are called the *directory keywords* of the record (file), because either the attribute-value pairs or their attribute-value ranges are kept in a directory for identifying the records (files). Those attribute-value pairs which are not kept in the directory are called *non-directory keywords*. In the example record above, the File, Record Number, and FacultyID attributes would be directory keywords, while the other attributes would be non-directory keywords. The identification of database records are by either directory or non-directory keywords. When

directory keywords are used, the search space is clearly defined using the indexing criteria. When non-directory keywords are used, the entire file of records must be searched. Note that null values are not allowed for directory attributes.

### The Functional Data Model (FDM)

An *entity* in a functional database is always associated with a collection of distinct functions that can be applied to the entity to return either individual data values or one or more objects. The term *object* is used to refer to the actual data values (values for the functions) for an entity. Thus, an object can be considered as an instantiation (occurrence, in the earlier terminology) of the entity. An object is analogous to a tuple in the RDM or a record in ABDM. The *functions* of an entity are applied to the entity to return a particular value that is associated with that entity, i.e., to return a portion of the object. There are two types of functions, scalar-valued functions and entity-valued functions. Each type of function may be either single-valued (returning one value) or set-valued (returning zero or more values). *Scalar-valued functions* return one or more typical database values (i.e., string, integer, and float). *Entity-valued functions* return one or more entity objects as their values.

Additionally, in the FDM, relationships between the entities may be defined that lead to one or more generalization hierarchies for the database. Generalization [19], is an abstraction technique that is used to organize commonalities from multiple entities into a single entity (e.g., the name and identifier for students and faculty can be organized in a person entity). Because of the presence of generalization hierarchies, there are two different categories of entities, entity subtypes and entity supertypes. An *entity subtype* exists at the inner and leaf nodes of a generalization hierarchy, and inherits all of the characteristics of its ancestors, i.e., inherits all of the functions of its ancestors. An *entity supertype*, or, more simply, an *entity type*, is at the root of a generalization hierarchy and has the property that no two instantiations (occurrences) of the entity type return the same values for all of the entity's functions. Additionally, one or more single-valued, scalar-valued functions in an entity type may be defined to be the key of the entity. Functional data is organized into *generalization hierarchies of entities*. A database is a collection of entities organized into generalization hierarchies. References among entities are accomplished when a function is entity valued, with self-referencing possible. In the following, the definition of the FDM version of the University database using the Daplex data definition language is presented.

```

DATABASE University IS

TYPE Person;
SUBTYPE Student;
SUBTYPE Faculty;
TYPE Courses;
TYPE CoursesOffered;

TYPE Person IS
    Name      : STRING(1..30);
    Identifier : STRING(1..9);
END ENTITY;

TYPE Courses IS
    Course#   : STRING(1..8);
    Title     : STRING(1..20);
    Descrip   : STRING(1..100);
    PCourse#  : SET OF Courses;
END ENTITY;

TYPE CoursesOffered IS
    Course#   : STRING(1..8);
    Section#  : INTEGER;
    TERM      : STRING(1..10);
END ENTITY;

SUBTYPE Student IS Person
    StTakes : SET OF CoursesOffered;
    GPA     : FLOAT;
    CampusAddress : STRING(1..100);
END ENTITY;

SUBTYPE Faculty IS Person
    Teaches : SET OF CoursesOffered;
    Ophone  : STRING(1..7);
    Department : STRING(1..30);
END ENTITY;

UNIQUE Course# WITHIN Courses;
UNIQUE Section# WITHIN Formats;
UNIQUE Identifier WITHIN Person;
UNIQUE Course#, Section#, Term WITHIN
    CoursesOffered;

END UNIVERSITY;

```

The FDM version of the University database has subtle differences from its RDM counterpart. Note that Student and Faculty in the RDM version have been reorganized into a generalization hierarchy with Person (root entity type with commonalities) and Student and Faculty (children or entity subtypes). Courses is similar, but the foreign key PCourse# has been replaced by an entity-valued function. The other two differences are: CoursesOffered no longer includes the faculty identifier and has been replaced by a Teaches entity-valued function in Faculty; and, Enrolled-Students has been replaced by a Takes entity-valued func-

tion in Student. The end of the definition of the database includes the various uniqueness constraints (akin to primary keys).

### The Extensible Markup Language (XML)

The eXtensible Markup Language (XML) [22] has emerged as a standard for information modeling and exchange for web-based applications, database interoperability, common software tool formats, patient record data, etc. XML allows information content to be hierarchically organized and tagged to highlight important/relevant content. The tags capture not only the content, but can be leveraged to represent the meaning of the information (semantics). In a web-based setting, XML is the successor to HTML to allow information content to be hierarchically organized and tagged to highlight important and relevant content. The tags can be exploited to capture both information content and the meaning of the information (semantics). In addition, XML allows the definition of templates to capture the known structure information for an application by the creating of XML schema files called Document Type Definitions (DTDs). The resulting XML document instances that are created contain both information content (data) as well as semantic notations (tags) that indicate the meaning of the information.

The Extensible Markup Language, XML [22] provides a flexible means to store and transmit data between different information systems and platforms, and has emerged as a dominant means for interoperability on the World Wide Web (as a successor to HTML) and as a standard for information format and exchange (e.g., the OpenDocument project for office applications). Both HTML and XML use tags to identify data/elements. However, while the HTML tags are predefined and specify the way the data within the tags is to be displayed, the XML tags are user-defined and can be employed to identify the data structure in a hierarchical fashion. For instance in HTML, `<b>` or `</b>` are all predefined tags to display the data/characters between them in a bold font. In XML, you will not have those kinds of predefined tags. All of the tags are defined by the users to indicate the structure of the data elements. An XML document for faculty data would be as simple as:

```
<?xml version='1.0' encoding='ISO-8859-1' ?>
<Faculty>
  <Name> John Smith </Name>
  <FacultyID> 12121212 </FacultyID>
  <Ophone> 5551111 </Ophone>
  <Department> Computer Science </Department>
</Faculty>
```

In the XML fragment, the tags `<Faculty>`, `<Name>`, `<FacultyID>`, etc., are defined, based on the application and/or modeling requirements. These tags capture data content (with attribute names) as well as data dependencies; they do not provide any clues on formatting or displaying the data elements between them. In addition to these structural capabilities of an XML file, each of the elements can be further quantified with tags to track semantic aspects of the data. For example, the Name could be augmented with a tag that indicates whether it is a married or maiden name, and the Department could be tagged with a title (e.g., Professor, Associate Professor, Head, etc.). If the XML structure contained numerical data (e.g., GPA), then the additional tags may denote the scale of the data (e.g., out of 4.0, 100, etc.).

The overall appearance and structure of an XML file can also be defined and validated by an XML schema file. An XML schema file defines the tags or attributes that can be used and where they appear in an XML file. There are two types of XML schema files: the Document Type Definition (DTD) and the XML Schema (a more sophisticated construct). A sample XML Schema for the previous XML fragment can be defined as:

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<xss:schema xmlns:xss=
  "http://www.w3.org/2001/XMLSchema">

<xss:element name="Faculty">
  <xss:complexType>
    <xss:sequence>
      <xss:element name="Name"
        type="xss:string"/>
      <xss:element unique name="FacultyID"
        type="xss:string"/>
      <xss:element name="Ophone"
        type="xss:string"/>
      <xss:element name="Department"
        type="xss:string"/>
    </xss:sequence>
  </xss:complexType>
</xss:element>
```

This sample XML Schema file defines that in the corresponding XML file, the Faculty element is a complex type (akin to a relation in RDM, file in ABDM, and entity type in FDM) must have one Name, one FacultyID, one Ophone, and one Department element. Uniqueness is achieved via the “unique” keyword attached to the FacultyID element. With a XML Schema, an XML parser can be used to validate whether the elements within a XML file are valid or not with respect to the XML Schema. Concep-

tually, a XML Schema is analogous to type declaration for a RDM relation, a ABDM file, or a FDM entity type (sub-type).

## Suitability of Information Modelsbreak for Granular Computing

This section, compares the four information models as given in Sect. “[Candidate Information Models](#)” (ABDM, RDM, FDM, and XML) for their suitability in support of granular computing. To facilitate the comparison, criteria are utilized that are relevant for comparing data models [3,4], in conjunction with various concepts proposed in [25]. The criteria chosen are:

- Formalism and Theory: This criterion is used to assess the degree to which a formal model exists for the data model that is capable of supporting GrC.
- Expressive Power: This criterion is used to evaluate the ability of each data model to support the definition, usage, and analysis of granules,
- Extensibility and Utility: This criterion is used to determine the usability of the data model, from many different perspectives.

In the remainder of this section, these criteria are explored against the four models, as a means to assess their suitability for GrC.

### Formalism and Theory

All of the data models given in Sect. “[Candidate Information Models](#)”, have a certain degree of formalism/theory upon which they are based. GrC relies on an information model that has a formal basis, in order to capture the required information within a data table [15] and to employ a theory of rough sets and fuzzy logic for granular analysis. As a result, it is vital that an underlying data model is rich enough to support the formalisms necessary for GrC. To begin, the attributed-based data model [6] was proposed by David K. Hsiao (founder of ACM Transactions on Database Systems) and Frank Harary (noted graph theorist) and has a formal basis that has been explored by other researchers in the 1970s [16], 1980s [4], and 1990s [9], with applicability in many different contexts including paging environments for OS, data-model transformations for federated databases, and polyinstantiation for security, respectively. Likewise, the relational data model [1] has a rich history and formal basis using set theoretic concepts, and is the dominant approach to date for GrC researchers [2,13,20].

The functional data model [17] was proposed in an era in the early-to-mid 1980s when semantic data mod-

els (successors to the entity-relationship data model) were advocated to break away from relatively flat approaches (e.g., relational, hierarchical, and network data models and databases systems) to an approach that was semantically rich and able to model the data as it occurs in the “real world”. In fact, in a workshop in 1993 later published as special journal issue [5], the overriding theme emphasized the functional model and its potential for a unifying paradigm, providing formalisms based on functional and logic programming that could be used to capture both relational data and object-oriented data in a formal way. Lastly, XML [22] has emerged as a de facto standard across many different disciplines, being used in a wide variety of contexts (web semantics, patient health records/standards, database interoperability, information exchange format, etc.). XML is based on detailed specifications [23] and has a grammar in Extended Backus–Naur Form. As a result, there are XML parsers that have been built for many purposes to parse XML files and XML Schemas (see Sect. “[The Extensible Markup Language \(XML\)](#)” again) in different ways. Such a formal-grammar basis means that XML can be utilized in a very rigorous manner to reason using context-free grammar concepts from automata theory. In summary, it is clear that from a formal basis, each of these four data models can realize the information/data table needs of GrC, offering different strengths in their support.

### Expressive Power

The expressive power needed to support information granules can be impacted by many factors; we choose three for our comparison:

- *Grouping capability* which involves the reason that objects are grouped into granules (equivalence classes) and can be based on object relationships, similarity, proximity, semantics, etc. [25].
- *Linguistic representation of granules* which assigns a linguistic value to each granule that can be accomplished using a naming convention or by providing sample representative objects [25].
- *Constraint and type checking* which involves the ability to enforce intra-granular and inter-granular dependencies (constraint checking) while simultaneously insuring that the granule always adheres to the defined structure (type checking).

From an information model perspective, these three granulation factors are directly related to the features and characteristics of the data models as presented in Sect. “[Candidate Information Models](#)”, and includes: aggregation

(records in ABDM, relations in RDM, types and subtypes in FDM, and XML Schemas in XML), generalization (inheritance in FDM), identity and uniqueness (directory attributes in ABDM, primary keys in RDM, uniqueness in FDM, and unique in XML), and relationships (record references in ABDM, foreign keys in RDM, entity-valued functions in FDM, and hierarchical structure in XML).

Specifically, for the grouping capability factor, the assembly of objects into granules (equivalence classes) is aggregation, which is provided by all four data models. Aggregation is simply a result of the creation of the equivalence class, and the aggregation can be based on similarity, proximity, etc. The process of grouping objects into granules is part of GrC, but once the granulation has been completed, then any of these four models have sufficient aggregation to represent a granule. If granules are related to one another hierarchically (or the grouping occurs based on a hierarchical relationship), then either FDM or XML is appropriate for this resultant aggregation. For the linguistic representation of granules factor, again, the same aggregation capability applies; all four models provide a naming convention, and the examples as given in Sect. “[Candidate Information Models](#)” for ABDM and XML illustrated data instances (sample representative objects).

Finally, for the constraints and type checking factor, each model offers different capabilities. For ABDM, uniqueness of records, record references across objects (instances), support for attribute-value pairs (used by many granulation researchers), and directory attributes (akin to keys), all provide the basis for constraint checking; the definition of the attributes (name and type) in each record provides type checking. For RDM, no duplicate tuples per relation, primary keys, foreign keys, and referential integrity are relevant for constraint checking; a relational table definition with its attributes and data types provides type checking. For FDM, the types and subtypes provide explicit support for type checking (i.e., these types and subtypes are akin to programming languages types and subtypes in Ada), while uniqueness constraints preserve properties (constraint checking). For XML, the XML Schema is a template that an XML instance must follow (type checking), and in its definition, provides the ability for dependencies (constraint checking); as such, for different granules, there would be different XML Schemas, and XML provides the ability to parse and map from one XML Schema to another via XSL Transformations (<http://www.w3.org/TR/xslt>) and XQuery (<http://www.w3.org/TR/xquery>). Thus, XML is much more powerful than the other models, particularly in type checking; the use of XML can facilitate the ability to transform one

granule to another while still preserving content and semantics.

### Extensibility and Utility

The final criterion considered is extensibility and utility, which is focused on the usability of the various models, both today and in the future. This can be examined from many different perspectives:

- Database Platform Support: While Sect. “[Formalism and Theory](#)” considered the theoretical basis for GrC, when one moves from theory to practice in GrC, a seamless transition to the corresponding database management system for a chosen data model would be very useful. In that regard, RDM has the definite advantage, with a wide variety of commercial (Oracle, SQL Server, Informix, etc.) and Open Source (MySQL, PostgreSQL, Ingres, etc.) products available. XML is also a strong player in this regard, since it is a dominant technique for information exchange (data interoperability) among databases, particularly across networked and distributed database solutions.
- Future Potential: XML has the greatest potential for future usage in GrC as an information model, due to its widespread and growing usage. In fact, there are many open source XML databases that are starting to be released, including: Apache Xindice, Senda XML DBMS, X-Hive/DB, etc. Many of these are in their earliest release stages. For RDM, SQL3, which will include object-oriented extensions, is still being considered from a standards perspective. These object-oriented capabilities are critical to allow granules to be formed in a representation similar to the way that they occur in the “real world”, rather than having the granule be flattened into a relational table. Once approved, the relational database vendors will have to implement SQL3 and release new versions. One can easily hypothesize that XML database systems may make SQL3 obsolete before it even is released.
- Data Model Translations: The ABDM as given in Sect. “[The Attribute-Based Data Model \(ABDM\)](#)”, has been shown to be capable of subsuming the features of RDM and FDM [4] at both a model and system level. This means that a single model can be used in place of other models without a loss of information.

Overall, the clear leader from a model perspective is ABDM and its ability to subsume RDM and FDM model capabilities. From a practical perspective, while RDM may have the edge in commercial and open source database platforms, the future seems pointed towards XML for sharing and database usage.

## Future Directions

In this chapter, candidate information models and their suitability for granular computing (GrC) have been explored. This chapter considered four different models, presented in Sect. “[Candidate Information Models](#)”: the attribute-based data model (ABDM) [6], the relational data model (RDM) [1], the functional data model (FDM) [17], and the extensible markup-language (XML) [22]. Two of these models (RDM and XML) are dominant from a database system perspective today, but all four models have different capabilities to offer when modeling, reasoning, and analyzing for GrC. To place the four models in their proper perspective, in Sect. “[Suitability of Information Models for Granular Computing](#)”, the models were explored using three different criteria: Formalism and Theory, Expressive Power, and Extensibility and Utility: For the first criterion (Sect. “[Formalism and Theory](#)”), each of the four models have a strong formal basis and long history of usage; to varying degrees all are suitable for a formal basis for GrC. For the second criterion (Sect. “[Expressive Power](#)”), we compared based on three factors: grouping capability, linguistic representation of granules, and constraint and type checking. In the analysis, ABDM, RDM, and FDM are very comparable in what they offer, but XML has the potential to transcend all three, with both its current capabilities and its emerging characteristics and features. Finally, for the third criterion (Sect. “[Extensibility and Utility](#)”), there were different strengths for different models: for ABDM, its strength was its ability to subsume RDM and FDM at the data model level [4]; for RDM, its strength was in the deployed database platforms (commercial and open source) and the future potential of SQL3 with object-oriented capabilities (which will improve granule representation); and, for XML, the usage in relational databases for information exchange, coupled with the arrival of XML database systems, may have the greatest potential for the future. This leaves us with two recommendations: from a theory perspective, while all four data models are appropriate to some degree for GrC, ABDM and RDM have to be strongly considered due to their long history of usage and formal basis, with FDM an alternative when there is a desire to have a functional/logic based basis to GrC; and, from a practical perspective, if one is to transition from GrC theory to actual usage, RDM is the choice today, but XML has an excellent chance to be the choice of the future.

## Bibliography

1. Codd E (1970) A relational model of data for large shared data banks. *Commun ACM* 13(6):377–387
2. Demchenko Y et al (2005) Security architecture for open collaborative environment. In: Sloot P et al (eds) *Advances in grid computing – EGC2005*, vol 3470. LNCS. Springer, Heidelberg
3. Demurjian S, Hsiao D (1987) The multi-lingual database system. In: Proceedings of the 3rd international conference on data engineering. IEEE Computer Society, Washington DC
4. Demurjian S, Hsiao D (1988) Towards a better understanding of data models through the multilingual database system. *IEEE Trans Softw Eng* 14(7):946–958
5. Gray P, King P, Kerschberg L (1999) Functional approach to intelligent information systems. *Special Issue J Intell Inf Syst (JIIS)* 12(2–3):107–111
6. Hsiao D, Harary F (1970) A formal system for information retrieval from files. *Commun ACM* 13(2), Corrigenda 13(4): 67–73
7. Lin T (1988) Neighborhood systems and approximation in relational database and knowledge bases. In: Proceedings of the fourth international symposium on methodologies of intelligent systems, poster session. Oct. 1989. IEEE Computer Society, Washington DC
8. Lin T (1989) Chinese wall security policy – an aggressive model. In: Proceedings of the fifth aerospace computer security application conference. Dec. 1989. IEEE Computer Society, Washington DC
9. Lin T (1992) Attribute based data model and polyinstantiation. In: Aiken R (ed) *Education and society – information processing '92*, vol 2, proceedings of the IFIP 12th world computer congress. North-Holland, 1992
10. Lin T (1997) Granular computing. In: Announcement of the BISC special interest group on granular computing
11. Lin T (1998) Granular computing on binary relations I: Data mining and neighborhood systems. In: Skowron A, Polkowski L (eds) *Rough sets in knowledge discovery*. Physica, Heidelberg, pp 107–121
12. Lin T (1998) Granular computing on binary relations II: Rough set representations and belief functions. In: Skowron A, Polkowski L (eds) *Rough sets in knowledge discovery*. Physica, Heidelberg, pp 121–140
13. Lin T (2005) Granular computing: A problem solving paradigm. In: Proceedings of 14th international conference on fuzzy systems, 2005. IEEE Computer Society, Washington DC
14. Lin T (2006) A roadmap from rough set theory to granular computing. In: Wang G et al (eds) *Proceedings of first international conference on rough sets and knowledge technology – RSkt 2006*, vol 4062. LNCS. Springer, Heidelberg pp 33–41
15. Pawlak Z (1991) *Rough sets – theoretical aspects of reasoning about data*. Kluwer, Heidelberg
16. Rothnie Jr JB (1974) Attribute based file organization in a paged memory environment. *Commun ACM* 17(2):63–69
17. Shipman D (1981) The functional data model and the data language DAPLEX. *ACM Trans Database Syst* 6(1):140–173
18. Skowron A, Stepaniuk J (2001) Information granules: Towards foundations of granular computing. *Int J Intell Syst* 16(1): 57–86
19. Smith J, Smith D (1977) Database abstractions: Aggregation and generalization. *ACM Trans Database Syst* 2(2):405–413
20. Wang D et al (2004) Medical privacy protection based on granular computing. *Artif Intell Medicine* 32(2):137–149
21. Wong E, Chiang T (1971) Canonical structure in attribute based file organization. *Commun ACM* 14(9):593–597

22. World Wide Web Consortium. <http://www.w3c.org/XML/>. Accessed 10 July 2008
23. XML 1.0 Fourth Edition Specification. <http://www.w3.org/TR/2006/REC-xml-20060816/>. Accessed 10 July 2008
24. Zadeh L (1996) Fuzzy logic = Computing with words. *IEEE Trans Fuzzy Syst* 4(2):103–111
25. Zadeh L (1997) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets Syst* 90(2):111–127
26. Zadeh L (1998) Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems. *Soft Comput* 2(1):23–25

## Granular Computing, Introduction to

TSAU YOUNG LIN  
 Department of Computer Science, San Jose State University, San Jose, USA

### Article Outline

- Granular Computing from Rough Set Theory
- Granular Computing in Database Theory
- Granular Computing in Social Networks
- Granular Computing and Fuzzy Set Theory
- Grid/Cloud Computing – A New Addition to GrC
- General Issues in Granular Computing
- Conclusions
- Bibliography

What is granular computing (GrC)? It is a shifting paradigm. Let us start with a few words about how the term was coined. In the academic year 1996–97, when Lin (this section editor) took his sabbatical leave at UC-Berkeley, Zadeh suggested granular mathematics (GrM) as his research area. To limit the scope, Lin proposed the term *granular computing* [14]. What was GrC at that time? Zadeh had outlined it in his 1997 seminal paper [15]. Lin took an incremental approach: he mapped his neighborhood system [5] to Zadeh's intuitive definition [12] and used it as his First GrC model [8,9,10]. It may be important to point out that the concept of neighborhood systems, which was motivated from approximate retrieval in databases [7], is a generalization of the topological neighborhood system that formalizes the ancient intuition of infinitesimal granules.

Much progress has been achieved since then. This section has been organized to represent this progress and to reflect the current state of GrC. We believe in the incremental approach, namely, that each new step is based on solid results and moves forward. So many special theories

and applications are gathered here. Jointly, they reflect the current state of GrC and may also implicitly hint at the ultimate goals of GrC.

To grasp the main idea from such a diverse collection of papers, a roadmap will be helpful: We suggest the reader start with the first four sections of T.Y. Lin's paper ► *Granular Computing: Practices, Theories, and Future Directions*. There, the reader may want to pay special attention to the first three examples:

**E1)** Human body is granulated into head, neck, etc.

So far, there is no flawless formal model. Obvious models do not work satisfactorily; we need a much more subtle theory.

**E2)** Space-time is granulated intuitively into infinitesimal granules.

**E3)** The Heisenberg uncertainty principle.

The last two examples have been fully digested by mathematicians and scientists. There are two solutions to the first example, namely, topology and non-standard analysis, so the readers may want to skim through Lin's Section "Second GrC Models and Modern Examples" (about neighborhood system/pre-topology) in ► *Granular Computing: Practices, Theories, and Future Directions* and the two articles, ► *Non-standard Analysis, An Invitation to* by Wei-Zhe Yang, and ► *Granular Computing and Modeling of the Uncertainty in Quantum Mechanics* by Kow-Lung Chang. These are classical subjects; the authors are a mathematician and a physicist, respectively.

Having digested these readings, the reader may proceed to Zadeh's ► *Fuzzy Logic* to gain some feeling about the modern view, and may use it to examine the very first example, E1. The article is full of revolutionary ideas; it is, we believe, one of his best papers in presenting an overview of his idea.

Here we paraphrase or quote some of his assertions from his article: "There are many misconceptions about fuzzy logic. A common misconception is that fuzzy logic is fuzzy. In reality, fuzzy logic is not fuzzy. Fuzzy logic deals precisely with imprecision and uncertainty. In fuzzy logic, the objects of deduction are, or are allowed to be, fuzzy, but the rules governing deduction are precise."

Fuzzy logic is much more than a logical system. More specifically, fuzzy logic has many facets. "The principal facets are: the logical facet, FLL; the fuzzy-set-theoretic facet, FLs, the epistemic facet, Fle; and the relational facet, FLR."

To gain a glimpse into the nature of fuzzy logic, we examine some examples taken from his article. The left-hand column is "the familiar example of deduction in Aristotelian, bivalent logic. In this example, there is no impre-

cision and no uncertainty.” On the other hand, the right-hand column is an example of fuzzy logic which is “in an environment of imprecision and uncertainty.”

all men are mortal	most Swedes are tall
Socrates is a man	Magnus is a Swede
Socrates is mortal	it is likely that Magnus is tall

To deduce the answer from the premises (for the right-hand example), “it is necessary to precisiate the meaning of “most” and “tall,” with “likely” interpreted as a fuzzy probability which, as a fuzzy number, is equal to “most.” This simple example points to a basic characteristic of fuzzy logic, namely, in fuzzy logic precisiation of meaning is a prerequisite to deduction.” In this example, “deduction is contingent on precisiation of “most,” “tall,” and “likely.” The issue of precisiation has a position of centrality in fuzzy logic.” “In fuzzy logic, the deduction is viewed as an instance of question-answering”.

[Digression] To this point, we may want to observe that there is some similarity to Lin’s context based reasoning method, which is a derivative of  $(\epsilon, \delta)$ -definition; see Lin’s article “Second GrC Models and Modern Examples” in ▶ [Granular Computing: Practices, Theories, and Future Directions](#) on the Meaning of “Near”.

To avoid mis-representation, we will not offer a summary here, but urge the reader to read Zadeh’s article.

There are five areas that have great interactions with GrC, namely, (in alphabetical order) databases (especially data mining), fuzzy theory (especially fuzzy control), grid/cloud computing, rough set theory (implicitly extended to topology) and Social Networks. We will group these articles (plus general issues) accordingly. We shall start from rough set theory.

### Granular Computing from Rough Set Theory

Briefly, rough set theory (RST) is a theory of equivalence relations (which are equivalent to partitions), so the RST community views GrC as a generalized partition theory. Recall that an equivalence relation is a reflexive, symmetric and transitive binary relation. Thus, the “next” generalizations are the tolerance relation (dropping transitive), the partial ordering relation (dropping symmetric), and the fuzzified equivalence relation.

We have collected the corresponding articles: ▶ [Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems](#) by Lech Polkowski, ▶ [Granular Computing and Data Mining for Ordered Data: The Dominance-Based Rough Set Approach](#) by Salvatore Greco, Benedetto Matarazzo1, and Roman Slowin-

ski, ▶ [Rough and Rough-Fuzzy Sets in Design of Information Systems](#) by Theresa Beaubouef and Frederick Petry and ▶ [Multi-Granular Computing and Quotient Structure](#) by Ling Zhang and Bo Zhang.

Thanks to Professor Polkowski’s diligent search, we are aware that the term “tolerance relation” first occurred much earlier than one might expect. However, in the rough set community, Nieminen (1988) and Lin [7] were probably the first to use the terms “tolerance equality” and “neighborhood of tolerance,” respectively. In fact, progress on the tolerance relation is quite far reaching; it is a complete Pawlak Theory (CPT); it is implicitly in [11].

Recall that a RST is called a CPT, if it includes analogous theories covered in Pawlak’s book (except logic), namely, a theory of approximations and a complete knowledge representation system; the latter notion is defined below (Definition 1).

There are more results in this directions: Let  $(U, \beta)$  be a Global GrC Model with  $\beta$  being a full covering. Then the model  $(U, \beta)$  is a CPT, if  $\beta$  is a semi-group under intersection. However, there are counter examples, namely, some GrC on partial ordering cannot be a CPT.

The quotient structures, addressed by Zhang-Zhang, have been neglected by the rough set community, except some brief results by Lin. Quotient structures are essential for information hiding and knowledge representations.

The quotient structure of a partition is simply a classical set. Zhang-Zhang actually has considered topological cases, and Lin has considered pre-topological cases (neighborhood systems). In general, it can be much more complicated. We shall conclude this paragraph with an example from algebraic geometry: Let the universe  $U$  be the polynomial ring over the complex numbers, and  $\beta$  be the collection of all prime ideals. Then the quotient structure is the complex line plus a generic point; see “Granular and Related Structures” of Lin’s article ▶ [Granular Computing: Practices, Theories, and Future Directions](#) for more details.

### Granular Computing in Database Theory

Relational databases can be related to GrC in two views, Fifth or Fourth GrC models (Relational or Multi-Binary GrC Models).

A Fourth GrC Model  $GDM_1 = (U, \beta)$  is called a Granular Data Model (GDM), if the granular structure  $\beta = \{R_1, R_2, \dots\}$  is a collection of equivalence relations. By giving each equivalence relation and its equivalence classes meaningful names, GDM becomes a relation instance, called an information table (IT) in RST. This naming process is called a knowledge representation. Let

$IT_1 = (U, A_1, A_2, \dots)$  be a knowledge representation of  $GDM_1$ , where  $A_i$ , called an attribute, denotes the meaningful names of  $R_i$ . Pawlak observed a beautiful phenomena:

**Theorem 1**  $(U, R_1, R_2, \dots) \iff (U, A_1, A_2, \dots)$ , up to isomorphism.

From this theorem, we define:

**Definition 1** A knowledge representation of the Fourth GrC model is said to be a complete knowledge representation if the representation (the process) induces an equivalence. In other words, the knowledge representation can re-capture the model.

Base on this theorem, RST can transform data mining on IT to granule-processing, namely, granular computing, on GDM. This view clarifies the nature of automated data mining and simplifies its computations. For example, the association rule mining is reduced to counting the cardinality of set theoretical expressions of granules. However, we should also observe that in such computations the semantic part of attribute values are ignored. In general, the semantics of attribute values are not utilized in the RST/GrC approach. So some data mining techniques, such as clustering, cannot be approached by RST, since clustering uses the metric of the ambient space of attribute values. In this respect, GrC may use techniques in computing with words to deal with semantic issues. However, we should also note that computing with words is still in its inception stage.

In this group, we have collected the following articles, ► [Granular Computing, Information Models](#) for by Steven A. Demurjian, ► [Rule Induction, Missing Attribute Values and Discretization](#) by Grzymala-Busse, ► [Co-operative Multi-hierarchical Query Answering Systems](#) by Zbigniew W. Ras, Agnieszka Dardzinska, ► [Dependency and Granularity in Data-Mining](#) by Tsumoto-Hirano, ► [Rough Set Data Analysis](#) by Shusaku Tsumoto, ► [Granular Model for Data Mining](#) by Anita Wasilewska, Ernestina Menasalvas. Here, we would like to note that Steven A. Demurjian's article views GrC from the database, not the RST, perspective.

### Granular Computing in Social Networks

By interpreting a granule as an “ordered set” or a tuple (in some relations), Lin formalizes Social Networks into a Fifth GrC model, called the Relational GrC Model. Articles ► [Granular Computing System Vulnerabilities: Exploring the Dark Side of Social Networking Communities](#) by Steve Webb, James Caverlee, Calton Pu, and ► [Social](#)

[Networks and Granular Computing](#) by Churn-Jung Liau investigate two issues in social networks: security and positional equivalence.

This area actually started much earlier. In the mid-1970s, Atkin started to investigate mathematical structure in human affairs [1]. In fact, he reached a simplicial complex in [2], which is symmetric Fifth GrC model or equivalently a Second GrC model(Global GrC model).

### Granular Computing and Fuzzy Set Theory

Basically, any assertion about classical sets can be fuzzified, so the intention here is to address GrC applications that extend beyond set theory.

A fuzzy set is defined by a membership function, which is a bounded non-negative real-valued function. So it is rather natural to investigate the generalization of fuzzy set theory to general function theory; even to more general cases, such as random variables (measurable functions) and to generalized functions (such as Dirac Functions). So, we have proposed a Sixth GrC model, which is a function-based GrC model.

In most known applications, we often need some extra properties on the granular structure, such as the universal approximation property. Using Banach space's language, it implies that the granular structure is a Schauder base. The collection of membership functions used in fuzzy control, and the activation functions in neural networks (e.g., Radial-Basis-Functions), all have such a property.

For this group, we have selected four papers: ► [Genetic-Fuzzy Data Mining Techniques](#) by Tzung-Pei Hong, Chun-Hao Chen and Vincent S. Tseng, ► [Fuzzy System Models Evolution from Fuzzy Rulebases to Fuzzy Functions](#) by I.B. Turksen, ► [Granular Neural Network](#) by Yan-Qing Zhang, and ► [Fuzzy Probability Theory](#) by Michael Beer.

We should note that many soft computing papers can be classified in this category; however, it has its own section, so we recommend that the readers visit that section. We specifically recommend reading some papers on Type II fuzzy sets, whose membership functions can be viewed as granule-valued (fuzzy-number-valued) membership functions.

We also recommend that readers compare Michael Beer's ► [Fuzzy Probability Theory](#) and non-standard probability theory (in Wei-Zhe Yang's article). From them, the reader might draw his own conclusions as to what a granular probability (granule-valued probability) theory might be. One could ask even further questions, such as, what is a granular function (granule valued function)?

## Grid/Cloud Computing – A New Addition to GrC

In the end, we would like to note that a new technology, called cloud computing (a subset of grid computing) is intrinsically a technology of granular computing. This is a most natural example of a Seveth GrC models (a Turing machine-based GrC Model), even though the name GrC is foreign to the experts in this field. We are expecting, in the near future, that this area will be the top group that influences GrC development.

### General Issues in Granular Computing

Two articles, ► [Granular Computing, Principles and Perspectives of](#) by Jianchao Han and Nick Cercone, and ► [Granular Computing, Philosophical Foundation for](#) by Zhengxin Chen, are about general views on GrC. However, they are basically addressing First, Second and Third models. Professor Chen's three-dimensional views, philosophical, technical, and social/application dimensions. are quite intriguing and may trigger more general investigations into all of GrC. In this group, we will add Kow-Lung Chang, Tsau Young Lin, Wei-Zhe Yang, Lotfi A. Zadeh Zadeh's articles.

### Conclusions

As GrC is an emerging field, many directions are possible (Zadeh [13,14,15]; Lin [6] pre-CrC, [7,8,9]; Bargiela and Pedrycz [3]; Jankowski and Skowron [4]). Some important works, due to lack of time and some other factors (including ignorance of this section editor), may not be included here. We apologize for the omissions, with the hope that many of them are referenced by some of the collected articles.

In editing this section, we have adopted an incremental approach: all papers are scientific papers, the only visionary paper is Lotfi Zadeh's "Fuzzy Logic"; and even in that paper, many of the visions are already accomplished.

Looking at the collection, we believe GrC is a promising field; here are some indications:

1. The recent emergence of cloud computing to GrC (our view) may give GrC a momentum to the real world applications.
2. From the abstract point of view (category theory based models) GrC and databases have the same abstract structures. This implies that GrC as an abstract concept can be technologically realized.
3. GrC has many links to various branch of mathematics, namely, algebraic topology (simplicial complex) homological algebra (extension functor), algebraic geometry ( $\text{Spec}(R)$ ). Topological spaces, probability and be-

lief functions. These facts indicate, that mathematically, GrC contains some important common structures that are touched by many distinct mathematical fields.

Finally, we would like to extend our thanks to all of the authors for their support of this project, to Professor Zadeh for his generous guidance and advice, to Dr. Robert Meyers for the opportunity to edit this section.

### Bibliography

1. Atkin RH (1974) Mathematical strucutures in human affairs. In: Heinemann educaton book. Pitman Press, Bath. ISBN 0 435 820257
2. Atkin RH (1977) Combinatorial connectivites in socail systems. Birkhäuser, Basel
3. Bargiela A, Pedrycz W (2002) Granular computing. Kluwer, Boston
4. Jankowski A, Skowron A (2007) Toward rough-granular computing. In: Proceedings of the 11th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, (RSFDGrC'07). Toronto, Canada. Lecture Notes in AI. Springer, pp 1–12
5. Lin TY (1988) Neighborhood systems and relational database. In: Proceedings of CSC'88. ACM Press, New York, pp 725
6. Lin TY (1989) Chinese wall security policy – an Aggressive Model. In: Proceedings of the Fifth Aerospace Computer Security Application Conference, December 4–8, 1989. Kluwer 2004, pp 286–293
7. Lin TY (1989) Neighborhood systems and approximation in database and knowledge base systems. In: Proceedings of the Fourth International Symposium on Methodologies of Intelligent Systems (Poster Session), pp 75–86
8. Lin TY (1998) Granular computing on binary relations I: data mining and neighborhood systems. In: Skowron A, Polkowski L (eds) Rough Sets In Knowledge Discovery. Physica-Verlag, Heidelberg, pp 107–121
9. Lin TY (1998) Granular computing on binary relations II: rough set representations and belief functions. In: Skowron A, Polkowski L (eds) Rough sets in knowledge discovery. Physica-Verlag, Heidelberg, pp 121–140
10. Lin TY (1999) Granular computing: fuzzy logic and rough sets. In: Zadeh L, Kacprzyk J (eds) Computing with words in information/intelligent systems. Springer Lecture Notes in AI. Physica-Verlag, Heidelberg, pp 183–200
11. Lin TY (2006) A roadmap from rough set theory to granular computing. RSKT 33–41
12. Zadeh LA (1996) The key roles of information granulation and fuzzy logic in human reasoning. In: IEEE International Conference on Fuzzy Systems, September 8–11, 1
13. Zadeh LA(1979) Fuzzy sets and information granularity. In: Gupta M, Ragade R, Yager R (eds) Advances in fuzzy set theory and applications. North-Holland, Amsterdam, pp 3–18
14. Zadeh LA (1998) Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems. Soft Comput 2:23–25
15. Zadeh LA (1998) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst 90:111–127

# Granular Computing and Modeling of the Uncertainty in Quantum Mechanics

KOW-LUNG CHANG

Physics Department, National Taiwan University,  
Taipei, Taiwan

## Article Outline

Glossary

Definition of the Subject

Introduction

Quantum Postulates and Associated Propositions

Dirac's Bra and Ket Notations and the Realization  
of Wave Function

Realization of  $q$ -Representation in Quantum Mechanics

Fourier Transformation and  $p$ -Representation

Quantum Uncertainty and Non-Compatibility  
of Observables  $P$  and  $X$

Conclusion

Future Directions

Bibliography

## Glossary

**Newtonian mechanics** The classical mechanics based on Newton's law of motion.

**Uncertainty** Also commonly referred to as error; deviation from the average value.

**Extrinsic uncertainty** The uncertainty due to the systematic error and random error.

**Intrinsic uncertainty** The uncertainty due to the nature of particle-wave duality in the quantum system.

$\psi(r)$   $q$ -representation of the quantum state, called state function or wave function in coordinate space.

$\varphi(p)$   $p$ -representation of the quantum state, called the state function or the wave function in momentum space.

$\rho(r)$  Probability density in coordinate space, defined as  $\rho(r) = \psi^*(r)\psi(r) = |\psi(r)|^2$ .

$\tilde{\rho}(p)$  Probability density in momentum space, defined as  $\tilde{\rho}(p) = \varphi^*(p)\varphi(p) = |\varphi(p)|^2$ .

$|\psi(r)|$  Probability amplitude of state  $\psi(r)$ .

**Particle-wave duality** A quantum object exhibits both the nature of a particle and a wave.

**Wave vector**  $k = p/\hbar$ .

**Wave number** One-dimensional wave vector,  $k = p/\hbar$ .

**Quantum state** The physical state in a subatomic quantum system.

**0th Postulate of quantum mechanics** Regarding quantum states as elements of Hilbert space  $\mathcal{H}$ .

**1st Postulate of quantum mechanics** Assigning each dynamical variable in a quantum system a unique linear Hermitean operator in Hilbert space  $\mathcal{H}$ .

**2nd Postulate of quantum mechanics** The set of eigenvectors of a given observable forms the bases of a Hilbert space.

**3rd Postulate of quantum mechanics** Poisson brackets in classical mechanics are replaced by commutators of the corresponding observables according to the relations in Eq. (5)

$[R, S]$  Commutator of operator  $R$  and operator  $S$  defined as  $RS - SR$ .

$\{R, S\}$  Anti-commutator of operator  $R$  and operator  $S$  defined as  $RS + SR$ .

**Eigenvalue and eigenvector** For operator  $A$  acting upon a particular state  $\psi_a$ , such that  $A\psi_a = a\psi_a$ , then the scalar number  $a$  and the state  $\psi_a$  are called respectively the eigenvalue and eigenvector of operator  $A$ .

**Inner product** A numerically valued function of the ordered pair of vectors  $\psi$  and  $\varphi$ , denoted by  $(\psi, \varphi)$  such that  $(\psi, \varphi) = (\varphi, \psi)^*$ . In Dirac's notation, it takes the forms  $\langle \psi | \varphi \rangle = \langle \varphi | \psi \rangle^*$ .

**Hilbert space** A complete vector space with norm defined as the inner product.

**Dual space** The space formed by the set of all functionals satisfying the linearity conditions.

**Hermitean operator** An operator which is self-adjoint; that is, an operator  $A$  equals its Adjoint conjugate,  $A = A^+$ .

**Adjoint conjugate operator** For an operator  $A$  and a pair of vectors  $\psi$  and  $\varphi$ , the Adjoint operator to  $A$ , denoted by  $A^+$  if it satisfies the relation

$$\langle \psi, A\varphi \rangle = \langle \varphi, A^+\psi \rangle^* = \langle A^+\psi, \varphi \rangle.$$

$|\psi\rangle$  A ket vector in Hilbert space.

$\langle\psi|$  A bra vector in dual space.

**Compatible observables** Physical observables that commute to each other.

**Projection operator**  $|a_i\rangle\langle a_i|$  a Hermitean operator that projects any vector in  $\mathcal{H}$  onto a subspace.

**Closure relation** Direct sum of all project operators equals identity operator.

**$\delta$ -function** A distribution function denoted by  $\delta(x - a)$  such that

$$\int_D f(x)\delta(x - a)dx = f(a) \quad \text{and}$$

$$\int_D f(x)\delta'(x - a)dx = -f'(a) \quad \text{if } a \in D,$$

$$\int_D f(x)\delta(x - a)dx = 0 \quad \text{if } a \notin D.$$

**Cauchy–Schwarz inequality** Given vectors  $|\alpha\rangle$  and  $|\beta\rangle$  such that

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq \langle\alpha|\beta\rangle\langle\beta|\alpha\rangle.$$

**G<sub>d</sub>** Deviation operator with respect to observable G defined as  $G_d = G - \langle G \rangle I$

### Definition of the Subject

For data collection or data acquisition in granular computing or modeling, particular attention should be paid to data that is involved with the position measurement and/or the momentum measurement of a particle in a physical system. For a system that is governed by Newton's law of motion, measurements of the positions and the momenta of a classical mass point can be made as precise as one wishes if the best measurement conditions are available. When we perform physical measurements in microscopic systems at subatomic level, the laws of Newtonian mechanics cease to apply and instead a new theory, quantum theory, is introduced to account for the quantum phenomena of the subatomic system.

Uncertainty in the measurement of a quantum system comes from two origins. The first is called the extrinsic uncertainty. Quantum measurements, just as in any scientific measurement, are never exact. There exists a systematic error or uncertainty as well as random ones during the processes of measurement for various reasons such as inaccurate calibration of instruments, limitation in resolution of apparatus or meters, variations in temperature or humidity or even the position adopted by the individual observer in reading data etc. The second is called the intrinsic uncertainty which arises only in the case of a quantum mechanical system.

This particular quantum uncertainty is inherent in the very nature of particle-wave duality [1] for a subatomic quantum system. The intrinsic uncertainty exists always even if the extrinsic uncertainty is eliminated or minimized to the level of negligibility.

In order to describe the particle-wave duality for a subatomic quantum system, one introduces the wave function, or state function,  $\psi(\mathbf{r})$  to represent the quantum state.  $\psi(\mathbf{r})$ , which provides all the dynamic quantities of the system, and is related to the probability density by  $\rho(\mathbf{r}) = \psi^*(\mathbf{r})\psi(\mathbf{r}) = |\psi(\mathbf{r})|^2$ . Therefore, the average position of the quantum object  $\langle\mathbf{r}\rangle$  can be obtained through the following integration

$$\langle\mathbf{r}\rangle = \int \rho(\mathbf{r})\mathbf{r}d^3\mathbf{r} = \int \psi^*(\mathbf{r})\mathbf{r}\psi(\mathbf{r})d^3\mathbf{r}. \quad (1)$$

The uncertainty in position, denoted by  $\Delta\mathbf{r}$ , can then be expressed by

$$\begin{aligned} (\Delta\mathbf{r})^2 &= \int \rho(\mathbf{r})(\mathbf{r} - \langle\mathbf{r}\rangle)^2 d^3\mathbf{r} \\ &= \int \psi^*(\mathbf{r})(\mathbf{r} - \langle\mathbf{r}\rangle)^2 \psi(\mathbf{r})d^3\mathbf{r}. \end{aligned} \quad (2)$$

Similarly, the uncertainty in momentum,  $\Delta\mathbf{p}$  can be written as

$$\begin{aligned} (\Delta\mathbf{p})^2 &= \int \tilde{\rho}(\mathbf{p})(\mathbf{p} - \langle\mathbf{p}\rangle)^2 d^3\mathbf{p} \\ &= \int \varphi^*(\mathbf{p})(\mathbf{p} - \langle\mathbf{p}\rangle)^2 \varphi(\mathbf{p})d^3\mathbf{p}, \end{aligned} \quad (3)$$

where  $\varphi(\mathbf{p})$  is the state function of the quantum system in momentum space.

**The quantum uncertainties of the system obey the following relation [2]:**

$$\begin{aligned} \Delta x\Delta p_x &\geq \hbar/2 \\ \Delta y\Delta p_y &\geq \hbar/2 \\ \Delta z\Delta p_z &\geq \hbar/2, \end{aligned} \quad (4)$$

where  $\hbar = h/2\pi$ , and  $h$  is called the Planck's constant.

### Introduction

Some basic mathematical tools are necessary in order to understand the uncertainty relations of quantum mechanics in a more rigorous sense. As discussed in the last section, it is only a vague answer to the statement that a quantum object possesses both the nature of a point particle and the nature of a wave at the same time, so that one visualizes the quantum object as a matter wave with some sort of distribution in coordinate space as well as in wave vector space. Therefore, neither can one find a localized point quantum object, nor a strictly monochromatic wave motion in the quantum system.

Next we shall discuss the relations of the quantum state and the element in Hilbert space as well as the physical observables and the linear Hermitean operator. The connection between the eigenvalue problems and the preparation of a quantum state for a particular measurement of physical observables are also explored in depth. Dirac's notations of the bra and ket vector in Hilbert space are also reviewed and their connection with the  $q$ -representation and  $p$ -representation of wave functions are derived. The uncertainty relations can therefore formally be proved as the consequence of non-compatibility of two physical observables of position and momentum.

## Quantum Postulates and Associated Propositions

### 0th Postulate of Quantum Mechanics

For every quantum system, there exists an abstract entity, called the state (or state function or wave function in  $q$ -representation) which provides the information of all the dynamical quantities of the system; such as positions, momenta, energy, angular momentum, spin, charge, ... etc. All the possible states  $\psi, \varphi, \chi, \dots$  etc. of a given quantum system are elements of Hilbert space  $\mathcal{H}$ .

### 1st Postulate of Quantum Mechanics

For the measurement of each dynamical variable in the quantum system, such as the total energy of the system, or 3rd component of the orbital angular momentum ... etc., there associates a unique linear Hermitean operator in Hilbert space  $\mathcal{H}$  corresponding to each dynamical variable. The set of linear Hermitean operators in  $\mathcal{H}$  are called physical observables, or simply observables.

Dynamical variable	Physical observable
$E$ (energy)	$\rightarrow H$ (Hamiltonian operator)
$\ell_3$ (3rd component of $\ell$ )	$\rightarrow L_3$ (3rd component of angular momentum operator $L$ )

The physical quantity  $a$  corresponding to the observable of the quantum state  $\psi$  is obtained from the inner product of the order pair  $\psi$  and  $A\psi$ , where  $A$  is the corresponding linear Hermitean operator associated with dynamical variable  $a$ . The inner product  $(\psi, A\psi)$  is called the expectation value, or the average value associated with the dynamical operator  $A$  for state  $\psi$  of the quantum system.

$$a = (\psi, A\psi).$$

In general, the operator  $A$  acting on  $\psi$  will change  $\psi$  into another element  $\varphi$  in Hilbert space  $\mathcal{H}$ , which implies that the action of the measurement usually would disturb the quantum system and brings the original quantum state  $\psi$  into a new state  $\varphi$  by the external perturbation during the process of measurement, i. e.

$$A\psi = \varphi.$$

In a particular case, if an operator  $A$ , acts on  $\psi_a$  such that

$$A\psi_a = a\psi_a,$$

i. e. when  $A$  acts upon a particular quantum state  $\psi_a$ , the resultant state is the same as the one before except multiplying by a scalar number  $a$ , then it is said that the quantum system is prepared for the measurement of the dynamical variable associated with the physical observable  $A$ .

This particularly prepared quantum state  $\psi_a$  is called the eigenstate of the operator  $A$ . The scalar numerical value  $a$ , called the eigenvalue of operator  $A$ , is the result of the measurement for the dynamical observable.

Since the value of the physical measurement is always a real quantity, we therefore conclude the first Proposition:

**Proposition 1** *The eigenvalues for a Hermitean operator are real.*

For  $A = A^+$ , and letting  $\psi_a$  be the eigenstate of  $A$ , then

$$\begin{aligned} a(\psi_a, \psi_a) &= (\psi_a, A\psi_a) = (A^+\psi_a, \psi_a) \\ &= (A\psi_a, \psi_a) = (\psi_a, A\psi_a)^* = a^*(\psi_a, \psi_a) \end{aligned}$$

Since  $(\psi_a, \psi_a) \neq 0$ , we have  $a = a^*$ .

It is also ready to show Proposition 2.

**Proposition 2** *Two eigenvectors of the same Hermitean operator are orthogonal if the corresponding eigenvalues are unequal.*

Let

$$A\psi_1 = a_1\psi_1, \quad A\psi_2 = a_2\psi_2.$$

And if  $a_1 \neq a_2$ , then

$$\begin{aligned} (\psi_1, A\psi_2) &= a_2(\psi_1, \psi_2) = (A^+\psi_1, \psi_2) \\ &= (A\psi_1, \psi_2) = a_1(\psi_1, \psi_2). \end{aligned}$$

Therefore,  $(a_1 - a_2)(\psi_1, \psi_2) = 0$ , that concludes that  $\psi_1$  is orthogonal to  $\psi_2$  if  $a_1 \neq a_2$ . Namely  $(\psi_1, \psi_2) = 0$ .

With these propositions, we can formulate the 2nd postulate of quantum mechanics.

### 2nd Postulate of Quantum Mechanics

The set of eigenvectors  $\psi_a$  for all possible value  $a$  of a given Hermitean operator corresponding to a physical observable form the bases of a Hilbert space  $\mathcal{H}$ .

If there exists a complete set of linearly independent vectors  $\psi_a$  which are eigenvectors of both operators  $R$  and  $S$ , then the two corresponding physical observables,  $R$  and  $S$  are said to be compatible, we then have:

**Proposition 3** *If two observables are compatible, their corresponding operators commute, i. e. if*

$$R\psi_a = r_a\psi_a$$

$$S\psi_a = s_a\psi_a$$

*Then  $(RS - SR)\psi_a = [R, S]\psi_a = 0$ .*

*Any vector  $\psi$  in  $\mathcal{H}$  can be expressed in terms of linear combination of  $\psi_a$ , i. e.*

$$\psi = \sum_a \alpha_a \psi_a,$$

one can easily verify that it also satisfies  $[R, S]\psi = 0$  for any state  $\psi$  in  $\mathcal{H}$ .

Therefore, the observable  $R$  and the observable  $S$  commute.

From the point of view of physical measurement, the compatibility of two observables implies that one is able to prepare a single quantum system for the precise measurements of two different dynamical quantities corresponding to  $R$  and  $S$  respectively at the same time.

Unfortunately, not all physical observables are compatible. The position and the momentum are just an example, and the non-compatibility of position operator  $X$  and momentum operator  $P$  is in fact the foundation of quantum theory that forms the 3rd Postulate of Quantum Mechanics.

### 3rd Postulate of Quantum Mechanics

Every Poisson bracket in classical mechanics for canonical variables  $(p_i, q_i)$  is replaced by the commutator of the corresponding operators in the quantum system with the following relations

$$\begin{array}{ll} \text{Poisson bracket} & \text{Quantum commutator operator} \\ [q_i, q_j] = 0 & \rightarrow [X_i, X_j] = 0 \\ [p_i, p_j] = 0 & \rightarrow [P_i, P_j] = 0 \\ [p_i, x_j] = \delta_{ij} & \rightarrow [P_i, X_j] = \hbar/i\delta_{ij}. \end{array} \quad (5)$$

The non-commuting of the dynamical observables  $P_i$  and  $X_i$  will in fact not only lead to the fundamental quantization of the microscopic physical system, it also leads to the deviations in the position measurement and the momentum measurement. The non-commutativity of  $P$  and  $X$  in the 3rd Quantum Postulate signifies that one is not able to prepare a quantum state for the simultaneous measurement of momentum and position with absolute precision. Therefore, one shall regard the uncertainty relations of Eq. (4) as the direct consequence of this Quantum Postulate.

### Dirac's Bra and Ket Notations and the Realization of Wave Function

To understand that quantum uncertainty is the direct consequence of two non-commuting physical observables, it is convenient to introduce Dirac's notations of bra vectors and ket vectors. A ket vector, written as  $|a\rangle$ , represents the abstract quantum state  $\psi_a$ , with  $a$  inside the notation  $|a\rangle$  stands for the eigenvalue of the observable  $A$ , and the

eigenvalue equation becomes

$$A|a\rangle = a|a\rangle.$$

Similarly we denote the states  $\psi, \varphi, \chi, \dots$  etc. respectively by the ket vector notations by  $|\psi\rangle, |\varphi\rangle, |\chi\rangle, \dots$  etc.

The bra vectors are the elements in another vector space that are dual to ket vector space.

The bra vector takes the form  $\langle\psi|$ . The inner product in Dirac's notation is expressed by taking a pair of dual vectors, i.e. a bra vector  $\langle\psi|$  and a ket vector  $|\varphi\rangle$ , and putting them side by side as  $\langle\psi| \bullet |\varphi\rangle$  of which one always simplifies the notation as  $\langle\psi|\varphi\rangle$ . The inner product is a liner functional that defined a complex number for every ket vector in Hilbert space  $\mathcal{H}$ , such that

$$\langle\psi| \bullet (\alpha|\varphi\rangle + \beta|\chi\rangle) = \alpha\langle\psi|\varphi\rangle + \beta\langle\psi|\chi\rangle.$$

The set of bra vectors that define all linear functionals associated with the ket vector in  $\mathcal{H}$  forms a dual Hilbert space  $\widetilde{\mathcal{H}}$ . The inner product is a complex number, i.e.

$$\langle\psi|\varphi\rangle = \langle\varphi|\psi\rangle^*,$$

because bra vectors and ket vectors are elements of two different Hilbert spaces, and they are dual to each other. The inner product of  $\psi$  and  $\varphi$  denoted previously by  $(\psi, \varphi)$  is then replaced by  $\langle\psi|\varphi\rangle$ . The expectation value of the physical observable  $A$  becomes

$$\langle\psi|A|\psi\rangle = (\psi, A\psi).$$

The advantage of using Dirac's notations is in the construction of the Projection Operator, of which a vector in the vector space is projected onto a subspace. Let us denote the projector operator  $\wp_a$  by putting the bra vector followed by a ket vector immediately as

$$\wp_a = |a\rangle\langle a|,$$

therefore the vector  $|\psi_a\rangle$  is automatically constructed when the Projection operator  $\wp_a$  applies upon any vector  $|\psi\rangle$ , that is

$$|\psi_a\rangle = \wp_a|\psi\rangle = |a\rangle\langle a|\psi\rangle.$$

Let  $|i\rangle$  be the set of the eigenkets of observable  $A$ , with the discrete eigenvalue  $a_i$ , then

$$A|i\rangle = a_i|i\rangle.$$

It is often normalized to the eigenvector such that

$$\langle i|j\rangle = \delta_{ij}.$$

Any vector  $|\psi\rangle$  in  $\mathcal{H}$  can be expressed in terms of the base vectors with the coefficients  $\langle i|\psi\rangle$  in the following linear combination,

$$|\psi\rangle = \sum_i (\langle i|\psi\rangle) |i\rangle = \sum_i |i\rangle\langle i|\psi\rangle.$$

The summation  $\sum_i |i\rangle\langle i|$  in fact is just the direct sum of projection operator  $\wp_i = |i\rangle\langle i|$ , which obeys the following closure relation

$$I = \sum_i \wp_i = \sum_i |i\rangle\langle i|.$$

In many cases, the eigenvalue of observables becomes continuous. Taking observable  $X$ , the position operator for instance in the quantum system of interest, and for the sake of simplicity, we shall consider only a one-dimensional case, the eigenvalue equation of this quantum system reads

$$X|x\rangle = x|x\rangle,$$

where  $|x\rangle$  is the eigenket of observable  $X$  with eigenvalue  $x$  that is a continuous parameter representing the position of the quantum object.

The normalization of an eigenvector with continuous eigenvalue is taken as a  $\delta$ -function

$$\langle x|x'\rangle = \delta(x - x'),$$

and the corresponding closure relation becomes

$$I = \int |x\rangle dx \langle x|.$$

### Realization of $q$ -Representation in Quantum Mechanics

Let us introduce a unitary operator

$$U(P; \xi) = e^{-\frac{i}{\hbar} \xi P}$$

where  $\xi$  is a real continuous parameter and  $P$  is the momentum operator. It is unitary because

$$U^+(P; \xi) = e^{\frac{i}{\hbar} \xi P}$$

and

$$U(P; \xi) U^+(P; \xi) = U^+(P; \xi) U(P; \xi) = I.$$

If one expands  $U(P; \xi)$  in terms of a power series in  $P$ , i.e.

$$U(P; \xi) = I + \left(-\frac{i}{\hbar} \xi P\right) + \frac{1}{2!} \left(-\frac{i}{\hbar} P \xi\right)^2 + \dots,$$

and calculates the commutator  $[X, P^n]$  team by team, and makes use of

$$[X, P^n] = i\hbar n P^{n-1},$$

which is based on the 3rd postulate of Quantum Mechanics, then one obtains

$$[X, U(P; \xi)] = \xi U(P; \xi)$$

or

$$XU(P; \xi) = U(P; \xi)(X + \xi I).$$

It can be easily verified that  $U(P; \xi)|x\rangle$  is also an eigenket of  $X$  with eigenvalue  $(x + \xi)$ , i.e.

$$\begin{aligned} XU(P; \xi)|x\rangle &= U(P; \xi)(X + \xi I)|x\rangle \\ &= (x + \xi)U(P; \xi)|x\rangle \quad \text{and} \\ U(P; \xi)|x\rangle &= |x + \xi\rangle, \end{aligned}$$

where we have chosen the phase factor of the new eigenket  $|x + \xi\rangle$  to be 1.

Consider the matrix element of the operator, defined as  $\langle x|U(P; \xi)|x'\rangle$  for infinitesimal value of  $\xi$ , then by keeping up to first order in  $\xi$ , one has

$$\langle X|U(P; \xi)|x'\rangle = \langle x|x' + \xi\rangle \doteq \left\langle x \left| \left( I - \frac{i}{\hbar} \xi P \right) \right| x' \right\rangle.$$

With the properties of the  $\delta$ -function, one concludes

$$\begin{aligned} \langle x|P|x'\rangle &= \frac{\hbar}{i} \lim_{\xi \rightarrow 0} \frac{1}{\xi} \{ \delta(x - x') - \delta(x - x' - \xi) \} \\ &= \frac{\hbar}{i} \frac{d}{dx} \delta(x - x') = -\frac{\hbar}{i} \frac{d}{dx'} \delta(x - x'). \end{aligned}$$

Therefore one is ready to establish the connection between the axiomatic quantum postulates and the  $q$ -representation in quantum mechanics by defining the wave function  $\psi(x)$  to be the inner product of  $|\psi\rangle$  and  $|x\rangle$ , i.e.

$$\psi(x) = \langle x|\psi\rangle,$$

and  $q$ -representation of observable  $P$ , denoted by  $\wp$  leads to

$$\begin{aligned} \wp\psi(x) &= \langle x|P|\psi\rangle = \int \langle x|P|x'\rangle dx' \langle x'|\psi\rangle \\ &= -\frac{\hbar}{i} \int \frac{d}{dx'} \delta(x - x') dx' \psi(x') = \frac{\hbar}{i} \frac{d}{dx} \psi(x), \end{aligned}$$

namely, the  $q$ -representation of the momentum operator becomes the differential operator with respect to coordinate  $x$ . The quantum theory formulated in the previous

sections can be translated into the  $q$ -representation as follows

Abstract formulation	$q$ -representation
State $\psi$	Wave function $\psi(x)$
Observable $X$	Coordinate $x$
Observable $P$	Differential operator
	$\wp = \frac{\hbar}{i} \frac{d}{dx}$
Any observable $F(X, P)$	$F(x, \wp) = F\left(x, \frac{\hbar}{i} \frac{d}{dx}\right)$

### Fourier Transformation and $p$ -Representation

As we have mentioned a quantum system can be described either in the coordinate space of  $x$ , or in the momentum space (or wave number space) of  $p$ . The  $p$ -representation of the quantum state  $\varphi(p)$  is similarly defined as the inner product of the order pair vectors  $|p\rangle$  and  $|\psi\rangle$ , i. e.

$$\varphi(p) = \langle p|\psi \rangle .$$

It is of interest to relate  $\varphi(p)$  with  $\psi(x)$  by inserting an Identity Projection operator before the state  $|\psi\rangle$ , i. e.

$$\varphi(p) = \int \langle p|x'\rangle dx' \langle x'|\psi \rangle = \int \langle p|x'\rangle dx' \psi(x') . \quad (6)$$

Since

$$\begin{aligned} p\langle p|x \rangle &= \langle p|P|x \rangle = \int \langle p|x'\rangle dx' \langle x'|P|x \rangle \\ &= \frac{\hbar}{i} \int \langle p|x'\rangle \frac{d}{dx'} \delta(x - x') = -\frac{\hbar}{i} \frac{d}{dx} \langle p|x \rangle , \end{aligned}$$

one finds, with proper normalization, that

$$\langle p|x \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar} px} ,$$

and Eq. (6) can be expressed as

$$\varphi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int e^{-\frac{i}{\hbar} px} \psi(x) dx .$$

Conversely,  $\psi(x)$  is the inverse Fourier Transform of  $\varphi(x)$ , i. e.

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int e^{\frac{i}{\hbar} px} \varphi(p) dp .$$

### Quantum Uncertainty and Non-Compatibility of Observables $P$ and $X$

According to the 3rd Quantum Postulate that  $[P, X] = \frac{\hbar}{i} I$ , where  $I$  is an Identity operator, one can show that

the deviation operators, or the uncertainty operators of  $P$  and  $X$ , defined by

$$\begin{aligned} P_d &= P - \langle P \rangle I , \\ X_d &= X - \langle X \rangle I , \\ \text{where } \langle P \rangle &= \langle \psi | P | \psi \rangle \equiv \hat{p} , \\ \text{and } \langle X \rangle &= \langle \psi | X | \psi \rangle \equiv \hat{x} \end{aligned}$$

are the average value of position and momentum respectively and the commutator of these deviation operators  $P_d$  and  $X_d$  will also follow the same quantization rule, i. e.

$$[P_d, X_d] = \frac{\hbar}{i} I .$$

Let us introduce the anti-commutator operator, defined by

$$P_d X_d + X_d P_d = \{P_d, X_d\} , \quad \text{and denote it by } \Sigma ,$$

which is a Hermitean operator. Therefore, one can express the product operator of  $X_d$  and  $P_d$  as

$$\begin{aligned} P_d X_d &= \frac{1}{2} [P_d, X_d] + \frac{1}{2} \sum , \\ X_d P_d &= -\frac{1}{2} [P_d, X_d] + \frac{1}{2} \sum . \end{aligned}$$

If we define the ket vectors  $|\alpha\rangle$  and  $|\beta\rangle$  as

$$\begin{aligned} X_d |\psi\rangle &= |\alpha\rangle , \\ P_d |\psi\rangle &= |\beta\rangle , \end{aligned}$$

then the square of position uncertainty  $(\Delta x)^2$  as well as the square of the momentum uncertainty  $(\Delta p)^2$  can be expressed respectively as

$$\begin{aligned} (\Delta x)^2 &= \langle \alpha | \alpha \rangle = \langle \psi | X_d^2 | \psi \rangle , \\ (\Delta p)^2 &= \langle \beta | \beta \rangle = \langle \psi | P_d^2 | \psi \rangle . \end{aligned}$$

The quantum uncertainty relation takes the expression

$$\begin{aligned} (\Delta x)^2 (\Delta p)^2 &\geq \langle \alpha | \alpha \rangle \langle \beta | \beta \rangle \\ &\geq \langle \alpha | \beta \rangle \langle \beta | \alpha \rangle = \left\langle \psi \left| \left( -\frac{\hbar}{2i} I + \frac{1}{2} \sum \right) \right| \psi \right\rangle \\ &\times \left\langle \psi \left| \left( \frac{\hbar}{2i} + \frac{1}{2} \sum \right) \right| \psi \right\rangle , \\ (\Delta x)^2 (\Delta p)^2 &\geq \frac{\hbar^2}{4} + \frac{1}{4} \left| \left\langle \sum \right\rangle \right|^2 , \end{aligned}$$

where Cauchy-Schwarz inequality was applied to obtain the right hand side of the first inequality sign. Since  $|\langle \sum \rangle|^2$

is also positive definite because of the Hermiticity property of the operator  $\Sigma$ , we conclude that

$$\Delta x \Delta p \geq \hbar/2,$$

which is the quantum uncertainty in a one-dimensional case. The equality sign in the last equation holds if the conditions  $|\beta\rangle = \lambda|a\rangle$  and  $\langle\psi|\sum|\psi\rangle = 0$  are met. By making use of

$$\langle\psi_{\min}|[P, X]|\psi_{\min}\rangle = \frac{\hbar}{i} \quad \text{and} \quad \left\langle\psi_{\min}\right| \sum |\psi_{\min}\rangle = 0,$$

one obtains  $\lambda = \frac{i\hbar}{2(\Delta x)^2}$ , and the condition for the state function of the minimal uncertainty in a one-dimensional quantum system satisfies the following equations

$$\left( \frac{\hbar}{i} \frac{d}{dx} - \tilde{p} \right) \psi_{\min}(x) = \frac{i\hbar}{2(\Delta x)^2} (x - \bar{x}) \psi_{\min}(x),$$

which can be solved with the solution behaving as a traveling Gaussian wave packet as follows,

$$\psi_{\min}(x) = \frac{1}{\sqrt[4]{2\pi(\Delta x)^2}} e^{-\frac{(x-\bar{x})^2}{2(\Delta x)^2} + \frac{i}{\hbar} \tilde{p}x}.$$

## Conclusion

The uncertainties in quantum mechanics exist not only in the simultaneous measurement of the positions and the momenta as we have treated previously, uncertainties also exist in the simultaneous measurement of any pair of the dynamical variables that are not compatible, or their corresponding physical observables do not commute. For example, one is not able to obtain the data with absolute precision in the simultaneous measurements of any two components of the angular momentum  $L_i$  and  $L_j$ , because of the non-vanishing commutator  $[L_i, L_j] = i\hbar \epsilon_{ijk} L_k$ . If we denote the deviation operator for the observable corresponding to angular momentum by

$$L_d = L - \langle L \rangle I,$$

one can easily show that

$$[L_{di}, L_{dj}] = i\hbar \epsilon_{ijk} L_k,$$

and the square of uncertainty in the  $i$ th and  $j$ th components of angular momentum, denoted by  $(\Delta\ell_i)^2$  and  $(\Delta\ell_j)^2$  respectively can be expressed as

$$(\Delta\ell_i)^2 (\Delta\ell_j)^2 = \langle\psi|L_{di}^2|\psi\rangle \langle\psi|L_{dj}^2|\psi\rangle.$$

Because of the Hermiticity of  $L$ , and making use of the

Cauchy–Schwarz inequality, one can derive that

$$(\Delta\ell_i)^2 (\Delta\ell_j)^2 \geq \frac{\hbar^2}{4} \epsilon_{ijk} \tilde{L}_k^2 + \frac{1}{4} \tilde{K}_{ij}^2,$$

where  $\tilde{L}_k = \langle\psi|L_k|\psi\rangle$  and  $\tilde{K}_{ij} = \langle\psi|\{L_i, L_j\}|\psi\rangle$ .

The uncertainty in simultaneous measurement of  $L_i$  and  $L_j$  depends upon the average value of the physical observable of the third component of angular momentum, as well as a positive definite number  $\tilde{K}_{ij}^2$ . Therefore, the minimum of the uncertainty

$$(\Delta\ell_i)(\Delta\ell_j) \quad \text{becomes} \quad (\Delta\ell_i)(\Delta\ell_j) \Big|_{\min} = \frac{\hbar}{2} \left| \epsilon_{ijk} \tilde{L}_k \right|,$$

Therefore,

$$\Delta\ell_i \Delta\ell_j \geq \frac{\hbar}{2} \left| \epsilon_{ijk} \tilde{L}_k \right|,$$

which indicates that the uncertainty relation for the product of two components of a given angular momentum in a quantum system can not be independent of the average value of the third component. Many peculiar quantum phenomena that would not be happening in classical physics do exist in the subatomic quantum system. Quantum uncertainties corresponding to a pair of non-compatible observables can only be obtained through detailed investigation based on quantum dynamical theory.

## Future Directions

The limitations in pursuit of accurate scientific measurement of some data in quantum mechanical systems are the direct consequence of the quantum dynamics in the commutation relations among various physical observables. One is able to minimize but not to eliminate the uncertainty in the simultaneous measurement in any pair of dynamical variables of which the commutator of the corresponding quantum operators does not vanish. For instance, a plane wave of one-dimension can be prepared in a quantum system to achieve the utmost monochromatic frequency only at the expense of the total accuracy in locality. As we have treated in the previous section Heisenberg's Uncertainty Relations in position measurement and momentum measurement can only be minimized by preparing the quantum state of a traveling Gaussian wave packet. In fact we can easily show that the state function in momentum space is also of a Gaussian traveling wave packet due to the property that the Fourier Transformation of a Gaussian distribution is again a Gaussian distribution. Therefore, it is imperative that one should execute uncertainty control or error management before performing the measurements for physical observables of non-compati-

bility in the quantum mechanical system. As we have already demonstrated in the simultaneous measurement of any two components of an angular momentum, the uncertainty can be minimized by preparing a quantum state with the least mean value of its third component.

## Bibliography

1. de Broglie L (1923) Nature 112:540. (1924) Thesis, Paris. (1925) Ann. Phys 3:22
2. Merzbacher E (1970) Quantum Mechanics, 2nd edn. Wiley, New York
3. Jordan TF (1982) Linear operators for quantum mechanics, 2nd edn. Wiley, New York
4. Dirac PAM (1958) The principles of quantum mechanics, 4th edn. Clarendon Press, Oxford

## Granular Computing, Philosophical Foundation for

ZHENGXIN CHEN

Department of Computer Science, University of Nebraska at Omaha, Omaha, USA

### Article Outline

- [Glossary](#)
- [Definition of the Subject](#)
- [Introduction](#)
- [The Road to Granular Computing](#)
- [The Nature of Granule and Granule Computing](#)
- [Granule Measurement](#)
- [Granular Structure](#)
- [Granulation Provides a Unified View for Intelligent Problem Solving](#)
- [Relationship with Soft Computing and Natural Computing](#)
- [Relationship with Fundamental Issues of Computing and Complex Systems Problem Solving](#)
- [Summary](#)
- [Future Directions](#)
- [Acknowledgments](#)
- [Bibliography](#)

### Glossary

**Fuzzy set and fuzzy logic** Unlike a conventional set, in a fuzzy set, a fuzzy membership function is used to define the degree of an element belonging to the set. Fuzzy logic is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth as defined by membership functions.

Fuzzy logic contributes to the machinery of granular computing.

**Granular computing (GrC)** In a broad sense, granular computing is the general term referring to any computing theory/technology that involves elements and granules, with granule, granulated view, granularity, and hierarchy as its key concepts.

**Granular structure** Granular structure is a collection of granules in which the internal structure of each granule is visible.

**Granularity** The granularity of a level refers to the collective properties of granules in a level with respect to their sizes.

**Granulation** Granulation refers to the process of forming granules.

**Granule** As the fundamental concept in granular computing, a granule is a clump of elements drawn together by various criteria such as indistinguishability, equivalence, similarity, proximity or functionality.

**Hierarchy** In granular computing, hierarchy captures the ordering of levels.

**Neighborhood system** A neighborhood system of a point (an element) in the universe is the nonempty family of subsets (referred to as the neighborhood of that point) associated to it.

**Rough set** Rough set is a formal approximation of a conventional set, using a pair of sets as the lower and the upper approximations of the original set. Rough sets provide a single-layered granulation structure of the universe.

### Definition of the Subject

In a little more than a decade, granular computing (GrC) has emerged as a major research field for further abstraction, generalization and unification of granule-based key concepts traditionally scattered throughout a wide range of scientific disciplines, offering new opportunities for systematic studies of challenging computational issues cross multiple disciplines. An examination of foundations of granular computing, particularly on its philosophical dimension, is extremely important because it should reveal the hidden nature of granular computing, shed useful light for future directions and provide guidelines for researchers working in this area. In order to make this article useful to the research community in granular computing, instead of using philosophical/epistemological jargon, we will stay with basic terminology used in granular computing.

An important advantage of studying the philosophical foundation of granular computing is to identify what

is still missing. In particular, we point out that the future of granular computing is inevitably tied to important features of complex adaptive systems such as self-organization, learning, evolution and adaptation, as well as key issues in contemporary artificial intelligence (AI) such as embodiment, emergence and emotion.

The objective of this article is to provide a fair review of certain important aspects related to philosophical foundation of granular computing. We summarize and analyze existing literature to find out what has led to GrC, which ideas have been explicitly stated, which are implied, and which are still lacking (or overlooked). Due to the huge amount of literature in this area and diverse viewpoints of researchers, comprehensive coverage of all perspectives and approaches is impossible. Only selected materials of influential work are included. Materials presented in this article (including the author's own opinions, some of which have never been published before) is not necessarily an indication that they have been accepted by the GrC community in general or approved by any specific individual.

## Introduction

In order to answer the question of what is meant by “foundations of granular computing”, we need to answer the question: What constitutes the foundations of granular computing? Just like in the examination of foundations of data mining (where granular computing can play an important role in discovery of hidden knowledge buried in huge amounts of data) [8], foundations of granular computing can be examined from philosophical, technical, and social/application dimensions.

In general, philosophy of science refers to the study of philosophical assumptions, foundations, and implications of science. Topics to be studied include the character and development of concepts and terms, propositions and hypotheses, arguments and conclusions as they function in science; the manner and types of reasoning, the formulation, scope and limits of scientific methods, the implication of scientific methods and models, etc. Philosophy of science is closely related to epistemology, as a field focusing on the theory of knowledge it studies the nature, methods, limitations, and validity of knowledge and belief [6].

In order to study the philosophy of science, either as a whole or as it applies to particular fields, we have to analyze the contents of theories as well as practices of scientists working in these fields, to identify the important philosophical thoughts are implied. Therefore, study of philosophy of science is not confined to “professional”

philosophers; researchers in various science fields can (and should) take the philosophical aspects of their research discipline in their own hands. This is our attitude toward the study of philosophical foundation of granular computing.

All the issues related to the philosophical foundation of granular computing boil down to one question: *What is granular computing?* (Or: What is the *nature* of granular computing?) A simple question, with complex, possibly conflicting, yet incomplete answers! The philosophical foundation of granular computing requires a holistic examination on the foundations of granular computing as a whole.

Before we attempt to answer the magic question of what granular computing is, we have to address the issue of *how to* study it. Note that this is a multi-faceted question. First, the general methodology: Just like the study of philosophy of science in general, we have to take a retrospective/reflective approach to study what researchers have done in published literature related to granular computing: What is explicitly stated (i. e., what researchers in granular computing states what it is), what is implied (i. e., what researchers assume or think what granular computing might be), as well as what is yet to be addressed.

The next aspect of dealing with the “how to” question is to identify specific issues that need to be investigated. The answers to the question of “what granular computing is” can be studied from several aspects, in light of interests from the perspective of philosophy of science but stated in terms of granular computing:

- Character and development of concepts and terms, propositions and hypotheses:
  - Where is granular computing from? How does this new paradigm emerge (or converge from multiple predecessors)? In order to understand what granular computing is, we need a historical examination of the roots of granular computing.
  - Now that granular computing has established itself as a research discipline, we need to study its basic terms such as granule, granularity, granulation, etc. What does each of them mean? Without an understanding of these terms and what lies behind them, we will not be able to understand the nature of granular computing.
- Manner and types of reasoning, the formulation, scope and limits of scientific methods:
  - What are the major research issues of concern to the granular computing community? These issues form a set of descriptive features which collectively characterize what granular computing is.

- What kinds of problems can granular computing solve, and where are concepts such as granularity and granulation are useful? The universality and diversity of these problems not only demonstrates the power of granular computing, but also reveal the nature of granular computing from yet another perspective.
- Implication of scientific methods and models:
  - In order to understand the nature of granular computing, it is equally important to examine how it distinguishes itself from other related disciplines. This leads us to examine its relationship with research fields such as soft computing and cognitive informatics.
  - Finally, it is time to examine a more fundamental issue: How is granular computing related to basic issues of computational modeling and complex problem solving? These more advanced issues help identify where granular computing stands in the larger picture and helps our understanding of granular computing.

Answers to the questions listed above form the bulk of the rest of this article.

Note that the study of the philosophical foundation of granular computing is inevitably intertwined with other dimensions of the foundations of granular computing. For example, one topic of study in the philosophy of science involves the implications of scientific methods and models, along with the technology that arises from scientific knowledge, for society at large [5]. This justifies the connection of the philosophical foundation with the other two dimensions of foundations. In the case of granular computing, the study of philosophical foundations requires a close examination of the models and algorithms developed (the task of the technical dimension) as well as the implication of various applications (the task of the social/application dimension).

### The Road to Granular Computing

Granule, granularity and granulation are basic concepts involved in all aspects of human intelligence. First, it is in the language of our daily life, such as “extra fine granulated sugar”. More profoundly, these concepts also form the foundation of science and technology, including biology, mathematics, etc. For example, mathematics, as the study of the measurement, properties, and relationships of quantities using numbers and symbols, has been dealing with computation problems on such granules for thousands of years.

For more recent history, we may consider computerized information systems, such as database management systems (DBMS) or information retrieval systems (IR). For example, in a relational database, tuples form the basic granularity level for information storage and retrieval, yet from the transaction processing perspective, it is the entire relation, rather than its constituent tuples, which serves as the basic granule for a commit. In fact, the theory and practice of transaction processing is largely built around concepts of granularity and granulation: In traditional DBMSs, transactions are the granules to commit and abort, and in order to assure this, ACID (atomicity, consistency, isolation and duration) properties are required. In more advanced DBMS systems such as object-oriented DBMSs, due to the significantly increased complexity of transactions, sub-transactions may be used as granules for commit, and accordingly, some of the rigid ACID properties may be relaxed somewhat. As a main issue of transaction processing, concurrency control is another showcase of granularity-driven activities: Although many concurrency control protocols use individual data items as the granules to perform synchronization, multiple granularity has also been used, where different locking modes are imposed on data at different granularity levels of hierarchy, such as the entire database, a partition of the database, a file, a record, and so on. As for the physical storage implementation, B-tree (and its variations) is a popular indexing structure for storing individual key values for records (here the individual data points serve as granules), while R-tree (and its variations) is B-tree’s counterpart dealing with spatial data with two-dimensional minimum bounding boxes (mbbs) as granules for operation, the concept shared by many other proposed data structures for handling spatial data. Similarly, in information retrieval (IR), granules include document, chapters, sections, paragraphs, sentences, phrases, words, etc. (For basics of DBMS and a very brief overview on IR, see [36]).

Even today, an examination of relational databases from an information theoretic perspective [24] has been considered as an important piece of work related to the pre-history of granular computing. Yet, what is really interesting in today’s granular computing community is what is now usually referred to as the *information granule* [38], a concept emphasizing its role in *intelligent* information problem solving. Another important aspect distinguishing granular computing from other areas in information technology is that it goes beyond concepts related to granularity and directly addresses *computational aspects* related to granulations.

As noted in [1,2], human-centered information processing has been pioneered by Zadeh through his intro-

duction of the concept of fuzzy sets in the mid-60s. GrC arose as a synthesis of insights into human-centered information processing by Zadeh in the later 1990s and the term was coined by T. Y. Lin in 1997. Below we take a closer look at contributions of several research fields which led to granular computing.

According to Zadeh [43], granulation refers to partitioning (crisp or fuzzy) of an object into a collection of granules, with a granule being a clump of elements drawn together by various criteria such as indistinguishability, equivalence, similarity, proximity or functionality. For example, in rough set theory partitioning is based on the criterion of equivalence, while in granular computing partitioning is based on similarity, proximity or functionality (note that a rough set is a formal approximation of a conventional set, using a pair of sets as the lower and the upper approximations of the original set).

A granular variable is a variable which takes granular values. Granulation of a variable (or granular variable), such as age, goes from continuous, to quantized, and finally, becomes granulated. Several principal types of granules can be distinguished. It has been noted that in fuzzy logic, granulation is achieved through graduation – “In fuzzy logic everything is or is allowed to be graduated, that is, be a matter of degree or, equivalently, fuzzy” [44]. In addition, granulation of a function is done by summarization, as shown in rules such as “If X is large then Y is small” [45]. On the other hand, developments in rough set theory provide another example of the universality of the basic concepts of granule and granulation.

As pointed out in historic remarks [30], while fuzzy set theory and Dempster–Shafer theory are considered as examples of non-partition theories (where the universe is not partitioned), rough set theory and algebraic theory of relational databases in the 1980s are examples of partition theory (where the universe is partitioned into subspaces). Rough set theory is about equivalence relations which is a subset of binary relations studied earlier. Rough set theory is about partition which is a subset of covering. Later, the concept of neighborhood system (NS) was proposed, which is the generalization of the rough set concept in that it incorporates various concepts of topological space, covering, binary relation, as well as  $\alpha$ -cuts from fuzzy sets. Furthermore, each neighborhood can be regarded as a granule, and because of this, early granular computing research has the neighborhood system as its mathematical model. Therefore, the key in a road map from rough set to granular computing lies in the mathematical concepts of coverings, binary relations, and neighborhood systems, where a covering is a collection of granules (subsets) whose union is the whole space. Neighborhood sys-

tems have been used for the study of relational databases, and the granulation structures induced by neighborhood systems have been studied. A binary relation is equivalent to a special type of neighborhood system.

## The Nature of Granule and Granule Computing

### The Nature of Granule

The concept of granule has been studied in the AI community for decades [20]. According to [44,45], a granule is a clump of elements drawn together by various criteria such as indistinguishability, equivalence, similarity, proximity or functionality. In addition, attributes of a granule include probability measure, possibility measure, verity measure, length, volume, etc. There are numerous assumptions implied by various researchers regarding basic notions in granular computing, and the concept of granule in particular. In order to have a better understanding about the nature of granules, below we examine some of the hidden assumptions and provide a brief comparison.

- *Granule vs. words:* Traditionally, artificial intelligence (AI) has been considered as a research field of reasoning with individual symbols (a symbol is a token of meaning). These days, many researchers in computational intelligence have put much emphasis on computing with words (CW). As Zadeh noted, granular computing serves as a basis for the methodology of computing with words (CW) [44]. Although efforts have been made in this aspect, few researchers in granular computing have focused on issues related to linguistics. In current study of granular computing, semantics is demonstrated not through individual granules, but rather on the grouping of granules. Reference [1] proposed an information processing “pyramid” consisting of three layers organized by the size of information granules, with the lowest one concerned with numeric processing, the intermediate one concerning large information granules, and highest one devoted to symbol-based processing. But this is just one perspective.
- *Granule vs. attribute or entity (in database modeling):* In the entity-relationship approach, although we talk about attributes of information granules, such as size, capacity and dimension, in general, granules are not distinguished from each other mainly based on descriptive features (or attributes). Granules have low individuality: A granule is a primitive concept while entity is a first order citizen in data modeling. Although entities of the same type can form an entity set, entity sets are not mainly used to form hierarchical levels (rather,

they are the “end” products of conceptual modeling), but granulation is.

- *Positive granule vs. object in object-oriented (OO) paradigms:* In OO, objects are related through super-class/class/subclass relationships and properties can be inherited or overridden. The OO paradigm has an anthropomorphic nature. On the other hand, granule is a general concept directly tied to computational aspects; although granules can also be related together through hierarchies, the relationship between various levels in the hierarchy is much more complex (see later discussion). In addition, although hierarchies play a key role in granular computing, other forms of relationships exist between granules (such as the volume of a cylinder and the surface of that cylinder).
- *Positive granule vs. cluster (in data analysis/data mining):* Similarity in cluster analysis is usually determined by topology which is largely absent in granular computing; i. e., distance is usually not used as the basic factor to relate individual granules.
- *Granule vs. system vs. graph:* Granularity is a measure of the size of the components, or descriptions of components, that make up a system. A system, either natural or man-made, is a set of interacting or interdependent components forming an integrated whole. A system can also consist of a number of subsystems. Ideas from systems theory have grown with diversified areas, exemplified by the work of B. Banathy, as well as others. Systems described in terms of large components (i. e., larger granularity) are referred to as coarse-grained, otherwise they are termed as fine-grained. Since graph data structure is a useful tool to describe systems, it can also be a useful modeling tool for the study of granules (see ► [Granular Computing: Practices, Theories, and Future Directions](#) for the modeling examples). However, since system theory is devoted to the study of inter-relationship of its components to optimize system performance, it does not address the majority of issues of concern in granular computing.

### The Nature of Granular Computing

According to [44,45], the rationale of granulation can be either imperative (forced) or intentional (deliberate), when precision is costly and there is a tolerance for imprecision. We can distinguish different forms of granulation, such as spatial (including interval), temporal granulation, etc. Granular Computing is an emerging conceptual and computing paradigm of information processing. It has been motivated by the urgent need for intelligent processing of empirical data, that is now commonly

available in vast quantities, into humanly manageable abstract knowledge. In this sense, granular computing offers a landmark change from the current machine-centric approach to information and knowledge to a human-centric approach [1,2,3]. According to [31] by T. Y. Lin who first coined the term granular computing, any computing theory/technology that involves elements and granules (subsets or generalized subsets) may be called granular computing. Intuitively, elements are the data, and granules are the basic knowledge. This simple definition has the advantage of unifying research activities of researchers from diverse backgrounds. In addition, this simple definition allows us to exploit the interesting idea of *granulate and conquer*, a softer version of classical divide and conquer; fuzzy granulate and conquer is the cornerstone of the success of fuzzy controls. A very common technique used in the classical “non-partitioning” recursive call is dynamic programming [17].

According to [1], the fundamental features of GrC consist of the following: allow for multiple abstraction levels (granularity levels), allow for several methods of traversing various levels of hierarchy (referred to as encoding-decoding mechanisms), and allow for non-homogeneous methods. In addition, Zadeh [45] noted that granulation is a core concept surrounded by natural language computation, rough set theory, computational theory of perceptions, as well as granular computing. Therefore, there is a simple formula of “granular computing = ballpark computing”.

As an example of recent development in regard to basic concepts in GrC, the study in [15] considered two granular worlds which are defined on the same universe of discourse but employ different granulations of this universe. In order to translate information from one granular world to the other, there is a need to regranulate the information so that it matches the information granularity of the target world.

### Granule Measurement

As indicated earlier, according to [44,45], a granule is a clump of elements drawn together by various criteria such as indistinguishability, equivalence, similarity, proximity or functionality. In addition, attributes of a granule include probability measure, possibility measure, verity measure, length, volume, etc. Therefore, granules can be considered as knowledge, compared with constituting elements which represent data [27]. In addition, granules may be interpreted from three semantic views: a granule is a unit of basic knowledge; a granule is a unit lacking precise knowledge; and a granule is a sub-problem of com-

puting (see ► [Granular Computing: Practices, Theories, and Future Directions](#)). We will get back to the last perspective in a later section. Here we examine the first two perspectives: Granules are knowledge units, yet consist of uncertainty. Together, these two perspectives imply an intrinsic property associated with granules: Measurement, which refers to the estimation of the magnitude of some attribute of an object, such as its length or weight, relative to a unit of measurement. Consequently, granularity is a measure of the size of the components. In the simplest form, measures can be made through the size of a granule, which indicates the degree of abstraction, concreteness, or detail, of the granule. In the set-theoretic setting, the size of a granule can be the cardinality of the granule. Various uncertainty theories related to granular computing, such as fuzzy set theory or rough set theory, all have specific ways to deal with measurement.

The need for measurement in studying granules implies the importance of dealing with topology, a key concept underlying many subfields in mathematics, such as mathematical analysis. Take a look, for example, at the well-known *Heine–Borel theorem*: Each open covering of a closed and bounded set of real numbers has a *finite* subcovering. The granules involved here include the closed and bounded set, the elements in the original open covering, the elements in the resulting subcovering, etc. Therefore, from a granular computing perspective, we may say that the Heine–Borel theorem is concerned with *reduction* from infinite to finite, or *mapping* from an infinite number of granules to a finite number of granules.

There are various kinds of granules, and they can be related through hierarchy or other relationships or through aspects such as measurement. For example, a granule could be 3-dimensional (3D) object, or could be the surface of this object. Yet the calculation of the volume of a 3D object can be converted to the calculation of the area of the surface through various theorems developed in mathematical analysis. Here the computation problem (i. e., the volume) on one granule (i. e., the 3D object) is *converted* to another computation problem (i. e., the area) on another granule (i. e., the surface of the 3D object). This is exemplified in *Green's theorem*, a special case of the more general Stokes' theorem.

As a more recent example concerning the importance of measures for granules, let us consider the case of relational database design theory as well as its contemporary extension to XML file design. Relations (or flat tables) are the most visible granules in relational database design. But what should comprise a relation? This issue has been addressed by normalization theory, where important normal forms such as Third Normal Form (3NF) or Boyce–

Codd Normal Form has been developed, based on the notion of functional dependency and related concepts. But how could this kind of result be generalized to the case of semistructured XML files? A recent study by Libkin and his research group [25] addressed this issue by proposing to use entropy-based measures to determine the quality of the granulation for relational database design, and to further apply this principle to XML file design. First the concept of relative information content (RIC) was proposed, subsequently other related measures were proposed, such as the Price. Furthermore,  $\text{Price} (3\text{NF}) = 1/2$ . Therefore, the conformation of a “good” granule (such as a relational table) is justified by a quantified approach. Although these studies are conducted by researchers outside of GrC community, these results can shed new insight for research in GrC.

## Granular Structure

### Quotient View – Information Hiding

As noted in ► [Granular Computing: Practices, Theories, and Future Directions](#), granular structure is the collection of granules in which the internal structure of each granule is visible. In this sense, granular structure is a collection of white box granules. In contrast, a quotient structure, as derived from the mathematical concept of quotient space, refers to the collection of granules in which each granule is regarded as an element (or point) of a set, and the “intersections” among granules are abstract to interactions among points (or elements). Therefore, a quotient structure is a collection of black box granules. This perspective provides a quotient view on a hierarchical structure consisting of granules at different levels of detail. With details encapsulated at lower levels while only abstract information is shown on higher levels, quotient view-based granular structure supports information hiding.

Earlier, in the Introduction section, we mentioned that DBMS is a major source behind GrC. In fact, database design is a perfect example of information hiding: details of the instances are hidden at the more abstract levels of a schema. The success of the well-known entity-relationship approach lies on the information hiding by focusing on schemas (while not completely throwing away the instances, such as the notion of mapping cardinality involved in relationship sets). More recent developments in management of semistructured data necessitates the need for XML schema discovery (see, e. g., [12,19]), which in essence is to identify the quotient structure and support information hiding.

The theory of hierarchy provides a multi-layered framework based on levels. Mathematically, a hierarchy

may be viewed as a partially ordered set. Two types of partial ordering can be distinguished, i.e., strong vs. weak dependencies: A neighborhood system X is weakly dependent on Y if every neighborhood of X is a subset of Y; or strongly dependent if every neighborhood of Y is a union of the neighborhood of X [26,27,33]. Note that strong = weak in the case of partitions. (More discussion on hierarchy will be given later in the section on emergence.)

A granule in a lower level schema may be a more detailed description of a granule in a higher level with added information. In the other direction, a granule in a higher level is a coarse-grained description of a granule in a lower level omitting irrelevant details. Note that although we are talking about granules of various granularities connected together in a hierarchical manner, granules are related to each other in a complex network. In this sense, what we call hierarchy in GrC is actually a kind of “information hiding” with detailed network information filtered out.

A taxonomy of types of granularity has been discussed [23]. First, one can identify the main differences in types of granularity based on:

- (1) Whether scale matters: Arbitrary scale vs. non-scale-dependent granularity;
- (2) How levels (and their contents) in a perspective relate to each other;
- (3) How perception and (mathematical) representation are related, such as based on set theory or mereology.

The philosophical implication of this taxonomy has also been examined; for example, it is noted that the difference between scale and non-scale dependency and their formal representations roughly fits with Sowa’s epistemic and intentional granularity [39] based on Peirce’s three categories.

### **Granular Computing and Ontologies**

The concept of hierarchy plays a key role in many other contexts studied in last decades, including frame systems in artificial intelligence and object-oriented approaches, just name a few. Here we take a look at the recent surge of research related to ontologies, an issue originated in philosophy [6] that has received ever-increasing attention in recent years from a much wider community.

Even though hierarchies play a significant role in both granular computing and ontologies, the relationship between these two research fields is not well-studied, and this is likely due to the fact that although granular computing sets place emphasis on both computation and semantics, the study of ontologies is largely “semantics without com-

putation”. Nevertheless, ontological aspects have been occasionally studied by granular computer researchers, including the work on taxonomy of types of granularity summarized in an earlier section [23]. In addition, granular computing can be used to aid the study of ontologies. One example is depicted in [35] where the concepts of domain granule and domain granule lattice were introduced, and an algorithm for generating the lattice was proposed to capture the ontology. The relationship between granular computing and ontology can take the opposite direction as well; for example, [48] described a granular computing model based on ontology.

In addition, [22] presented a rough-granular approach which incorporates domain knowledge. This additional knowledge, represented by ontology of concepts, is used to make it feasible to search for features (condition attributes) relevant for the approximation of concepts on different levels of the concept hierarchy defined by a given ontology. Yet, the question of how ontologies of concepts can be discovered from sensory data remains as one of the greatest challenges for many interdisciplinary projects on learning of concepts.

### **Information Hiding and Emergence**

A basic task of granular computing, granularity conversion, is to change views with respect to different levels of granularity (with changing details). However, it would be naive to believe that relationships between different levels of granulations can be described by mappings.

Information hiding as demonstrated in a granulation hierarchy may shed important light on revealing the nature of emergence [21], where complex phenomena are produced by interaction of simple, tiny things. Think about the operations of zoom in and zoom out: Just like using zoom lenses on a camera, we can show a smaller area of an image at a higher resolution or magnification or a larger area at a lower resolution or magnification. Note that there are apparently two kinds of “zoom in”: For example, when we are examining an online interactive map, we may click a particular country on the initial world map, then click a particular region or state, then click a particular city, and so on, to take a close look at particular streets or buildings in which we are interested. On the other hand, sometimes the so-called “zoom in” only gives us larger prints of the original map without any additional information added: Although this latter type of “zoom in” is useful to senior citizens with vision problems, it is useless for most people. Similarly, we may have two types of zoom out: We can get the entire picture as a whole rather than focusing on a particular part of the picture (with all the details kept),

or we summarize or abstract information not available at the more detailed level.

Apparently, zoom in and zoom out are directly concerned with getting information at different granulation levels. When we “zoom out” from the lower level, we may not necessarily just lose the details at the lower level, but we may also gain insight not demonstrated at the lower level. As a typical example, think about the painting “Gala Contemplating the Mediterranean Sea which at Twenty Meters becomes a Portrait of Abraham Lincoln” by Spanish surrealist Salvador Dali. As the title suggests, at the detailed level is Gala’s picture and the surroundings, while at a coarse granulation level, Lincoln’s portrait emerges. This is where the excitement of emergence comes from, and granular computing can play a critical role in finding hidden rules which produce the emergent phenomena.

Due to the complexity of the networks in which granules are interrelated to each other, it would be questionable to assume that an algorithmic study of procedures for constructing granules and related computation is a universal and realistic approach in GrC. Just like the more general study of complex adaptive systems, the tasks faced in the GrC community are closely related to the task of understanding the nature of emergence. As the movement from low-level simple rules among granules to higher level sophistication, emergence is a ubiquitous feature of the world around us. The hallmark of emergence is the sense of much coming from little. Complex systems present behaviors by drawing upon populations of relatively “unintelligent” individual agents, rather than a single, “intelligent” agent. Yet, as John Holland noted, regardless of how mysterious it might appear, emergence can be studied as a scientific discipline because low-level rules and laws can be discovered, and the movement from low-level rules to higher level sophistication is what we call emergence. This view suggests a hierarchy of levels from which emergence takes place, and the phenomena of emergence are typically bottom-up in nature. Holland underlines a number of features of emergent (or complex) systems and suggests that emergence is a product of coupled, context-dependent interactions. These interactions, along with the resulting system, are nonlinear, where the overall behavior of the system cannot be obtained by summing up the behavior of its constituent parts. Persistent patterns at one level of observation can become building blocks for persistent patterns at still more complex levels. At each level of observation the persistent combinations of the previous level constrain what emerges at the next level. As Holland put it, “This kind of interlocking hierarchy is one of the central features of the scientific endeavor. It will lead us into, and out of, the thorny thicket known as reduction – roughly, the

idea that we can reduce explanations to the interactions of simple parts”. Yet, there are no systematic ways described there on how to discover much needed laws or rules so that emergence can take place.

Think about one of the most mysterious things we can ever have, namely, life. Regardless whether we are talking about how the first life appeared on the Earth millions or billions of years ago, or how a particular biological species (such as rhino or elephant) first evolved from other species, or how a particular baby named John J. Johnson was conceived by his parents in the last year, a life always emerges from a hierarchy of extremely complex interactions, starting from the lowest level (where the granules are the chemical elements) all the way up along the hierarchy.

An important lesson learned from making a life is: Emergence is a bottom-up process, and top-down is not symmetrical to bottom-up. A top-down analysis of a new born baby (using a partitioning approach) is the task of anatomy, but this will not reveal how he or she is made. In other words, although top-down lets us examine the product, bottom-up reveals the process. This justifies the need for both top-down and bottom-up directions of research in GrC.

Just like the study of emergence, the study of granular computing is intended to provide a unified approach of studying complex phenomena across various domains. In addition, just like the case of emergence, granular computing models are made concrete in specific application domains. Although emergence is a complex issue and has not caught much attention by researchers in the GrC community, certain aspects have been addressed under the notion of information integration ► [Granular Computing, Introduction to](#). More systematic research along this direction is needed. Below we discuss several notable points.

- *Interaction among granules:* The interactions of granules were the beginning of GrC. The issues in knowledge representations, additive measures, belief functions (non-additive measures) and approximations were discussed in [26,27]. In addition, although the Choquet integral [37] usually is not considered within GrC realm, it is also based on belief functions. A Choquet integral is a way of measuring the expected utility of an event where the event is uncertain, and is applied specifically to capacities. The Choquet integral focuses on interactions among attributes, resulting in non-additive measures. Therefore, the contribution of two attributes to a classification variable is more than the sum of the contributions made by these two attributes alone.
- *Relationships of levels within a hierarchy:* In granular computing, granules in a higher level may have greater

integrity and higher bond strength than those in a lower level. The structures need to be fully explored to establish a basis of granular computing. However, in the study of emergence, levels in the hierarchy interact in a much more complex way; for example, a completely new form may take place, dramatically different from what can be found at a lower level (such as life coming from non-living chemical materials). In this respect, the relationship should be examined between granular computing and mereology, which is a collection of axiomatic formal systems dealing with parts and their respective wholes and the parts of the parts in a whole.

- *Other important features:* Other important features demonstrated in complex systems, such as self-organization, evolution, adaptation, etc., are not well-supported by current studies in granular computing. However, this situation seems to be changing. For example, [16] described a self-learning algorithm for uncertain information processing.
- *Building blocks:* Both granular computing and the study of emergence have emphasized the importance of building blocks for model construction but for different purposes.

### **Granulation Provides a Unified View for Intelligent Problem Solving**

Granular computing and related concepts play an important role in a wide range of intelligent problem solving tasks, including data mining, which is intended to discover implicit, interesting knowledge patterns from massive data [18]. We now take a look at this issue, which is addressed through three tiers, from the most basic form to the most advanced form, and accordingly this section is divided into three subsections. We first present how to *interpret* problem domains in terms of concepts in granular computing, and then discuss how to *cast* issues of intelligent problem solving in terms of granular computing. Neither of these two tiers involves any specific algorithms or methods developed in granular computing. Finally, we take on the issue of granular computing algorithms using granular computing for data mining as an example, which is further illustrated by an implemented case study. Our discussion through these tiers demonstrates that granular computing and its related concepts serve as a unified view for intelligent problem solving in various degrees.

### **Re-Examination of Existing Studies from a GrC Perspective**

Although granular computing should not be simple re-statement of existing results, re-examining or re-interpret-

ing existing results from the standpoint of granularity can help reshape the problem definition and establish a new start of granular computing-oriented solutions – so long as our investigation does not stop at the re-interpretation.

Here we use the case of bioinformatics as an example (note this is not intended as a general discussion on the relationship between granular computing and bioinformatics). Various subjects studied in bioinformatics, such as sequence, block, pattern, motif, and protein, can all be viewed as granules at different levels. Below we describe some sample scenarios to illustrate the applicability of GrC in bioinformatics. The key idea behind our discussion is that the relationship between GrC is twofold: GrC can be applied to shed light on bioinformatics problem solving, and bioinformatics also poses interesting challenges for GrC. Materials presented here are mainly taken from resources cited in [11].

- *Granulation construction under constraints:* The complexity involved in structural bioinformatics problem solving can take advantage of techniques developed in granular computing, and demand new solutions yet to be developed in granular computing. As discussed before, one of the basic concepts in granular computing is granulation. An essential question related to granulation is how to construct higher-level granulation from lower-level granules. As an example, here we examine the fragment assembly problem, which arises when a DNA sequence is broken into small fragments. These fragments must then be assembled to reconstitute the original molecule. (In a sense, the fragments form a covering in the context of GrC). The importance of this problem lies in the fact that with current technology it is impossible to sequence directly contiguous stretches of more than a few hundred bases. On the other hand, there is a technology to cut random pieces of a long DNA molecule and to produce enough copies of the pieces to sequence. Therefore, a typical approach to sequencing long DNA molecules is to sample and then sequence fragments from them. Solving this complex problem consists of assembling a collection of fragments coming from a long, unknown DNA sequence into a correct order and orientation. This can be cast as a problem in GrC, which is concerned with construction of larger granulation from smaller granules *under certain constraints* (such as imposing or preserving a certain order and orientation).

- *Characterization of granulation:* Once granulations have been constructed, the next question is to uncover the basic features of these granulations (namely, find

an *interpretation* of the constructed granulations). As an example, consider the following. Reference [37] described a data-mining algorithm for discovering regions of locus control, i.e., these regions that are instrumental for activating genes. One type of such elements of locus control is the matrix attachment regions (MARs). The discovery of MARs is based on the observation that a group of patterns are bonded together by the virtue of their similar function. After such a grouping, a search for the patterns in a given group can be performed to identify three regions in the query DNA sequence. If a large subset of members of a functionally related group of patterns is found in a given region of the DNA sequence, one can justifiably classify it as an MAR. The detection problem is to identify observation within a given region. Therefore, discovery of MARs is actually done by grouping for granulation (here in the form of patterns) and then finding the statistical characteristics of those granulations – somewhat similar to solving a mathematical equation. A more general indication of this approach is the task of studying *characterizing the granulations*, namely, identifying the significant features of the constructed granulations. Although characterization rule mining has been studied by many researchers in the data mining community [16], the task described in the above example goes beyond the reach of many existing approaches in characterization of rule mining, because the granulations considered here may be constructed in a dynamic manner. Therefore, a study in characterization of granulation will enrich both GrC and basic theory for characterization rule mining.

### Forming Intelligent Problem Solving Tasks in Terms of Concepts in Granular Computing

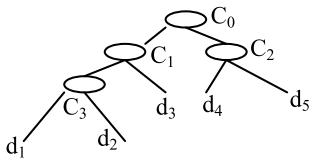
We now examine several issues of intelligent problem solving.

**Intelligent Query Answering** Intelligent query processing has very rich contents. GrC (integrated with appropriate AI and data mining techniques) can contribute to intelligent query answering involving aggregate data, mainly due to its power of grouping data in a dynamic and flexible manner. The concepts of constructing equivalent classes, partitioning and covering, as widely used in GrC literature, can be incorporated into the discovery of query-relevant rules from the database. In addition, GrC can be more directly absorbed into presentation techniques applied on original extensional answers to produce final answers with an intensional flavor. Here we take a look at the issue of

answering queries at the right granulation levels as originally presented in [10].

Although discovery of query-relevant rules is an interesting direction of study, in many other scenarios we may want to generate aggregate answers from existing extensional answers. Comparing with the previous topic, this direction can be considered as a “low key” solution; nevertheless, it deserves equal attention due to its practicality. As the online availability of data is exploding, business firms have begun to use such data to help to *make better decisions* for their business. For example, an auto manufacturer may want to know which type of vehicles made the highest profit in the previous year, or which generation of people could be the ideal target population for auto purchase and what their “tastes” may be. For such inquiries, conventional answers, usually extensional answers consisting of all the involved individual tuples, are not the most appropriate answers. Frequently, people feel the need for better decision support based on data analysis and knowledge discovery from the primitive data. As a more concrete example, suppose the Board of Directors of a retail firm wants to know something about the salary levels of its employees. A question could be: “*In our company, what kind of people are making more than \$ 100 000 a year?*” We may expect to receive several types of answer to this query. An answer could be “All directors and most managers”. It gives some qualitative description about the common characteristics of those qualified individuals. An alternative way is using an aggregate expression. An *aggregate expression* is a sequence of terms that are in the format of “ $r/t C$ ”, where  $C$  represents a concept with a total number of  $t$  individuals, while  $r$  is the number of these individuals who belong to the answer. As to the above query, an aggregate answer could be “8/8 directors + 27/32 managers”. From this example it can be seen that an aggregate answer is superior to the qualitative answer “all directors and most managers”, in that the aggregate answer not only covers all the information released in the latter, but also provides more quantitative information towards the overall picture of the database. In general, there exists more than one answer to a query. The question is how to obtain the best (optimal) one from those candidate answers.

An approach of using *query entropy* (or  $Q$ -entropy for short) for intelligent query answering concerning aggregate answers (with a prototype experiment developed) has been developed for answering aggregate queries at the appropriate data granulation levels [10]. The particular aspect we are looking for is that, given a conceptual hierarchy  $C$  and a query  $Q$  against a database  $D$ , when multiple answers exist at different granulation levels, how do we evaluate them and select the best answer? This ap-



**Granular Computing, Philosophical Foundation for, Figure 1**  
A bi-level taxonomy To of concepts

proach can be examined from a granular computing perspective.

An expression is a collection of terms. A term is represented by the ratio of qualified individuals and total individuals in that concept. The length of an expression measures conciseness, while the entropy of the expression measures preciseness. An expression is said to be the optimal answer to a given query if it has the lowest entropy value. For a real world problem, to obtain the optimal answer to a query, a hierarchy has to be built first from the primitive information stored in a given database.

Shannon entropy and associated conditional entropy are widely used today in information theory. Although the intuition to apply the entropy concept in information theory to intelligent database query answering is well-known, practical application of entropy theory is hindered by the inherent computing burden of Shannon and conditional entropy. One has to be very clear with taxonomy in order to obtain the probability of the event that an individual in a concept belongs to the extensional answer and the probability of the event that a randomly chosen individual belongs to a concept.

To take advantage of the approximation of Shannon entropy and the simplification of conditional entropy, the concept of  $Q$ -entropy was introduced. Let  $R$  be a finite space consisting of mutually disjointed subsets  $R_1, R_2, \dots, R_k$  with cardinalities  $|C_1|, |C_2|, \dots, |C_k|$  respectively. The probability of the event that an individual in set  $R_i$  is a qualified individual is denoted as  $r_i$ . Then the  $Q$ -entropy is defined as:

$$Q = \sum_{i=1}^k |C_i| * G_i = \sum_{i=1}^k -|C_i| * \left(r_i - \frac{1}{2}\right)^2.$$

In addition to concepts  $C_i$ , we also use  $d_i$  to denote extensional answers (which are actual tuples). As an example, consider a taxonomy  $T$  of four concepts with  $C_0 = \{d_1, d_2, d_3, d_4, d_5\}$  as the root concept, and let  $C_1 = \{C_3, d_3\}$ ,  $C_2 = \{d_4, d_5\}$  and  $C_3 = \{d_1, d_2\}$ , as shown in Fig. 1. (As a more concrete example,  $C_0$  would be employee,  $C_1$  could be worker,  $C_2$  could be manager, etc.).

There are three candidate expressions:

- I) “ $1/2 C_3 + 2/2 C_2$ ”;

- II) “ $1/3 C_1 + 2/2 C_2$ ”;  
III) “ $3/5 C_0$ ”.

Calculating  $Q$ -entropy using the formula introduced above, we have  $Q_I < Q_{II} < Q_{III}$ . Therefore the best answer, (I), is selected.

The  $Q$ -entropy approach for query answering was developed prior to our study of GrC. Nevertheless, we can re-examine this approach from a granular computing (GrC) perspective.

We first note that partition can be viewed as a special kind of granulation. However, in classification rule mining, partition is conducted on training samples. In contrast, in  $Q$ -entropy approach partition is conducted on extensional answers for queries. We also note that in the  $Q$ -entropy approach for any two aggregate answers of the same query, one answer may not necessarily be a refinement or coarsening of the other. For example, suppose in Fig. 1  $C_2$  also has a child concept  $C_4$ . Then the answer “ $1/2$  of  $C_3$  and  $1/2$  of  $C_4$ ” is not a refinement or coarsening of the answer “ $1/4$  of  $C_2$  and  $1/4$  of  $C_2$ ”.

The approach of using  $Q$ -entropy for selecting appropriate levels of query answering also differs from classification rule mining in several other aspects, including size of granules, the construction of the tree, the way in which entropy is used, and so on. Because of these features, a GrC perspective can be incorporated into presentation techniques applied on original extensional answers and can produce final answers in an intensional flavor.

## Granular Computing and Data Mining/OLAP

**Relationship Between These Two Fields** The interesting relationship between granular computing and data mining can be further examined in many different aspects; here we take a look at just one example of granular operators for property preservation. Granulation allows different representations of the same problem in different levels of detail. It is naturally expected that the same problem must be consistently represented. Granulation and its related computing methods are meaningful only if they preserve certain desired properties. For example, [46] studied the “false-preserving” property, which states that if a coarse-grained space has no solution for a problem then the original fine-grained space has no solution. Such a property can be explored to improve the efficiency of problem solving by eliminating a more detailed study in a coarse-grained space. In the context of hierarchical planning, one may impose similar properties, such as the upward solution property, the downward solution property, monotonicity, etc. This property resembles a priori principle as discussed in the context of asso-

ciation rule mining [18], but significantly differs from the view of new properties appearing at higher levels of hierarchies which are not found in lower levels in the study of emergence [20].

**Granular Computing as a Basis for Data Mining and OLAP** Data mining from databases [18], a key step in the knowledge discovery from databases (KDD), is intended to discover hidden knowledge buried under the ocean of data. As indicated in [33], “In essence data mining may be viewed as a form of summarization of very large datasets, while granular computing may be viewed as operations on summaries of small datasets. The common rule of summarization in data mining and granular computing is the principle reason why granular computing is of high relevance to data mining”. For example, in [29], Lin showed that generalized association rules can be expressed as union of basic granules. More recently, [32] further explored the issue on deductive data mining using granular Computing.

In a sense, data mining is finding hidden equivalence relations. For example, cluster analysis aims to find clusters, which are a sort of equivalence relation, so at the cluster level only clusters (rather than original data points) are seen – a typical case of information hiding. Similarly, characterization or discrimination analysis [18] aim to identify dominant features to characterize target class or to discriminate target class versus comparison class (with details of attribute values removed) – again a typical example of information hiding. An example of the impact of granulation size to data mining is multi-level association rule mining [18], which results in revised algorithms and reduced support/confidence level at coarser granulation levels.

Related to data mining is On-Line Analytical Process (OLAP) technology. Unlike data mining, however, OLAP does not resort to a reasoning process; rather, it is more closed tied to database and data warehouse technology for analyzing summary and historical data. Since information granules are formed from data granules, perspectives from GrC also contribute directly to OLAP. In fact, OLAP operations such as rollup and drill down are directly concerned with granulation sizes. In addition, OLAP techniques have been coupled with data mining, giving GrC more room to maneuver. For example, as discussed in [18], complex aggregations using multi-featured data cubes can facilitate data mining type queries to allow computation of aggregates at different granularity levels. In addition, a unified On-Line Analytical Mining (OLAM) has been proposed [18]. Additional discussion about this integrated data mining/OLAP and the role of GrC can be found in [7,9,47].

## Relationship with Soft Computing and Natural Computing

*Relationship with soft computing:* Probably, the most ancient granule is the intuitive infinitesimals (which refers to the idea of objects so small that there is no way to see them or to measure them). It was the inspiring source of calculus and Zeno’s stationary and moving paradox. A granule is an atom of uncertainty (see ► [Granular Computing: Practices, Theories, and Future Directions](#) and [31]). Yet research related to uncertainty in computational research has focused on the technical side rather than on the philosophical side; therefore, the bulk of research on uncertainty has been conducted in the realm of soft computing rather than granular computing. Soft computing has been defined as an association of computing methodologies which includes its unique, complementary and symbiotic constituent members fuzzy logic, neurocomputing, evolutionary computing and probabilistic computing. Soft computing differs from conventional (hard) computing in that, unlike hard computing, it is tolerant of imprecision, uncertainty, partial truth, and approximation. In particular, the primary contribution of fuzzy logic is the machinery of granular computing, which serves as a basis for the methodology of computing with words [43]. Suffice it to say that even granular computing and soft computing have significant overlap in contents and in research individuals, and are both driven by human-centered intelligent problem solving. They also differ in the perspectives used in problem solving: Unlike granular computing which places emphasis on granulation-based reasoning aspects of human information processing, soft computing sets emphasis on the “soft” side of human information processing which tolerates imprecision and uncertainty.

*Relationship with natural computing:* Since granular computing is closely related to soft computing and since soft computing is also closely related to natural computing, the relationship between granular computing and natural computing should not be overlooked. Natural computing is the term used to encompass all AI approaches based on some inspiration from nature, soft computing, man-made complex systems, biologically inspired computing, and novel computing paradigms rooted in nature in a higher or lesser degree. However, as noted by Zadeh [43], in effect, the role model for soft computing is the human mind. There is a significant overlap between natural computing and soft computing in content; yet they differ in *perspectives* and *emphases*: Natural computing is more concerned with how to *model nature* while soft computing focuses more on how to *achieve “softness”*. A brief examination of the contents of natural computing can re-

veal the relationship between granular computing and natural computing more directly: Recall that natural computing refers to three types of approaches [14]: *Computing inspired by nature* (such as generic algorithms) which makes use of nature as inspiration for the development of problem solving techniques; *simulation and emulation of natural phenomena in computers*, which is basically a synthetic process aimed at creating patterns, forms, behaviors, and organisms that (do not necessarily) resemble “life-as-we-know-it”; and *computing with natural material*, such as quantum computing. From a GrC perspective, a quantum can be viewed as a granule, and granule is organized in such a way that it could grow exponentially, so granulation may provide a new framework of attacking classical theory of computing [17]. In addition, genetic algorithms can be interpreted as producing new granules from existing granules through specific operations (such as crossover or mutation).

Taking a step further, we note that natural computing is an important field to demonstrate the pervasive concepts of emergence [20], embodiment [34], as well as other key concepts such as interactivity, adaptation, learning, evolution, and self organization. Granular computing can play an important role in the study of natural computing as in the case of learning theory. As noted in [17], learning is interpolation of data based on background knowledge (granules). Neural network, support vector machine, and Smale’s mathematical foundation of learning [13] can all be so formulated.

### Relationship with Fundamental Issues of Computing and Complex Systems Problem Solving

We now examine a number of more advanced perspectives relating granular computing to basics of computing and complex systems of problem solving.

### Granular Computing as Infrastructure for AI-Engineering

As endorsed by Zadeh [42] at the very beginning, granular computing should serve as foundation of human problem solving. As such, GrC serves as the infrastructure for AI-engineering: uncertainty management, data mining, knowledge engineering, and learning [30]. Structures, representations, and applications of granular computing have been discussed in this context [28]. Note that unlike the human problem solving perspective, this loose definition is computation-driven – even the key notions of granular computing were originally motivated by human-centered information processing.

It has been noted that a simpler and more effectively computable view of knowledge is needed. Rough set theory takes a courageous step and assumes that partitions (classifications) are the essence of human knowledge. Granular computing takes a softer view: generalized subsets are basic knowledge [17].

This perspective seems to serve as the de facto standard of “what granular computing is” for the majority of researchers who are active in this area. For example, [40] apparently echoes this perspective, where several learning paradigms for granular computing are discussed, although from a fuzzy set perspective.

Note also according to this definition, there is no urgency calling for a new computer model or AI architecture; we will stay with the von Neumann architecture, as traditional AI has done. Nevertheless, proponents of this perspective do relate granular computing to new computational models, such as quantum computing (as already mentioned earlier).

### Granular Computing and New Computational Models

References [1,2] proposed to interpret granulation in the context of axiomatic set theory. They define information granulation as a *semantically* meaningful grouping of elements based on their indistinguishability, similarity, proximity or functionality. The semantics of granules are derived from the domain that has, in general, higher cardinality than the cardinality of the granulated sets. The next question is then how the meaning (semantics) are instilled into real-life information granules. It has been argued that Turing’s computer on its own is unable to respond to external, physical stimuli. To overcome the problem, we may have to advocate the idea of embodiment in AI [34], and embrace Bain’s approach [4] of using inherent computational ability of physical phenomena in conjunction with the numerical information processing ability of a Universal Turing Machine (UTM). It has been argued that in a broad context a UTM implementing the virtual computational intelligence function can be referred to as computing with perceptions or computing with words.

### Summary

As a new research discipline, granular computing (GrC) is a subject of study for further abstraction, generalization and unification of granule-based key concepts. However, as the discussion of this articles shows, due to the complexity of this subject researchers differ in the scope, methodology, and perspective on the study of GrC. Although there has been a widespread enthusiasm for GrC, a unified theory of GrC is still missing. Nevertheless, there are a num-

ber of basic elements serving as the backbone of GrC such as granular structure, quotient structure and information hiding. We have also addressed certain issues not being emphasized by some researchers in GrC, such as granule measurement, as well as the relationship between hierarchy and emergence.

In the main text, while recognizing differences among different researchers, we have tried to present the philosophical foundation of granular computing as a coherent whole. As we are closing this article, we want to point out a few major differences in regard to philosophical foundations observed from various researchers.

As it currently stands, granular computing seems to be a loosely structured, somewhat interdisciplinary field: A big umbrella which accommodates researchers in rough set theory, fuzzy set theory, as well as soft computing and data mining – *so long as we all respect granulation as the core of human-centered information processing*. As for the theoretical core, much of today's study in granular computing is still largely an extension of rough set theory. Should this situation continue? Some may argue that this multidisciplinary nature is not all bad – after all, researchers with different backgrounds and different interests can now gather together, exchange their thoughts and foster new ideas – for example, the increasing publication of hybridization of fuzzy/rough set approaches seems to be a solid proof. Ironically, such successes may also have hampered the development of a unified theory of granular computing because there seems to be no urgency of bothering with it at all.

For those who have a long-term vision of granular computing, viewpoints vary significantly. Although seemingly everybody agrees granular computing is about human-centered information processing, what is its ultimate objective? As shown in the previous sections, answers could be:

- Assisting human problem solving in new ways of thinking,
- Providing infrastructure to AI-engineering, and
- Serving as the starting point toward new computational models.

Our examination has also addressed important issues related to embodiment and emergence. In addition, to cite Zadeh: “In coming years, granular computing is likely to play an increasingly important role in scientific theories—especially in human-centric theories in which human judgment, perception and *emotions* are of pivotal importance” (from [3], our italics). Therefore, as with the study of human-centered information processing, granu-

lar computing should also incorporate the role of emotions into its research agenda.

### Future Directions

Born in 1997, granular computing is now a happy teenager, enjoying a wide range of research fruits and publicity. But will its multiple personality survive adolescence to its adulthood? The answer is stated (or implied) in other articles on granular computing in this encyclopedia, and is thus not in the scope of this current article.

As for the research on philosophical foundations of granular computing, what we want to emphasize is that the future of granular computing lies in the development of algorithms. Yet a continued examination of philosophical foundations will benefit this direction of research. In addition, as noted in [41], in the past 10 years, granular computing publications have experienced a linear growth rate; however, the granular computing community has less interaction with other research communities other than fuzzy sets and rough sets. The granular computing community should be aware of this and take steps to promote interdisciplinary research activities.

The research agenda of GrC includes a series of methodological and algorithmic issues [1], including construction of information granules, characterization of dimension (granularity) of information granules (to provide better insight as to the essence of the granulation process and its implications), development of the encoding and decoding mechanisms between levels of hierarchy; research on interoperability which is crucial to the design of systems operating within the realm of various formalisms of information granularity, as well as others.

There is a need for focusing on key ideas behind granular computing that have been overlooked (or at least not emphasized enough), such as computing with words. In addition, as indicated earlier, important features needed to support complex systems such as self-organization, adaptation, evolution, etc. are currently not well-supported by granular computing research, and this situation should be changed.

### Acknowledgments

The author thanks Dr. T. Y. Lin's useful comments for the improvement of the paper.

### Bibliography

#### Primary Literature

1. Bargiela A, Pedrycz W (2002) Granular computing as an emerging paradigm of information processing. In: Bargiela A, Pedrycz W (eds) *Granular computing: An introduction*. Kluwer, Boston

2. Bargiela A, Pedrycz W (2006) The roots of granular computing. Proc IEEE Conf Granul Comput, pp 806–809
3. Bargiela A, Pedrycz W (2008) Toward a theory of granular computing for human-centered information processing. *IEEE Trans Fuzzy Syst* 16(2):320–330
4. Bains S (2004) Physical computation and the design of anticipatory systems. Proc IEEE SMC 2004
5. Boyd R, Gasper P, Trout JD (1991) The philosophy of science. Blackwell Publishers, Cambridge
6. Burkhardt H, Dufour CA (1991) Part/Whole I: History. In: Burkhardt H, Smith B (eds) Handbook of metaphysics and ontology. Philosophia, München
7. Chen Z (1999) An integrated architecture for OLAP and data mining. In: Bramer M (ed) Knowledge discovery and data mining: Theory and practice. IEE, London, Chap 6, pp 114–136
8. Chen Z (2002) The three dimensions of data mining foundation. In: Proc ICDM 2002 workshop on the foundation of data mining and discovery, Dec 2002, pp 119–124
9. Chen Z (2003) On Granular computing for aggregate data mining. In: 7th international conference on computer science and informatics, Proc JCIS, pp 431–434
10. Chen Z (2003) Discovering rules and answering queries at the right granulation levels. In: 7th international conference on computer science and informatics, Proc JCIS, pp 435–438
11. Chen Z (2003) Bioinformatics as a study of structured granules. In: International conference on computational intelligence and natural computing, Proc JCIS, pp 1629–1632
12. Chidlovskii B (2001) Schema extraction from XML data: A grammatical inference approach. In: KRDB'01 Workshop (knowledge representation and databases) Rome, Italy, 2001. <http://citeseer.ist.psu.edu/chidlovskii01schema.html>
13. Cucker F, Smale S (2002) On the mathematical foundations of learning. *Bull Amer Math Soc* 39:1–49
14. De Castro LN (2006) Fundamentals of natural computing: Basic concepts, algorithms, and applications. Chapman and Hall/CRC, Boca Raton
15. Dick S, Schenker A, Pedrycz W, Kandel A (2007) Regranulation: A granular algorithm enabling communication between granular worlds. *Inf Sci* 177(2):408–435
16. Gan Q, Wang G, Hu J (2006) A self-learning model based on granular computing. In: Proc 2006 Int Conf Granul Comput (IEEE GrC 2006), pp 530–533
17. Granular computing information center. [http://www.cs.sjsu.edu/~grc/grcinfo\\_center/grcinfo\\_index.php](http://www.cs.sjsu.edu/~grc/grcinfo_center/grcinfo_index.php)
18. Han J, Kamper M (2006) Data mining: Concepts and techniques, 2nd edn. Morgan Kaufmann, San Francisco
19. He J (2002) Schema discovery of the semi-structured and hierarchical data. In: Proc IDEAL (intelligent data engineering and automated learning) 2002. LNCS, vol 2412. Springer, Berlin, pp 129–134
20. Hobbs JR (1985) Granularity. In: Proceedings of the 9th international joint conference on artificial intelligence, pp 432–435
21. Holland JH (1999) Emergence: From chaos to order. Addison-Wesley, Reading
22. Jankowski A, Skowron A (2007) Toward rough-granular computing. In: Rough sets, fuzzy sets, data mining and granular computing (Proc RSFDGrC 2007). LNCS, vol 4482. Springer, Berlin, pp 1–12
23. Keet CM (2006) A taxonomy of types of granularity. In: Proc 2006 Int Conf Granul Comput (IEEE GrC 2006), pp 106–111
24. Lee TT (1987) An information-theoretic analysis of relational databases – Part I: Data dependencies and information metric. *IEEE Tran Softw Eng* 13(10):1049–1061
25. Libkin L (2007) Normalization theory for XML. In: Proc XSym 2007, pp 1–13
26. Lin TY (1998) Granular computing on binary relations I: Data mining and neighborhood systems. In: Skowron A, Polkowski L (eds) Rough sets in knowledge discovery. Physica, Heidelberg, pp 107–121
27. Lin TY (1998) Granular computing on binary relations II: Rough set representations and belief functions. In: Skowron A, Polkowski L (eds) Rough sets in knowledge discovery. Physica, Heidelberg, pp 121–140
28. Lin TY (2003) Granular computing: Structures, representations, and applications, rough sets, fuzzy sets, data mining, and granular computing. In: Proc RSFDGrC 2003. LNCS, vol 2639. Springer, Berlin, pp 16–24
29. Lin TY (2004) Mining associations by linear inequalities. ICDM 2004:154–161
30. Lin TY (2006) A roadmap from rough set theory to granular computing. LNAI, vol 4062, pp 33–41. Updated version: Granular computing on coverings: A roadmap from rough set theory to granular computing
31. Lin TY (2006) Granular computing II: Infrastructure for AI-engineering. In: Proc 2006 Int Conf Granul Comput (IEEE GrC 2006), pp 2–7
32. Lin TY (2003) Deductive data mining: Mathematical foundation of database mining. LNCS, vol 2369, pp 403–404
33. Lin TY, Zadeh LA (2004) Special issue on granular computing and data mining. *Int J Intell Syst* 19(7):565–566
34. Pfeifer R, Bongard J (2007) How the body shapes the way we think: A new view of intelligence. MIT Press, Cambridge
35. Qiu T, Chen X, Huang H, Liu Q (2006) Ontology capture based on granular computing. In: Proc 6th Int Conf Intell Syst Design Appl (ISDA'06), pp 770–774
36. Silberschatz A, Korth H, Sudhshan S (2006) Database system concepts, 5th edn. McGraw-Hill, Boston
37. Singh GB (2002) Discovering matrix attachment regions (MARs) in genomic databases. SIGKDD Explor 1(2):39–45
38. Skowron A, Stepaniuk J (2001) Information granules: Towards foundations of granular computing. *Int J Intell Syst* 16:57–85
39. Sowa JF (1999) Knowledge representation: Logical, philosophical, and computational foundations. Brooks/Cole Publishing, Pacific Grove
40. Yager RR (2006) Some learning paradigms for granular computing. In: Proc 2006 Int Conf Granul Comput (IEEE GrC 2006), pp 25–29
41. Yao JT (2007) A ten-year review of granular computing. GrC 2007:734–739
42. Zadeh LA (2006) Granular computing – The concept of generalized constraint-based computation. RSCTC 2006 (LNCS, vol 4259), 12–14
43. Zadeh LA (1998) Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems. *Soft Comput* 2:23–25
44. Zadeh LA (2007) Granular computing – Computing with uncertain, imprecise and partially true data. In: Proc 5th Int Sym Spat Data Qual (ISSDQ 2007). [http://www.itc.nl/ISSDQ2007/Documents/keynote\\_Zadeh.pdf](http://www.itc.nl/ISSDQ2007/Documents/keynote_Zadeh.pdf)
45. Zadeh LA (2007) Granular computing and rough set theory. In:

- Kryszkiewicz M, Peters JF, Rybinski H, Skowron A (eds) Rough sets and intelligent systems paradigms. Proc RSEISP 07, pp 1–4
46. Zhang L, Zhang B (2003) The quotient space theory of problem solving. LNCS, vol 2639. Springer, Berlin, pp 11–15
  47. Zhang X, Chen Z (2002) Algorithms for finding influential association rules. In: Proc 1st IEEE international conference on fuzzy systems and knowledge discovery, pp 499–503
  48. Zhou G, Liang J (2006) Granular computing model based on ontology. IEEE Int Conf Granul Comput 321–324

## Books and Reviews

- Johnson S (2001) Emergence: The connected lives of ants, brains, cities, and software. Scribner, New York
- Lin TY (1997) Granular computing. In: Announcement of the BISC special interest group on granular computing.
- Minsky M (2006) The emotion machine: Commonsense thinking, artificial intelligence, and the future of human mind. Simon and Schuster. Online draft at Minsky's website <http://web.media.mit.edu/~minsky/>
- Pawlak Z (1991) Rough sets: Theoretical aspects of reasoning about data. Kluwer Academic Publishers, Dordrecht
- Pawlak Z (1998) Granularity of knowledge, indiscernibility and rough sets. In: Proceedings of 1998 IEEE International Conference on Fuzzy Systems, pp 106–110
- Wang Z, Leung K-S, Klir GJ (2008) Integration on finite sets. Int J Intell Syst 21:1073–1092
- Zadeh LA (1997) Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Set Syst 19:111–127

# Granular Computing: Practices, Theories, and Future Directions

TSAU YOUNG LIN

Department of Computer Science,  
San Jose State University, San Jose, USA

## Article Outline

- Glossary
- Definition of the Subject
- Introduction
- Classical Examples of Granulation
- Formal Models of Granulation
- Granular and Related Structures
- Integration – A Dual of Granulation
- Semantic Views
- Future Directions
- Bibliography

## Glossary

All terms are explained in classical sets, but implicitly, we are assuming all terms and assertions do include fuzzified versions (if fuzzifiable).

**Granulation** Granulation is an operation or a process of forming granules, with a granule being a collection of objects (points) that are drawn together by some constraints, such as indistinguishability, similarity or functionality.

**Granular structure** Granular structure is the collection of granules, in which the internal structure of each granule is visible as a sub-structure. Informally speaking, granular structure is a collection of white box granules.

**Quotient structure** A quotient structure is the mathematical structure of the collection of granules, in which each granule is regarded as an element (point) of a set, but the interactions among granules are preserved. Informally speaking, a quotient structure is a collection of black box granules.

The collection of  $\{\dots, -2, 0, 2, \dots\}$  and  $\{\dots, -3, 1, 3, \dots\}$  is a granular structure. Let  $E$  be the first subset (even integers) and  $O$  be the second subset (the odd integers). We write the two subsets by  $[E]$  and  $[O]$ , when we think of them as points. Then the collection of  $[E]$  and  $[O]$  (as points) is the quotient structure.

**Neighborhood system (local granular model abb. local GrC model)** A domain of interests (a classical set)  $U$  is called the universe. To each point  $p$  in the universe, a family of subsets is assigned. Such a family (could be empty) and each such subset is called a neighborhood system  $NS(p)$  at  $p$  and a neighborhood at  $p$ , respectively. The collection  $\beta$  of such a family at every point of the universe is called a neighborhood system  $NS(U)$  of the universe. Neighborhood and neighborhood system are pre-GrC language; in granular computing, they are called granule and the granular structure, respectively. The pair  $(U, \beta)$  is called a local granular model, since each granule is associated with some points.

**Topological neighborhood system** A neighborhood system is called a topological neighborhood system, if it satisfies the axioms of topology.

**Binary neighborhood system (binary granular model; binary GrC model)** A binary neighborhood system is a neighborhood system defined by a binary relation  $R$ . A (right) neighborhood is defined as follows:  $B(p) = \{x \mid (p, x) \in R\}$ . The collection  $B$  of  $B(p)$  at each  $p$  is the (right) binary neighborhood system. Similarly, we can define a left version: A left neighborhood system  $L$  is defined by the  $L(p) = \{x \mid (x, p) \in R\}$  at every point  $p$ . Note that the right and left neighborhood systems determine each other. The pair  $(U, \beta)$  is called a binary granular model, where  $\beta$  is right (left) neighborhood system or  $R$ .

**pre-Topology** Pre-topology is a general term referring to

the neighborhood system, which includes the topological neighborhood system and binary neighborhood system as special cases. Technically, it is equivalent to the neighborhood system.

**Bag** A bag is similar to a set, but allows an element to appear more than once. For example  $\{1, 2, 1, 2, 1\}$  is a bag, but not a set. If a bag contains  $n$  elements, we may say it is an  $n$ -bag. For example, the previous bag is a 5-bag.

**Relational structure (relational granular model; relational GrC model)** A family of classical sets is called a universe and denoted by  $U$ . A Cartesian product of an  $n$ -bag of  $U$  is called an  $n$ -product set. A  $n$ -ary relation is a subset of an  $n$ -product set. A collection  $\beta$  of  $n$ -ary relations ( $n$  could vary) is called a relational structure. The pair  $(U, \beta)$  is called a Relational GrC Model. Note that this relational structure, except the size, is similar to the relational structure (without functions) of the First Order Logic; the Relational GrC Model permits  $n$  to be any cardinal number.

**Partial covering (global granular model; global GrC model)** Let  $U$  be a classical set, called the universe. Let  $\beta = \{F^1, F^2, \dots\}$  be a family of subsets. Such a  $\beta$  is a partial covering, and (full)covering, if the union of  $\beta$  is the whole universe. The pair  $(U, \beta)$  is called a Global Granular Model (Global GrC Model).

**Equivalence Relation** A binary relation  $R$  is called an equivalence relation, if it has the following properties: Let  $u, v$ , and  $w$  be elements of  $U$ .

**reflexive:**  $uRu$

**symmetric:**  $uRv$  implies  $vRu$

**transitive:**  $uRv$  and  $uRw$  implies  $vRw$

**Partition** A partition  $P$  of a classical set  $U$  is a collection of subsets that are mutually disjoint and their union is  $U$ . Each subset is called an equivalence class. This name is derived from the fact that partition is equivalent to the following equivalence relation: We define  $uRv$ , if and only if  $u$  and  $v$  belong to the same equivalence class. Such  $R$  is the equivalence relation corresponding to the partition  $P$ . Note that a partition is a special type of a granular structure, so an equivalence class is a special granule.

### Definition of the Subject

Granular Computing (GrC) is still in its inception stage, we use motivation as its statements of importance.

### How Important is Granular Computing?

Granulation seems to be a natural methodology deeply rooted in human thinking. Many daily “things” are rou-

tinely granulated into sub“things”; for example, the human body has been granulated into the head, neck, and so forth. The notion is intrinsically fuzzy, vague, and imprecise. Formalization is difficult, mathematicians idealized/simplified it into the notion of partitions (= equivalence relations), and have developed it into a fundamental part of mathematics, for example, congruence in Euclidean geometry, quotient structures (groups, rings, etc) in algebra, the concept of “a. e.” (almost every where) in analysis. Nevertheless, the notion of partitions (see glossary), which absolutely does not permit any overlapping among its granules, seems to be too restrictive for real world problems. Even in natural science, classification *does* permit a small degree of overlapping; there are beings that are both appropriate subjects of zoology and botany. So a more general theory, namely, Granular Computing is needed.

### What is Granular Computing?

It has been a changing paradigm. We believe in incremental developments. Its development may be similar, though not as glorious, to that of classical geometry. Specific geometries, such as, Euclidean, hyperbolic, elliptic geometries, had appeared before the the unified theory was proposed in Klein’s Erlangen program. Except some details, we are on the final line; in GrC2008 keynote, we (this section editor) verbally had proposed to regard the category theory-based model (Eighth GrC Model) as the final GrC model; see Sect. “[Formal Models of Granulation](#).”

Granular Computing is a recent label coined by Lin and Zadeh (see Sect. “[Introduction](#)”) to denote a set of common, even ancient, concepts and practices. The subject will be presented from various angles: What are the target concept (defined by examples), key constituents, and the current interpretations or semantic views? How far has it been formalized? What are the important applications?

1) The “Definition” of Target Concept: The key concept in Granular Computing is the concept of granulation, which has been “defined” implicitly by a set of intuitive examples (see Sect. “[Classical Examples of Granulation](#)” for more), including Zadeh’s intuitive view. Here is a list of representative examples. In the following, we will call the domain of interests the universe and denote it by  $U$ .

1. Many daily things have been routinely granulated into “sub”things; for example, the human body is granulated into head, neck, etc. The notion is intrinsically fuzzy, vague and imprecise. A formal model just had been proposed by this writer in the keynote

- of GrC2008. Its final form should appear soon in the International Journal of Granular Computing, Rough Sets, and Intelligent Systems.
2. Space and time has been granulated intuitively into infinitesimal granules; a circle was viewed as a polygon with infinitesimal sides. The idea was known to Zeno (490 BC, implicitly in his paradox), Archimedes (287–212 BC) etc. It led to the invention of calculus, topology, and nonstandard analysis; we have modeled it in the First GrC Model.
  3. The simplest kind of granulation is partition (see glossary). Its algebraic correspondence, equivalence relation, has played an important role in Euclidean Geometry. (300 BC)
  4. The Heisenberg uncertainty principle states that, in general, neither the momentum nor the position of a particle can be determined simultaneously with arbitrary great precision. In other words, a precise measurement of the momentum can only determine a “neighborhood”(granule) of positions and vice versa. The idea is abstract into the Third GrC Model.
  5. A collection of fuzzy sets (in fuzzy control) or functions (e.g. Radial-Basis-Functions) which has the universal approximation property is useful granular structure in a function space or a set of fuzzy sets. The idea is modeled in the Sixth GrC Model.
  6. A committee in a human society is a granule in a social network. Observe that each member may play different roles. By viewing the collection of roles as a relational schema, a committee is a tuple, not necessary a subset. The idea is modeled in the Fifth GrC Model and the Second GrC Model.
  7. A mathematical proof or computer program often contains some lemmas or subprograms. These lemmas or subprograms are granules. These are conceptual examples. They are modeled in the Seventh GrC Model for computable domain, and in the Ninth GrC Model for general cases.
  8. Computers or clusters of computers in Grid/Cloud computing are granules. These are hardware examples. This also belongs to the Seventh GrC Model.
  9. Zadeh's informal definition [38]: “information granulation involves partitioning a class of objects (points) into granules, with a granule being a clump of objects (points) which are drawn together by indistinguishability, similarity or functionality.”
  - 2) A Category Theory based Formal model (Eighth GrC Models) is proposed to be the Formal Model for GrC. It realizes all classical examples given above; see Sect. “[Formal Models of Granulation](#).” By specifying the abstract category to various ones, eight common models have been explained in this article. Two models are illustrated at the end of this section;
  - 3) The Key Constituents: two operators, three semantic views, and four structures.
    - 3.1) Two Operators: information granulation and integration. In granulating a problem, a dual action, namely, integrating the solutions of sub-problems is triggered. Here, we highlight some unusual points; see Sect. “[Integration – A Dual of Granulation](#).”
      1. Recursively granulating a problem may take  $N_p$  hard time to terminate. Dynamic programming technique has been used in the classical cases.
      2. Two distinct problems may have the same granulation. Divide/granulate and conquer may have deeper meanings. The EXTENSION Functor of Homological Algebra of Commutative group is illustrated. This is an unexplored area.
    - 3.2) Three Semantic Views: Granules may be interpreted from:
      1. Uncertainty Theory: A granule is a unit of lacking precise knowledge.
      2. Knowledge Engineering: A granule is a unit of Basic Knowledge (Information).
      3. How-to-Solve/Compute-it: A granule is a subproblem or software unit. It is a special type of basic knowledge.
 Each view may have its own GrC theory; for example, Concept Approximations are useful in the Second View, while Information Hiding is in Third View; see Section [Semantic Views](#).
    - 3.3) Four Structures: Granular, Quotient, Knowledge and Linguistic Structures:
      1. Granular Structure (GrS): It is the collection of all granules. In the case of partition, GrS is the collection of equivalence classes.
      2. Quotient Structure (QS): If each granule is abstracted into a point and the intersections of granules are abstract to the interactions of points, then such a collection of points is called a quotient structure. In the case of partition, the quotient structure is a classical set, called quotient set. Informally, GrS is a collection of white boxes (the content of granules are visible), while the quotient set is a collection of black boxes (the contents of granules are hided). The process of abstracting a granular structure into a quotient structure is called information hiding; see Examples in Sect. “[Granular and Related Structures](#).”
      3. Knowledge structure: By giving each granule (point) in the quotient structure a meaning-

ful symbol then the named quotient structure is called a knowledge structure. The knowledge structure provides an intuitive view of the quotient structure; the symbols and interaction among symbols are in sync with the granules (points) and interactions among granules (points).

In the case of  $n$  partitions (equivalence relations), the knowledge structure can be arranged into an  $n$ -column relational table. In the case of  $n$  binary relations, the table has been called binary information table, granular table or topological table [15,22].

4. Linguistic structure: By giving each granule in the granular structure a word that reflects its meaning. The interactions among these words are reflected implicitly in precisiated natural language (in a knowledge structure, the interactions among symbols are explicitly reflected from the quotient structure). The linguistic structure is the domain of computing with words
- 4) Applications to Computer Security, Web Technology, and Complex Data.
  1. Discretionary Access Control Model (DAC) [8,17]. This structure has been captured in the Third GrC Model. On the basis of this, information flows on DAC can be analyzed; this has been considered a very “difficult” area. As a consequence the Aggressive Chinese Wall Security Policy can be enforced, namely, the system can guarantee that a company’s data will never flow into “enemy” hands, where “enemy” is a granule of companies that are in conflict.
  2. Documents can be clustered into a simplicial complex (a common structure in Combinatorial Topology [34]) of keywords and co-occurring keyword sets. The set of keywords can be regarded as a set of vertices, and the collection of co-occurring keywords (within a small neighborhood) is a set of simplexes. Together, they form a simplicial complex [21,24]. Simplexes are granules; a simplicial complex is a Second GrC model.
  3. Granular computing has been used to solve the modeling problem of complex architectures [25]. It uses the Fifth GrC model.

### Illustration of Working Formal Models

The most general model is expressed in the category theory. However, for easiness, we explain the Second GrC model first.

### Granular Computing: Practices, Theories, and Future Directions, Table 1

#### Generalize 2nd to 5th GrC Models

2nd GrC	Generalized to	5th GrC
$U$	$\rightarrow$	$U = \{U_j^h, h, j, = 1, 2, \dots\}$
$F^1 \subseteq U$	$\rightarrow$	$R^1 \subseteq U_1^1 \times U_2^1 \times \dots$
...	...	...
$F^l \subseteq U$	$\rightarrow$	$R^l \subseteq U_1^l \times U_2^l \times \dots$
...	...	...

1) **Second GrC Model** Let  $U$  be a classical set, called the universe. Let  $\beta = \{F^1, F^2, \dots\}$  be a family of subsets. Then the pair  $(U, \beta)$ , called the Global GrC Model or the Second GrC Model. The  $\beta$ , some time, is called Partial Covering (PCov).

A *granule* can be intuitively defined as a clump of objects that are drawn together by the constraints  $X$ , where  $X$  can be indistinguishability, similarity or functionality and etc [14,38]. In the Second GrC model, a granule is a set, namely, we have implicitly assumed that the constraints are uniform. In general, each object may receive distinct constraints, so in the Fifth GrC model, a granule is a tuple, not necessarily a set: One can regard the constraints as a schema (of a relational database), and a granule is a tuple under such a schema. In this case the collection of granules are tuples from various relations.

To understand the next model, Table 1, that explains the generalization process, may be helpful.

#### 2) Fifth GrC Model

1. Let  $U = \{U_j^h, h, j, = 1, 2, \dots\}$  be a given family of classical sets, called the universe. Note that distinct indices do not imply that the sets are distinct.
2. Let  $U_1^j \times U_2^j \times \dots ; j = 1, 2, \dots$  be a family of Cartesian products of various lengths.
3. Recall that an  $n$ -ary relation is a subset  $R^j \subseteq U_1^j \times U_2^j \times \dots U_n^j$ .
4. Let  $\beta = \{R^1, R^2, \dots\}$  be a given family of  $n$ -ary relations for various  $n$ .

Then the pair  $(U, \beta)$ , called the Relational GrC Model, is the formal definition of the Fifth GrC Model.

### Introduction

What is Granular Computing (GrC)? The best approach is to trace how the intuitions have been evolved. For this purpose, let us recall the event. In the academic year 1996–97, when Lin had his sabbatical leave at Berkeley, Zadeh suggested granular mathematics (GrM) to be his research area. To limit the scope, Lin proposed the term granular

computing [40]. So, at the beginning, roughly GrC is the computable part of granular mathematics.

*What is Granular Mathematics (GrM)?* Zadeh in his 1979 paper [37] had implicitly explained his view. Here, we take a simpler view: It is a new mathematics, in which “points” are replaced by or associated with “granules.” We call this process granulation, and the collection of granules the granular structure.

*What is a Granule?* This is the main topic. There is an obvious candidate, namely, an equivalence class of a *partition* that is an ancient notion in mathematics. In general, a granule can be a crisp/fuzzy subset, a function, an algorithm, a random variable (measurable function), a generalized functions etc.

Traditionally, “how to solve it” [31] cannot be any part of formal mathematics, however, “how to compute it” is an integral part of computing. So GrC has included the mathematical/computational problem solving practices.

## Classical Examples of Granulation

### Ancient Examples

**E1 Granulation of the Human Body: The Granules Are Head, Neck, Body, Hand etc.** Many daily things are routinely granulated into “sub”things, probably since ancient time. Human body is granulated into head, neck, etc. Currently, there are no flawless formal models to capture such intuitive concept yet. The obvious one does not work well: One can easily write down a fuzzy membership function to represent a head, a neck or a body. However, the likelihood of any two persons to write down the same membership function for the same granule (e.g., head) is extremely low. Hence we need a much more subtle theory. A new proposal on qualitative fuzzy set theory is in preparation by this author, this example can be realized.

**E2 Granulation of Space and Time** The space and time has been granulated intuitively into infinitesimal granules by early scientists; this notion was known to Zeno, Archimedes, etc. This intuitive notion led to the invention of calculus by Newton and Leibniz. However, its formalizations, theory of limit (18th century), topology (early 20th century [28]) and nonstandard analysis (mid 20th century [32]) are relatively recent. This ancient example, inspired two models, Local GrC model (First GrC Model) and Global GrC Model (Second GrC Model).

**E3 Partition** The simplest kind of granulation is partition (see glossary). Its algebraic correspondence, equivalence relation, has played an important role in the Euclidean Geometry (300 BC).

### Modern Examples

**E4 Local Granules of Uncertainty – from Quantum Mechanics** Heisenberg’s uncertainty principle states that, in general, neither the momentum nor the position of a particle can be determined simultaneously with arbitrary great precision. In other words, a point of the momentum (a precise measurement) can determine only a “neighborhood” of positions and vice versa. The idea is simplified into the Third GrC Model (Binary GrC Model).

**E5 Local Granules of Basic Knowledge – Discretionary Access Control Models in Computer Security** This example is a common data structure in many computer systems. For each user, there is a set of users (friends) who can access his files, or a set of users (foe), who cannot access his files; this set is called explicitly denied access list.

Examples E4 and E5 are “serious” examples. The idea is modeled in the third GrC Model. Mathematically, it is equivalent to a binary relation. Geometrically, a binary relation is a graph, or network.

**E6 Global Granules of Knowledge – Simplicial Complexes** A simplicial complex consists of two objects: One is a finite set of vertices, another one is a family of subsets, called simplexes, of vertices that satisfy the closed condition, namely a subset of a simplex is a simplex. It is a common mathematical structure in combinatorial topology. Currently, it is finding its way to web technology [21].

### Further Examples

Next, we give examples of non-commutative granules, which is a generalization of a binary relation (Binary GrC Model; Third GrC Model).

**E7 A committee in a Human Society Is a Granule in a Social Network** A committee in a human society (a set of human beings) is a granule. Observe that each member may play different roles, so the committee may not consist of homogenous members; so the members cannot exchange their roles. By viewing the collection of roles as a relational schema, a committee is a tuple. This idea is modeled in the Fifth GrC Model. Observe that different types of committees have different schema. The set of committees under the same schema forms a relation. A collection of  $n$ -nary relations for various  $n$  in a society can be viewed as a granulation of the society.

In these examples, granules can be fuzzy sets. Since a fuzzy set is characterized by a membership function, so we have generalized the idea further.

### E8 A Granule Can Be a Function, a Random Variable (A Measurable Function) or Even a Generalized Function (Such as the Dirac Delta Function)

This class of examples led to the Sixth GrC model.

Traditionally, “how to solve it” [31] has not been any part of formal mathematics, however, “how to compute it” is an integral part of computing. So GrC includes some mathematical/computational practices.

**E9 A granule can be a lemma in a mathematical proof, a subprogram in a program, or a machine/cluster of machines in a grid/clouding computing environment.** Formally, within the computable domains, a granule is a sub-Turing machine. This is modeled in the Seventh GrC Model. For general cases, they are included in the Ninth GrC Model

**E10 Computers or clusters of computers in Grid/Cloud computing are granules.** These are hardware examples. This also belongs to the Seventh GrC Model

### Formal Models of Granulation

On the basis of these examples, nine models are discussed. The category theory-based model (Eighth GrC Model) has been proposed by this writer in the keynote of GrC2008 as *the formal model of GrC*. The rest of the eight models are basically “convenient models” in the sense that they can be derived from the most general model, but for convenience, they are modeled independently.

### First GrC Models and Ancient Examples

This model is derived from the Ancient Example E2, which probably is the most lively example in the early notion of granules. It led to the invention of calculus by Newton and Leibniz. Actually the idea was much more ancient; it was in the mind of Archimedes, Zeno, etc. Yet the solutions were in modern time. Two formalizations had emerged. One was the concept of limit (19th century) that led to the notion of topology (early 20th century). The other one was the nonstandard analysis (mid 20th century), which formally realized the original intuition.

These developments conclude that the following concepts are “equivalent” (not in the formal mathematical sense, as the first one is merely an intuitive notion):

1. The ancient intuitive notion of infinitesimal granule.
2. The formal infinitesimal granule in the non standard analysis.
3. Topological Neighborhood System (TNS) in the standard world.

In other words, the ancient intuition of infinitesimal granules (with the required properties) is realized, not by a set, but by a family of subsets, that satisfies the axioms of topology. Nevertheless, in this paper a (modern) granule will refer to a neighborhood, but, not to the whole neighborhood system.

The notion of topology can be defined in two ways:

1. A topology  $\tau$  is a family of subsets, called open sets, that satisfies the (global version) axioms of topology.
2. A topology, called topological neighborhood system (TNS), is an assignment that associates to each point  $p$  a family of subsets,  $TNS(p)$ , that satisfies the (local version) axioms of topology.

These two definitions lead us to First and Second GrC Models (Local and Global GrC Models). In this subsection, we will focus on the First GrC Model: Let  $U$  and  $V$  be two classical sets. Let NS be a mapping, called neighborhood system (NS),

$$\text{NS}: V \rightarrow 2^{P(U)},$$

where  $P(X)$  is the family of all crisp/fuzzy subsets of  $X$ .  $2^Y$  is the family of all crisp subsets of  $Y$ , where  $Y = P(U)$ . In other words, NS associates each point  $p$  in  $V$ , a family  $NS(p)$  of crisp/fuzzy subsets of  $U$ . Such a subset is called a neighborhood (granule) at  $p$ , and  $NS(p)$  is called a neighborhood system at  $p$ .

**Definition 1 (First GrC Model)** The 3-tuple  $(V, U, \beta)$  is called the **Local GrC Model**, where  $\beta$  is a neighborhood system (NS). If  $V = U$ , the 3-tuple is reduced to a pair  $(U, \beta)$ . In addition, if we require NS to satisfy the topological axioms, then it becomes a TNS.

**Some Intuition Behind the NS** The following arguments are adopted from my pre-GrC paper [13]

#### (1) The Meaning of “Near”

The notion of near is rather difficult to formalize. Let us examine the following two examples.

1. Is Santa Monica “near” Los Angeles? Answers could vary. For local residents, who have cars, answers are often “yes.” For visitors, who have no cars, answers may be “no.”
2. Is  $1.73$  “near”  $\sqrt{3}$ ? Again answers vary; it depends on what should be the appropriate tolerance radius.

Intrinsically “near” is a subjective judgment. One might wonder whether there is a scientific theory for such subjective judgments? Mathematicians have offered a nice solution. They simply include all contexts into its formalism.

Here is the formalism of the second question: Given the radius of an acceptable error, say, radius of errors 1/100 (a given context)

Is 1.73 “near”  $\sqrt{3}$ ?

With the agreement 1/100 is acceptable, then 1.73 is near  $\sqrt{3}$ ! In this case, all possible contexts are  $\varepsilon$  that represents all positive real numbers. Similarly, if a neighborhood system has been assigned to each city in the Los-Angeles-area. For example, based on car driving, public transportation, walking etc., we assign a neighborhood to *each* city for *each* context. Under such a concept of a neighborhood system, we could have a definite answer for Example 1. So a proper formulation for such a question is:

Assuming that we are taking public transportation (a given context), Is  $p$  near  $q$ ?

Therefore, a neighborhood system is a good infrastructure for addressing the concept of “near”! These analysis leads to the following conclusions.

1. In Modeling, a neighborhood system is a good infrastructure for providing all possible contexts.
2. Under this model, in an application, selecting a context means selecting *a fixed* neighborhood as a unit of tolerance (uncertainty).

Now, under this concept, we will re-examine previous examples.

*Example 1* If we have chosen “driving half an hour” as acceptable distance, then Santa Monica is “near” Los Angeles.

*Example 2* Let the collection of  $\varepsilon$ -neighborhoods be the neighborhood system for the real numbers  $R$ ; then  $(R, \varepsilon\text{-neighborhoods})$  is a First GrC Model, where  $\varepsilon$  could take any real value. Now, we re-state the previous example using this First GrC Model

1. Assuming we have agreed  $\varepsilon = 1/100$  is acceptable, then 1.73 is “near”  $\sqrt{3}$ .
2. But, if we have only agreed  $\varepsilon = 1/1000$  then 1.73 is not “near”  $\sqrt{3}$ .
3. Next let us consider a deeper question:

Is the sequence  $1, 1/2, 1/3, \dots, 1/n, \dots$  “near” zero?

Then, it is possible, we can have a “yes” answer for all contexts: For any given context, namely,  $\varepsilon > 0$ , there is a number  $N = [1/\varepsilon] + 1$ , such that, for all  $n > N$ ,  $1/n$  is “near” zero, where  $[1/\varepsilon]$  denotes the biggest integer  $\leq 1/\varepsilon$ .

Readers who are familiar with the standard  $(\varepsilon, \delta)$ -definition of limit can spot the origin of neighborhood systems. Such a context-free (all possible contexts) answer is precisely the classical notion of limits,  $\lim_{n \rightarrow \infty} 1/n = 0$ . Using our language, we may say that limit is the context free answers of “near”.

Perhaps we should also point out here that there are no context free answers for the question whether two points are “near.”

### Brief pre-GrC Historical Notes

1. In 1988–1989, Lin generalized TNS to the Neighborhood Systems (NS) by simply dropping the (local version) axioms of topology [7,9] and apply it to approximate retrievals. Each neighborhood was treated as a unit of uncertainty.
2. In the same year (1989), Lin also examined a non-reflexive and symmetric binary relation (conflict of interests) for computer security from the view of NS [8].
3. Abstractly, Lin imposed the NS structure on the attribute domains for approximate retrieval. Taking this view, we should mention that earlier D. Hsiao imposed equivalence relations on the domain for access precision in early 1970 [6,36]. In 1980, S. Ginsburg and R. Hull had imposed partial ordering on attribute domains [4,5].
4. In much earlier, NS was studied in [33] as a generalization of topology. Note that however, there are fundamental differences, for example, the concept of closures are different. The term pre-topology also has been used for referring NS and TNS.
5. In the early GrC period, Lin, by mapping the NS onto Zadeh’s intuitive definition, used NS as his first mathematical GrC model [14,15,16].

### Second GrC Models and Modern Examples

As in the previous case, by dropping the global axioms of topology, we have the Second GrC model.

**Definition 2 (Second GrC Model)** The Pair  $(U, \beta)$ , where  $\beta$  is a family of subsets of  $U$ , is called **Global GrC Model**. The  $\beta$ , some time, is referred to as a partial covering (PCov).

Note that the Second GrC model is a special case of the First GrC model: If we regard the sub-collection of all members of the partial covering  $\beta$ , that contains  $p$ , as a neighborhood system at  $p$ , then this Second GrC model is an example of the First GrC model.

The modern example, simplicial complexes, is an important example of such a model: A simplicial complex

consists of a set of vertices and a family of subsets, called simplexes, that satisfies the closed condition [34].

[Digression] Perhaps, it is worthwhile to note that

- The closed condition of the simplicial complex is the a priori principle in association (rules) mining.

This observation plays an important role in document clustering [24].

### Third and Fourth GrC Models and Modern Examples

In this section, we will build a new model that realizes modern example E4 and E5. Recall that E4 concludes that a precise measure of the momentum can only determine a (probabilistic) “neighborhood” of positions; and E5 concludes that in computer security, the Discretionary Access Control Model (DAC) assigns to each user  $p$  a family of users,  $Y_i$ ,  $i = 1, \dots$ , who can access  $p$ 's data. In other words, each  $p$  is assigned a granule of friends.

To formalize these examples, let  $U$  and  $V$  be two classical sets. Each  $p \in V$  is assigned a subset,  $B(p)$ , of “basic knowledge” (a set of friends or a “neighborhood” of positions).

$$p \rightarrow B(p) = \{Y_i, i = 1, \dots\} \subseteq U.$$

Such a set  $B(p)$  is called a (right) binary neighborhood and the collection  $\{B(p) \mid \forall p \in V\}$  is called the binary neighborhood system (BNS).

**Definition 3 (Third GrC Model)** The 3-tuple  $(U, V, \beta)$ , where  $\beta$  is a BNS, is called a **Binary GrC Model**. If  $U = V$ , then the 3-tuple is reduced to a pair  $(U, \beta)$ .

Observe that BNS is equivalent to a binary relation (BR):

$$BR = \{(p, Y) \mid Y \in B(p) \text{ and } p \in V\}.$$

Conversely, a binary relation defines a (right) BNS as follows:

$$p \rightarrow B(p) = \{Y \mid (p, Y) \in BR\}.$$

So, both modern examples give rise to BNS which was called a binary granular structure in [14]. We would like to note that based on this (right) BNS, the (left) BNS can also be defined:

$$D(p) = \{Y \mid p \in B(Y)\} \text{ for all } p \in V.$$

Note that BNS is a special case of NS, namely, it is the case when the collection  $NS(p)$  is a singleton  $B(p)$ . So the Third GrC Model is a special case of the First GrC Model.

The algebraic notion, binary relations, in computer science, is often represented geometrically as graphs, networks, forests etc. So the Third GrC Model has captured most of the mathematical structure in computer science.

Next, instead of a single binary relation, we consider the case where  $\beta$  is a set of binary relations. It was called a [binary] knowledge base [14]. Such a collection naturally defines an NS.

**Definition 4 (Fourth GrC Model)** The Pair  $(U, \beta)$ , where  $\beta$  is a set of binary relations, is called Multi-Binary GrC Model. This model is most useful in data bases; hence it has been called Binary Granular Data Model (BGDM), in the case of equivalence relations, it is called Granular Data Model (GDM)

Observe that a Fourth GrC Model can be converted, say by a mapping  $G$ , to a First Model. Conversely, a First GrC Model induces, say by  $F$ , to a Fourth Model. So First and Fourth models are equivalent, but not naturally, namely,  $G$  and  $F$  are not inverse to each other.

### Models for Further Examples

We have observed in Sect “[Definition of the Subject](#)” that the collection of  $n$  objects that are “drawn together” is, not necessary a subset, but is a tuple in an  $n$ -ary relation. For example, if the universe is a human society then a group of people may be drawn into a committee with distinct roles, such as the chair, vice chair, secretary, treasurer, etc. As every member has a different role, they can not be swapped around. So the committee is not a set; it is a tuple under the schema that consists of distinct roles.

### Definition 5 (Fifth GrC Model)

1. Let  $U = \{U_j^h, h, j = 1, 2, \dots\}$  be a given family of classical sets, called the universe. Note that distinct indices do not imply that the sets are distinct.
2. Let  $U_1^j \times U_2^j \times \dots$  be a family of Cartesian products of various length.
3. Recall that an  $n$ -ary relation is a subset  $R^j \subseteq U_1^j \times U_2^j \times \dots \times U_n^j$ .
4. Let  $\beta = \{R^1, R^2, \dots\}$  be a given family of  $n$ -ary relations for various  $n$ .

The pair  $(U, \beta)$ , called the Relational GrC Model, is a formal definition of the Fifth GrC Model.

Note that this granular structure is the relational structure (without functions) in the First Order Logic, if  $n$  only varies through finite cardinal number.

For the next two models, we will use the language of category theory in next sub-section. We may note that we have not committed ourselves to every specific details yet.

**Definition 6** The Sixth GrC Model is in the categories of functions, random variables, and even generalized functions.

Fuzzy sets are described by membership functions, so granules can be regarded as membership functions; note that the First to Fifth GrC Models include fuzzy sets. Hence, we consider further generalizations: granules are functions, random variables (measurable functions), generalized functions (e.g. Dirac delta functions).

In the case, a granule is a function, we may require that the granular structure (the collection of granules) has the universal approximation property, namely, any function in the universe can be approximated by the functions in the collections. The membership functions selected in fuzzy controls do have such properties. In neural networks, the functions generated by the activation functions also have such a property [29].

In the case of probability/measure theory, quantum mechanics may be a good guiding example.

**Definition 7** The Seventh GrC Model is in the category of Turing machines.

For example, a collection of lemmas in a mathematical proof (mechanizable), a set of subprograms in a computer program, or a computer or cluster of computers in grid/cloud computing are granules in the model.

**Definition 8** The Ninth GrC Model is in the category of qualitative fuzzy sets.

This model was proposed after the Eighth GrC model. It has not been published in printing form yet. The idea is similar to the model that we have called it sofset (this is not a typo) [12]. It associates to each “real world” fuzzy set, a collection of membership functions; please watch for new development.

### Category Theory Based Models

Now we generalize the category of sets to general categories. It is somewhat a surprise that this is basically the same as the category of relational databases [10]. In other words, the abstract structures of data and knowledge are similar. After analysis, it seems reasonable; because in GrC approach, the basic unit of knowledge is a granule of data.

Let us set up some language for the Category Theory. A category consists of

1. A class of objects, and

2. A set  $\text{Mor}(X, Y)$  of morphisms for every ordered pair of objects  $X$  and  $Y$ , which satisfies certain properties. For this paper, the formal details are not important; we only need the language loosely.

Here are some examples.

1. The Category of Sets: The objects are classical sets. The morphisms are the maps.
2. The Category of Sets with binary relations as morphisms: The objects are classical sets. The morphisms are binary relations. This is the Category of Entity Relationship Models.
3. The Category of Power Sets: The object  $U_X$  is the power set  $P(X)$  of a classical set  $X$ . Let  $U_Y$  be another object, where  $Y$  is another classical set. The morphisms are the maps,  $P(f): U_X \rightarrow U_Y$  that are induced by maps  $f: X \rightarrow Y$ .

Let  $CAT$  be a given category.

### Definition 9 (Category Theory Based GrC Model)

1.  $C = \{C_j^h, h, j = 1, 2, \dots\}$  is a family of objects in the category  $CAT$ .
2. There are families (which are bags; see the glossary) of Cartesian products,  $C_1^j \times C_2^j \times \dots$  of objects,  $j = 1, 2, \dots$  of various lengths. They are called product objects.
3. An  $n$ -ary relation object  $R^j$  is a sub-object of the product object  $C_1^j \times C_2^j \times \dots \times C_n^j$ .
4.  $\beta = \{R^1, R^2, \dots\}$  be a family of  $n$ -ary relations ( $n$  could vary).

The pair  $(C, \beta)$ , called the Categorical GrC Model (Eighth GrC model), is the formal model of granulation.

By specifying the general category to various special cases, we have all models: By specifying the category to be the category of sets, we have the Fifth GrC model. Further by limiting  $n$  to 2, we have the First GrC Model and the Fourth GrC Model. By assuming the symmetry for all  $n$ -ary relations, we have the Second GrC Model. By restricting the number of relations to be one and  $n = 2$ , we have the Third GrC Model. For the Sixth and the Seventh, further researches are needed.

### Overview of Early GrC Models

Schematically we summarize the relationships based on the Granular Structures of early GrC Models as follows: (the diagram will be different, if it is based on their approximation spaces). “ $\Rightarrow \Leftarrow$ ” is a two-way generalization but they are not inverse to each other. “ $\Rightarrow$ ,  $\uparrow$ , and  $\Downarrow$ ” are one

way generalizations. “GM” means GrC Models and RST means Rough Set Model.



### Granular and Related Structures

Four structures are related to each other, they are all briefly explained in Sect. “[Definition of the Subject](#).“ We will focus on a quotient structure here.

A granule can be examined from three states: We will use the following example to illustrate the idea. Let  $U$  be the set  $(\mathbb{Z}, +)$  of the integers as an additive group, let  $E$  and  $O$  be the set of even and odd integers respectively.

1. Isolated State (Internal State): Let us consider the case  $E$  is in the isolated state, that is, independent from the universe  $U$ . In this case, only the internal structure of  $E$  is available to us: we can only know that the number 2 is playing the role of the identity in an additive group, but is *not* aware of the existence of 1 in the outside of  $E$ . So in isolated state,  $E$  is the additive group  $(\mathbb{Z}, +)$ . Some time, this state may also be called internal state as only the internal structure is involved.

2. Embedded State (Conceptual State): In this state,  $E$  is a subgroup of  $U$ . We not only know that the number 2 is playing the role of the identity in  $E$  and we also *do know* that 2 is an even integer in  $U$ . [Digress] In category theory, a subobject (in this case a subgroup) is often represented by a pair (a group, the mapping of the group into the subgroup),

$$(Z, +) \rightarrow_{\text{homomorphism}} E \subseteq (Z, +).$$

The pair (the first  $(Z, +)$ ,  $\rightarrow_{\text{homomorphism}}$ ) is the subobject (subgroup). The first component  $(Z, +)$  represents the internal structure of (or isolated state of) even integers, the middle  $E$  is the embedded state of even integers. Since this state involves with every aspect of the concept of a granule, one may call it conceptual state.

3. Quotient State (External State): Note that the quotient structure  $Q$  consists of two elements, namely, the two subset of even and odd integers ( $E$  and  $O$ , respectively). But the two subsets appear as two elements,  $[E]$ , and  $[O]$  in  $Q$ . Note the bracketed symbols emphatically denote their roles as points (elements) of a set; no contents of subsets are visible. The interactions between the

two subsets induce interactions between these two elements. So  $Q$  is the integer mod 2 (as additive group). In notations,

$$Q = (\mathbb{Z}_2, +) \equiv \{[0]_2, [1]_2\}$$

as an additive group. So the quotient state of  $E$  is the element  $[0]_2$  in  $Q$ . Since the quotient state only involves the external relationships with other granules, this state may also be called external state.

With these preparations, we can define the concepts of granular structure (GrS), quotient structure (QS) and knowledge structures (KS): GrS is the collection of granules in embedded states. The quotient structure is the structure of granules in quotient states, namely, each granule is abstracted to an element (point) and the interactions among granules are abstracted to the interactions among elements (points).

Among these four structures only the quotient structure is a difficult concept. Here, we will illustrate only the quotient structure. We will start from the simplest case, namely, the GrS is a partition. By definition, a partition is a collection of equivalence classes that are mutually disjoint and their union is the whole universe. Hence, by abstracting each equivalence class to a point, we have a set of points that have no interactions. This implies:

**Proposition 1** *The quotient structure of a partition is a classical set.*

Next, we need a lemma.

**Lemma 1** *The collection of complete inverse image of a given mapping  $f: U \rightarrow V$  defines a partition. Namely, the collection  $\{f^{-1}(f(p)) \mid p \in U\}$  forms a partition.*

So NS and BNS induce partitions on  $U$ . These partitions have been called the derived partition (or derived equivalence relation); see [14].

Let  $B: U \rightarrow P(U)$  and  $NS: U \rightarrow 2^{P(U)}$  be a binary neighborhood system (BNS) and neighborhood system (NS), respectively. Then  $(U, NS)$  and  $(U, B)$  are First and Third GrC Models, respectively. By the previous Lemma, the complete inverse images of the two mappings,  $B$  and  $NS$ , are partitions (equivalence relations). The two partitions, denoted by  $E_B$  and  $E_{NS}$ , induce two quotient sets  $U/E_B$ ,  $U/E_{NS}$ , respectively. Moreover, the pre-topologies, BNS and NS, of  $U$  naturally induce the respective pre-topologies on the quotient sets  $U/E_B$  and  $U/E_{NS}$ , respectively [9,11].

**Proposition 2** *Such pre-topological spaces  $U/E_B$  (BNS-space) and  $U/E_{NS}$  (NS-space) are the quotient structures of*

*Binary and Local GrC Models (Third and First GrC Models) respectively.*

This example indicates that there are additional pre-topological structures on the information tables [2,22].

Let  $(U, \beta)$  be a Global GrC Model, where  $\beta = \{F^1, F^2, \dots\}$  be a partial covering, We will use  $[F^{j_1}], [F^{j_2}], \dots$  to denote a set of points (when we think of  $F^1, F^2, \dots$  as points).

**Proposition 3** *The granular structure,  $\beta$ , generates a semi-group  $S(\beta)$  under set theoretical intersection. Then the quotient structure of  $(U, S(\beta))$  is a semi-group generated by  $[F^{j_1}], [F^{j_2}], \dots$  under “intersection”  $[F^{j_1}] \circ [F^{j_2}] = [F^{j_1} \cap F^{j_2}]$ .*

This example indicates that there are additional algebraic structures on the information tables [20].

We observe that each quotient structure is not easy to determine. It depends on what are the mathematical structures under consideration. Let us recall some works from pure mathematics. Let the universe  $U$  be the ring  $Z$  of integers. In ring theory, the collection  $\beta$  of all prime ideals is the center of attentions. Let  $p$  be a prime number, then the prime ideal consists of following set

$$\{\dots, -2p, -p, 0, p, 2p, \dots\},$$

together with some algebraic structure.

By regarding each prime ideal as a point, the collection of prime ideals is a set, often denoted by  $\text{Spec}(Z)$  in algebraic geometry. The structure of prime ideals turns  $\text{Spec}(Z)$  into a topological space under Zariski topology.

*Example 3 (Two granular and quotient structures)*

1. The quotient structure of  $(U, \beta)$ , where  $\beta$  is the collection of the prime ideals, is a topological space  $\text{Spec}(Z)$  [27].
2. The quotient structure of  $(U, S(\beta))$ , where  $S(\beta)$  is the semi-group generated by the intersection, is isomorphic to the semigroup of positive integers.

In RST, the approximation spaces of the first and second examples are different. However in GrC, the approximations (under the knowledge engineering view) of the two examples are the same. In GrC, granules represent known basic units of knowledge (known concepts), hence the intersections of known concepts are known concepts. So in GrC we take all possible intersections of granules to approximate unknown concepts. This example clearly indicates the effectiveness of this view.

## Integration – A Dual of Granulation

Classical Divide (partitioning) and Conquer is extended to Granulate and Conquer. In other words, we are considering the cases, in which the sub-problems may not be logically independent from each other. So the standard thinking may not work, for example

- It may take Np-hard time to granulate a problem, from top to bottom.

This issue is not new. Some cases have been addressed in the topic of “dynamic programming” in data structure courses. We have used “topological divide” to solve the cases in the Third GrC model [18].

In this section, we will explore the following problem: For convenience, we will use the term, sub-structure, to denote (1) the collection of the internal structures of each granule and (2) the quotient structure. Now, we raise the following questions:

- Can we find more than one universes (granular models) that have the same sub-structure?
- Equivalently: Could two distinct problems, after granulation, have the same sub-structure?

In homological algebra, this is called the extension problem; in computer science, we feel that integration is a better name.

We will illustrate the concept in the following two cases

- 1) Simple Integration: no information structure
- 2) Integration with information structure.

## Simple Examples of Integration

This is a truly unexplored area; we will explain the case when granulations are partitions.

**A Simple Partition** We will illustrate the key concepts by simple examples. Let the universe  $U$  be the set of all integers,  $Z = \{\dots, 1, 0, 1, \dots\}$  with the following granular structure, namely, a given collection of two granules

$$\{\{\dots, -2, 0, 2, \dots\}, \{\dots, -3, 1, 3, \dots\}\}.$$

We name the granules  $E$  and  $O$  (the even and odd integers). Using the languages of Sect. “[Granular and Related Structures](#),” we say  $E$  and  $O$  are in the embedded state. We will use  $[E]$  and  $[O]$  to denote the quotient state. So the set  $Q = \{[E], [O]\}$  is the quotient set. Intuitively,  $Q$  is a set of black boxes; the internals of black boxes are invisible.

Now we will summarize the important structures of  $U$  so that integration can be formulated: Let  $\text{Int}(X)$  denote

the internal structure of a granule  $X$ , that is, the structure of  $X$  in isolation (and not a subset of  $U$ ).

1.  $A = \text{Int}(\{\dots, -2, 0, 2, \dots\})$  is the set of integers  $Z$ .
2.  $B = \text{Int}(\{\dots, -3, 1, 3, \dots\})$  is the set of integers  $Z$ .
3. Two copies of integers,  $A$  and  $B$ , are mapped to even integers and odd integers respectively.
4.  $Q$  is the quotient set that consists of  $[E]$  and  $[O]$  as two elements (points).

Schematically, we give the following situation:

$$Z = \begin{cases} A = \text{Int}(E) \rightarrow E \\ B = \text{Int}(O) \rightarrow O \end{cases} \quad (\subseteq) U \rightarrow Q = \{[E], [O]\} .$$

**Integrations on Partitions** From the point of view of problem solving,

- The quotient structure represents the recipe (a set of higher level instructions) of integrating the sub-solutions. It is the “Main Program” that integrates the returns of sub-program calls.

Here a sub-solution means the solution of a sub-problem that has been solved in isolation. Now the integration problem can be formulated as follows:

1. We only know that two copies of integers,  $A$  and  $B$ , are mapped to the two subsets of unknown universe.
2. The images of the mappings form a partition in the unknown universe  $U$
3. But we do know its quotient structure  $Q = \{[E], [O]\}$ , namely a set of two elements, where  $[E]$  represents the point that is abstracted from the image  $E$  of  $A$  in  $U$ ; similarly  $[O]$  is that of  $B$ .

Schematically, we can summarize it as follows: The unknown  $U$  has two unknown subsets, image  $A$  and image  $B$ , that form a partition on unknown  $U$ , but with a known quotient set which is equivalent to the integers mod 2.

$$Z = \begin{cases} A \sim \rightarrow \text{Image } A \\ B \rightarrow \text{Image } B \end{cases} \quad (\subseteq) \text{ Unknown } U \rightarrow Q = Z_2$$

*Can we construct the unknown universe? And is it unique?* The answer is “yes” and it is unique. It is the Cartesian product  $Q \times Z$ . Observe that from a set-theoretical point of view,  $Q \times Z$  is equivalent to  $Z$ , so the constructed one is the old friend.

### Integration on Partition with Information Structure

Let us consider a second view on the set of integers. But this time, the universe carries additional information,

namely, the additive structure of integers,  $(Z, +)$ . This universe is denoted by  $(U, +)$ . Then

1.  $\text{Int}(E)$  is the *additive group*  $(Z, +)$  of integers, and  $\text{Int}(O)$  is a set  $Z$  of integers.
2. The quotient structure  $(Q, +) = \{[E], [O]; +\}$  is an additive group:  $[E] + [E] = [E]$ ,  $[E] + [O] = [O] + [E] = [O]$ ,  $[O] + [O] = [E]$ . This  $(Q, +)$  is often called the integer mod 2, and denoted by  $(Z_2, +)$ .

Again, we give a similar situation

$$\begin{array}{ccc} (Z, +) = \text{Int}(E) & \xrightarrow{\text{(homomorphism)}} & E \\ Z = \text{Int}(O) & \xrightarrow{\text{(map)}} & O \\ & & \subseteq (U, +) \rightarrow (Q, +) . \end{array}$$

The upper arrow

$$(Z, +) = \text{Int}(E) \rightarrow_{\text{(homomorphism)}} (U, +) \rightarrow (Q, +)$$

is very similar to a short exact sequence (SES) in homological algebra.

Now, we will extend the SES to granules:

$$\text{Int}(\text{GrS}) \rightarrow_{\text{map}} U \rightarrow \text{QS} ,$$

where (1)  $\text{Int}(\text{GrS})$  is the collection of the internal structures of all granules, (2) its image in  $U$  (under the map  $\rightarrow_{\text{map}}$ ) is  $\text{GrS}$ , the granular structure in  $U$ , and (3)  $\text{QS}$  denote the quotient structure of  $\text{GrS}$ . The map,  $\rightarrow_{\text{map}}$ , is on-to-one within each granule, but two granules may have overlapping images in  $U$ . We will call this a *granular exact sequence*.

Can  $(U, +)$  be reconstructed back?

The answer is more than yes; there are *two* solutions! The two solutions are:  $(Z_2 \times Z, +)$  and  $(Z, +)$ . They are not equivalent as additive groups. This fact can be expressed by the extension functor, namely,  $\text{EXT}(Z, Z_2) \neq 0$ .

The important question is

Could such a functor be extended to formal granular models?

The answer may be “yes.” The granular exact sequences shall play a similar role as short exact sequences. There are some preliminary results in [19]; for example, the fact that the knowledge representation of a symmetric binary relation is complete, implies that the integration is unique.

### Integrations on Non-Partition Case

We will consider a simple example using the concept of granular exact sequence. Let  $U = \{a, b, c\}$ , let  $\beta$  be the set

of all subsets of  $U$ . So the internal structure of the granules are: (we will use  $n$ -granule to denote a granule of  $n$  elements) One copy of 3-granule; three copies of 2-granule, and three copies of 1-granule. This is a Global GrC Model. The quotient structure is a collection of 7 elements (to find the quotient structure is not easy, but easy to verify). These elements are related by partial ordering, called “a face of.” We will use the geometry to represent this quotient structure; it is a triangle spanned by  $\mathbf{i}(1, 0, 0)$ ,  $\mathbf{j} = (0, 1, 0)$ ,  $\mathbf{k} = (0, 0, 1)$ . The triangle can be viewed geometrically (a simplicial complex of closed triangle) as ONE open triangle, THREE open segments, and THREE vertices. Its a partial ordered set of 7 elements.

Now the granular exact sequence (three steps) can be described as follows:

1. 3-granule  $\{g_1^3, g_2^3, g_3^3\} \rightarrow$  a subset of unknown universe  $\rightarrow$  open triangle  $\Delta(\mathbf{ijk})$
2. first 2-granule  $\{g_1^2, g_2^2\} \rightarrow$  a subset of unknown universe  $\rightarrow$  first open segment  $\Delta(\mathbf{ij})$
3. second 2-granule  $\{g_3^2, g_4^2\} \rightarrow$  a subset of unknown universe  $\rightarrow$  second open segment  $\Delta(\mathbf{ik})$
4. third 2-granule  $\{g_5^2, g_6^2\} \rightarrow$  a subset of unknown universe  $\rightarrow$  third open segment  $\Delta(\mathbf{jk})$
5. first 1-granule  $\{g_1^1\} \rightarrow$  a subset of unknown universe  $\rightarrow$  first vertex  $\Delta(\mathbf{i})$
6. second 1-granule  $\{g_2^1\} \rightarrow$  a subset of unknown universe  $\rightarrow$  second vertex  $\Delta(\mathbf{j})$
7. third 1-granule  $\{g_3^1\} \rightarrow$  a subset of unknown universe  $\rightarrow$  third vertex  $\Delta(\mathbf{k})$

We will use  $I(g_j^i)$  to denote the image in the unknown universe. We do know many distinct  $g_j^i$  mapped to the same point in the quotient structure. Here are the mappings

1. 3-granule  $\{I(g_1^3), I(g_2^3), I(g_3^3)\} \rightarrow \Delta(\mathbf{ijk})$
2. first 2-granule  $\{I(g_1^2), I(g_2^2)\} \rightarrow \Delta(\mathbf{ij})$
3. second 2-granule  $\{I(g_3^2), I(g_4^2)\} \rightarrow \Delta(\mathbf{ik})$
4. third 2-granule  $\{I(g_5^2), I(g_6^2)\} \rightarrow \Delta(\mathbf{jk})$
5. first 1-granule  $\{I(g_1^1)\} \rightarrow \Delta(\mathbf{i})$
6. second 1-granule  $\{I(g_2^1)\} \rightarrow \Delta(\mathbf{j})$
7. third 1-granule  $\{I(g_3^1)\} \rightarrow \Delta(\mathbf{k})$

From the maps above, we may make some identifications: Observe that from the maps, 1st, 2nd, 5th, we can identify, without losing generality, that  $I(g_1^3) = I(g_1^2) = I(g_1^1)$ . By similar arguments, we have

1.  $I(g_1^3) = I(g_1^2) = I(g_3^2) = I(g_1^1) \rightarrow \Delta(\mathbf{i})$
2.  $I(g_2^3) = I(g_2^2) = I(g_5^2) = I(g_2^1) \rightarrow \Delta(\mathbf{j})$
3.  $I(g_3^3) = I(g_4^2) = I(g_6^2) = I(g_3^1) \rightarrow \Delta(\mathbf{k})$

So the 12 points are actually three points; they will be denoted by  $\{I(g_1^3), I(g_2^3), I(g_3^3)\}$ . Thus, the granular structure of the unknown universe is: one 3-granule  $\{I(g_1^3), I(g_2^3), I(g_3^3)\}$ , three of its 2-subgranules  $\{I(g_1^3), I(g_3^3)\}$ ,  $\{I(g_2^3), I(g_3^3)\}$ ,  $\{I(g_1^3), I(g_2^3)\}$  and three of 1-subgranules  $\{I(g_1^3)\}$ ,  $\{I(g_2^3)\}$ ,  $\{I(g_3^3)\}$ . So we have re-captured the unknown  $U$  and  $\beta$ . – This is the integration. The key question is: Is this  $U$  unique? The answer is “no”; we will skip it here.

### Semantic Views

Granules may be interpreted from three views.

1. Uncertainty Theory: A granule is a unit of lacking precise knowledge. Both L. A. Zadeh and T. Y. Lin, who coined the label, started from the uncertainty theory. Lin took his neighborhood system as a system of uncertainty. Zadeh has a grand project [40].
2. Knowledge Engineering: A granule is a unit of basic knowledge (Information). In their book, D. Stanat and D. McAllister state “Knowledge varies in sophistication from simple classification to ...” [35]. A classification is a partition, so an equivalence class is a (unit of) basic knowledge; this view is also promoted by Pawlak [30]. So we believe: a granule is a (unit of) basic knowledge.
3. How-to-Solve/Compute-it: A granule is a sub-problem or a software unit. It is a special type of basic knowledge.

Each view may have its own GrC theory: Some fundamental operators are:

- 1) *Information Hiding*: It is a transformation of granular structures into quotient structures (see glossary).

A quotient structure is the mathematical structure of the collection of granules, in which each granule is regarded as an element (point), and the interactions among granules are transformed into the interactions among elements. For example, in group theory a quotient group is a collection of cosets, in which each coset is regarded as an element and the multiplication of cosets is abstracted into the multiplication of elements in the quotient group. Using the software engineering language, a granule in a quotient structure is a black box, while in a granular structure, it is a white box. These are easy cases, for granulations that have nonempty overlappings, the quotient structure may not be easy to determine; see Section [Granular and Related Structures](#).

- 2) *Information Integration*: see Sect. “[Integration – A Dual of Granulation](#).”

- 3) *Knowledge Representation*: This amounts to give each point in the quotient structure (Sect. “[Granular and Related Structures](#)”) a meaningful name; we will call it naming map. So knowledge representation is a composition of information hiding (a map from a granular structure to a quotient structure) and the naming map.

Knowledge representation is another way to discover new knowledge by organizing the knowledge structure in an appropriate fashion, such as, an information table (= a relation in database theory). In GrC, the table often has additional algebraic and/or topological structures [18,22].

- 4) *Concept Approximation*: Among three semantic views, the knowledge engineering view is the most suitable view for this operation. Under this view, concept approximation is to express approximately the unknown concepts (arbitrary subsets of the universe) in terms of known basic knowledges (granules).

It is reasonable to regard that “and” ( $\cap$ ) or “or” ( $\cup$ ) operations of two basic units of known knowledge are also a known knowledge. So, we take finite intersection and any number of union as acceptable knowledge operations; the last one is derived from the topological spaces. However, we believe a negation of a known knowledge is not necessarily a piece of known knowledge; so negation is not an acceptable operation.

Note that this approximation is different from *Rough Set Approximations*, which, including its generalizations, are based on the sole operation “or” ( $\cup$ ). In other words, RST does not regard the “and” of two known concepts as also a known concept. So strictly speaking, rough set approximation is not a concept approximation.

GrC has many models; each model has a slightly different approximation theory. First, we will explain

The *Second GrC Model (Global GrC Model)*: Let  $C_1$  be a given collection. Let  $\mathcal{G}_1$  be the collection of all possible finite intersections of  $C_1$ . Then, by definition, the pair  $(U, \beta)$  is a Second GrC Model (Global GrC Model), where  $\beta = C_1$ . Let  $G$  be a variable that varies through the collection  $\mathcal{G}_1$ , then we define

**Definition 10** Three approximations

1. Upper approximation:

$$C[X] = \beta[X] = \{p : \forall G, \text{such that}, p \in G \ \& \ G \cap X \neq \emptyset\}.$$

2. Lower approximation:

$$I[X] = \underline{\beta}[X] = \{p : \exists \ a \ G, \text{such that}, p \in G \ \& \ G \subseteq X\}.$$

3. Closed set-based upper approximation: [33] used *closed* closure operator. It applies closure operator repeatedly

(transfinitely many steps) until the results stop growing. The space is called Frechet(V)-space or (V)-space.  $Cl[X] = X \cup C[X] \cup C[C[X]] \cup C[C[C[X]]] \dots$  (transfinite). For such a closure, it is a closed set.

**Theorem 1** *The concept approximation space of the Global GrC Model (Second GrC Model) is a topological space.*

We should also note that under the rough set approximation, this model is not a topological space.

Next, let us consider the approximation theory of The *First GrC Model (Local GrC Model)*: Before, we proceed, let us examine the classical case, the topological space  $(U, \tau)$ . A subset  $N(p) \subseteq U$  is a neighborhood of  $p$ , if  $N(p)$  contains an open set that contains  $p$ . The union of all such open sets is the interior points of  $N(p)$ , which is the largest open set in  $N(p)$ ; let us denote it by  $O(p)$ . Observe that  $O(p)$  consists of every point that regards  $N(p)$  as its neighborhood.

Now, we will generalize this idea to the First GrC Model. Let  $NS(p)$  be the neighborhood system at  $p$ . Let  $G(p)$  be the collection of all finite intersections of all neighborhoods in  $NS(p)$ . Let  $G$  be a variable that varies through  $G(p)$ .

**Definition 11** With such a  $G$ , the previous equations given above define the appropriate notions of  $C[X]$ ,  $I[X]$ ,  $Cl[X]$  for the First, Third and Fourth GrC Models.

We should caution here that in the Third GrC Model, there is at most one neighborhood at each point, so there is no “true” intersection.

Let  $N(p)$  represent an arbitrary neighborhood of  $NS(p)$ . Let  $C_N(p)$ , called the center set of  $N(p)$ , consist of all those points that have  $N(p)$  as its neighborhood. (Note that  $C_N(p)$  is the generalization of  $O(p)$  in TNS).

Now we will observe some harder question: Do the intersections of neighborhoods at distinct points belong to some  $G(p)$ ?

**Proposition 4 (Theorem of Intersections in NS)**

1.  $N(p) \cap N(q) \text{ is in } G(p) = G(q),$   
iff  $C_N(p) \cap C_N(q) \neq \emptyset$ .
2.  $N(p) \cap N(q) \text{ is not in any } G(p) \forall p,$   
iff  $C_N(p) \cap C_N(q) = \emptyset$ .

If we regard  $N(p)$  as a known basic knowledge, then we should define the knowledge operations: Let  $\circ$  be the “and” operation of the basic knowledge (a neighborhood). For technical reasons, the  $\emptyset$  is regarded as a piece of the given basic knowledge.

### Definition 12 (New operations in NS)

1.  $N(p) \circ N(q) = N(p) \cap N(q)$ , iff  $C_N(p) \cap C_N(q) \neq \emptyset$ .
2.  $N(p) \circ N(q) = \emptyset$ , iff  $C_N(p) \cap C_N(q) = \emptyset$ .

Observe that BNS is a special cases of NS. So we have:

**Definition 13** Let  $B$  be a BNS, then

1.  $B(p) \circ B(q) = B(p) = B(q)$ , iff  $C_B(p) = C_B(q)$ .
2.  $B(p) \circ B(q) = \emptyset$ , iff  $C_B(p) \cap C_B(q) = \emptyset$ . Note that  $B(p) \cap B(q)$  may not be empty, but it is not a neighborhood of any point.

Observe that in Binary GrC Model, two basic knowledges are either the same or the set theoretical intersection does not represent any basic known knowledge.

### 5) Higher Order Concept Approximations

In the Fifth GrC model, we consider the relations (subsets of product space) as basic knowledge. Any subset in a product space is a new unknown concept. We will illustrate the idea in the following case:  $U^j$  is either a copy of  $V$  or  $U$ . Moreover, in each product space, there is at most one copy of  $V$ , but no restrictions on the number of copies of  $U$ . If a Cartesian product has no  $V$  component, it is called a  $U$ -product space. If there is one and only one copy of  $V$ , it is called a product space with unique  $V$ .

1.  $u$  and  $u^1$  are said to be directly related, if  $u$  and  $u^1$  are in the same tuple (of a relation in  $\beta$ ), where  $u \in U$  and  $u^1$  could be an element of  $U$  or  $V$ .
2.  $u$  and  $u^2$  is said to be indirectly related, if there is a finite sequence  $u_i, i = 1, 2, \dots, t$  such that (1)  $u_i$  and  $u_{i+1}$  are directly related for every  $i$ , and (2)  $u = u_1$  and  $u^2 = u_t$ .
3. An element  $u \in U$  is said to be  $v$ -related ( $v \in V$ ), if  $u$  and  $v$  are directly or indirectly related.
4.  $v$ -neighborhood,  $U_v$ , consists of all the  $u \in U$  that are  $v$ -related.

In such a relational GrC model ( $U^j$  with unique  $V$ ) induces a map:

$$B: V \rightarrow 2^U; v \rightarrow U_v.$$

Such a map defines a binary neighborhood system (BNS), where  $U_v$  is a  $v$ -neighborhood in  $U$ , and hence induces a *binary GrC model* ( $U, V, B$ ). Next, we will consider the case  $U = V$  and define

**Definition 14** The high order approximations of the Fifth GrC model are the approximations based on the  $v$ -neighborhood system.

### 6) Concept Approximations in other Categories

In a category of functions, we will be interested in those granular structures that have the universal approximation property. For a category of Turing machines (algorithms), it is still unclear how to define the concept approximations.

### Future Directions

Granular Computing is still in its inception stage; possible directions are wide open. Here we will focus only on those issues that are touched in this article.

#### 1) Developments of Categories

In this paper, a category based model is proposed as the Formal Model for GrC. It can be specialized into various models to realize all the classical examples, including the first example, the granulation of the human body. We should note that the claim on the realization of the first example is not in printing form yet. However, the author feels that it is important to inform the readers that is occurring.

The key to realize the first example is based on the category of qualitative fuzzy sets or softsets (this is not a typo); please watch for new development. The categories of functions, random variables (measurable functions) and Turing machines must be developed, too.

#### 2) Developments of Granular Structures

Given a granular structure, we associate it with four structures (including itself). Among them, quotient and knowledge structures are mathematical consequences of a granular structure (if it is given mathematically). However, the linguistic structure is not a mathematical formalism but is a natural language formulation. In this paper, there is no report on this direction. We urge the readers to read Zadeh's article.

#### 3) Imported Concepts

For information integration (this may correspond to Zadeh's term, "organization"), we have illustrated the idea imported from homological algebra. It is unclear if we have imported the correct thinking; but it does point out essential problems in granulate and conquer.

#### 4) GrC and RST

RST has been served as the "model" of GrC developments. So, even there are a lot of similarities here, we would like to caution the readers that there are fundamental differences. For example, the fundamental views of uncertainty are quite different; Pawlak used "unable to specify" as the base of uncertainty, while GrC regards a granule as a unit of uncertainty (such as

uncertainty in quantum mechanics) Also the approximation theories are different. Of course, there are other differences; we skip.

### 5) GrC, Databases and Data Mining

As we have pointed out that the categorical structures of databases and GrC are similar; at the same time, we need to point out the differences in semantics. Nevertheless, we are looking forward to the transfer of database technology to GrC. For data mining, please see the database section on the articles by this author on deductive data mining using GrC, and mining decision rules using RST.

### 6) GrC and Fuzzy Logic

Most of the expositions have been based on classical sets (and fuzzifiable concepts). For more intrinsic fuzzy view, we strongly recommend the readers to read Zadeh's article.

### 7) GrC and Clouding Computing

Theoretically, cloud computing can be related to the GrC on the category of Turing machines. We expect some strong interactions in near future.

As we have observed that GrC is deeply rooted in human thinking, we expect GrC will have many interactions with wide variety of areas.

## Bibliography

1. Baliga P, Lin TY (2005) Kolmogorov complexity based automata modeling for intrusion detection. In: 2005 IEEE international conference on granular computing, Beijing, 25–27 July 2005. pp 387–392
2. Chiang IJ, Lin TY, Liu Y (2005) Table representations of granulations revisited. In: Proc 10th international conference, RSFD-GrC 2005, Regina, Canada, 31 Aug–3 Sept 2005, part I. Lecture notes in computer science, vol 3641. Springer, Berlin, pp 728–737
3. Dubois D, Prade H (1992) Putting rough sets and fuzzy sets together. In: Slowinski R (ed) Intelligent decision support: Handbook of applications and advances of the rough sets theory. Kluwer, Dordrecht, pp 203–232
4. Ginsburg S, Hull R (1981) Ordered attribute domains in the relational model. In: XP2 workshop on relational database theory, June 22–24, Pennsylvania State University, USA
5. Ginsburg S, Hull R (1983) Order dependency in the relational model. *Theor Comput Sci* 26:149–195
6. Hsiao D, Harary F (1970) A formal system for information retrieval from files. *Commun ACM* 13(2):67–73 (Corrigenda 13(4))
7. Lin TY (1988) Neighborhood systems and relational database. In: Proc of the 16th ACM annual conference on computer science, Atlanta, 23–25 Feb 1988. p 725
8. Lin TY (1989) Chinese wall security policy – An aggressive model. In: Proc of the 5th aerospace computer security application conference, Tuscon, 4–8 December, pp 286–293
9. Lin TY (1989) Neighborhood systems and approximation in database and knowledge base systems. In: Proc of the 4th international symposium on methodologies of intelligent systems (poster session), Charlotte, 12 Oct 1989. pp 75–86. (Available to the public from the National Technical Information Services, U. S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161. NTIS price codes rated copy: A11 Microfiche A01. Available to DOE and DOE contractors from the office of scientific and technical information P. O. Box 62, Oak Ridge, TN 37831; prices available from (615) 576-8401)
10. Lin TY (1990) Relational data models and category theory (abstract). In: CSC'90, Proc of the ACM 18th annual computer science conference on cooperation, Sheraton Washington Hotel, Washington DC, 20–22 Feb 1990. p 424
11. Lin TY (1992) Topological and fuzzy rough sets. In: Slowinski R (ed) Decision support by experience – Application of the rough sets theory. Kluwer, Dordrecht, pp 287–304
12. Lin TY (1996) A set theory for soft computing. In: Proceedings of 1996 IEEE international conference on fuzzy systems, New Orleans, 8–11 Sept 1996, pp 1140–1146
13. Lin TY (1997) Neighborhood systems – A qualitative theory for fuzzy and rough sets. In: Wang P (ed) Advances in machine intelligence and soft computing, vol IV. Duke University, Durham, North Carolina, USA, pp 132–155
14. Lin TY (1998) Granular computing on binary relations I: Data mining and neighborhood systems. In: Skowron A Polkowski L (eds) Rough sets in knowledge discovery. Physica, Heidelberg, pp 107–121
15. Lin TY (1998) Granular computing on binary relations II: Rough set representations and belief functions. In: Skowron A, Polkowski L (eds) Rough sets in knowledge discovery. Physica, Heidelberg, pp 121–140
16. Lin TY (1999) Granular computing: Fuzzy logic and rough sets. In: Zadeh L, Kacprzyk J (eds) Computing with words in information/intelligent systems. Physica, Heidelberg, pp 183–200
17. Lin TY (2003) Chinese wall security policy models: Information flows and confining trojan horses. In: De Capitani di Vimercati S, Ray I, Ray I (eds) Data and applications security XVII: Status and prospects, IFIP TC-11 WG 11.3 17th annual working conference on data and application security, Estes Park, 4–6 Aug 2003. Kluwer, Boston, pp 275–287
18. Lin TY (2005) Divide and conquer in granular computing: Topological partition. In: Proc of the 2005 North American fuzzy information processing society annual conference: Computing for real world applications, Ann Arbor, 22–25 June 2005
19. Lin TY (2006) A roadmap from rough set theory to granular computing. In: Proc of the 1st international conference rough sets and knowledge technology, RSKT 2006, Chongqing, 24–26 July 2006. Lecture notes in computer science, vol 4062. Springer, Berlin, pp 33–41
20. Lin TY (2006) Granular computing on partitions, coverings, and neighborhood systems. *J Nanchang Inst Technol* 25(2):1–7 (Special issue as the Proc of the international forum on theory of GrC from rough set prospective.)
21. Lin TY, Chiang IJ (2005) A simplicial complex, a hypergraph, structure in the latent semantic space of document clustering. *Int J Approx Reason* 40(1–2):55–80
22. Lin TY, Liau CJ (2005) Granular Computing and Rough Sets. In: Maimon O, Rokach L (eds) The data mining and knowledge discovery handbook. Springer, New York, pp 535–561
23. Lin TY, Huang KJ, Liu Q, Chen W (1990) Rough sets, neighborhood systems and approximation. In: Proc of the 5th international symposium on methodologies of intelligent systems, se-

- lected papers, Knoxville, 25–27 Oct 1990, pp 130–141. Library of Congress Catalog Card Number:90-84690
24. Lin TY, Sutojo A, Hsu JD (2006) Concept analysis and web clustering using combinatorial topology. In: Workshops Proc of the 6th IEEE international conference on data mining (ICDM 2006), Hong Kong, 18–22 Dec 2006. pp 412–416
  25. Liu Y, Lin TY, Xu C, Zhang Q, Huang L, He P (2008) Modeling complex architectures based on granular computing on ontology. Submitted to IEEE Trans Fuzzy Syst
  26. MacLane S (1995) Homology. Classics in mathematics, reprint of the 1975 edn. Springer, Berlin, x+422 pp
  27. Mumford D (1988) Introduction of algebraic geometry. In: The red book of varieties and schemes, mimeographed notes from the Harvard mathematics department 1967. Lecture Notes in Mathematics, vol 1348. Springer, Berlin
  28. Munkres J (2000) Topology, 2nd edn. Prentice-Hall
  29. Park JW, Sandberg IW (1991) Universal approximation using radial-basis-function networks. Neural Comput 3:246–257
  30. Pawlak Z (1991) Rough sets. Theoretical Aspects Of Reasoning About Data. Kluwer, Dordrecht
  31. Polya G (1957) How to solve it, 2nd edn. Princeton University Press, Princeton
  32. Robinson A (1966) Non-standard analysis. North-Holland, Amsterdam
  33. Sierpenski W (trans by Krieger C) (1952) General Topology. Mathematical exposition, no 7. University of Toronto Press, Toronto
  34. Spanier EH (1966) Algebraic topology. Springer, New York
  35. Stanat D, McAllister D (1977) Discrete mathematics in computer science. Prentice-Hall, Englewood Cliffs
  36. Wong E, Chiang TC (1971) Canonical structure in attribute based file organization. Commun ACM 14(9):593–597
  37. Zadeh LA (1979) Fuzzy sets and information granularity. In: Gupta M, Ragade R, Yager R (eds) Advances in fuzzy set theory and applications. North-Holland, Amsterdam, pp 3–18
  38. Zadeh LA (1996) The key roles of information granulation and fuzzy logic in human reasoning. In: 1996 IEEE international conference on fuzzy systems, New Orleans, 8–11 Sept, p 1
  39. Zadeh LA (1997) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst 90:111–127
  40. Zadeh LA (1998) Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems. Soft Comput 2:23–25
  41. Zimmerman H (1991) Fuzzy set theory – and its applications. Kluwer, Dordrecht

## Granular Computing, Principles and Perspectives of

JIANCHAO HAN<sup>1</sup>, NICK CERCONE<sup>2</sup>

<sup>1</sup> Department of Computer Science,  
California State University,  
Dominguez Hills, Carson, USA

<sup>2</sup> Faculty of Science and Engineering, York University,  
Toronto, Canada

## Article Outline

### Glossary

### Definition of the Subject

### Introduction

### Neighborhood System:

#### A Mathematical Structure of Granular Computing

### Rough Set System:

#### An Equivalence Neighborhood System

#### Fuzzy Set System: A Fuzzy Neighborhood System

#### Granular Computing in Data Mining

#### Granular Computing in Software Engineering

#### Future Directions

#### Bibliography

## Glossary

**Granular computing** Granular computing is a strategy of problem solving. The basic idea of granular computing comes from the strategy of divide-and-conquer. It is a twofold process: First it granulates a complex program into granules and then it computes these granules and integrates results to form a solution to the complex problem. Granulation of problems into granules is of different forms such as chunking, clustering, partitioning, division, or decomposition, while granules are clumps of objects or points. Computations with granules are either within granules or granule with environment.

**Neighborhood system** Neighborhood system is a mathematical structure of granular computing to model granules, and can be used to compute structure of granules and/or between granules and ambient spaces. A neighborhood system at a point is a framework to capture the concept of “near” objects, and any subset of objects can be approximated by a set of neighborhoods. A neighborhood system defines a set of binary relations, and a set of binary relationships can be used to define a neighborhood system.

**Fuzzy set theory** Unlike the classic set theory where a set is represented as an indicator function to specify if an object belongs or not to it, a fuzzy set is an extension of a classic set where a subset is represented as a membership function to characterize the degree that an object belongs to it. The indicator function of a classic set takes value of 1 or 0, whereas the membership function of a fuzzy set takes value between 1 and 0.

**Rough set theory** The rough set theory deals with inexact information systems. In an information system, a decision table consists of a set of objects which are characterized by a set of condition attributes and decision at-

tributes. Objects in the decision table can be classified into equivalence classes using an *indiscernibility* relation, and equivalence classes are explored to approximate crisp sets of objects.

**Data mining** Data mining is a very important step of knowledge discovery in databases to extract nontrivial, previously unknown, and potentially useful patterns that are hidden from large data sets. Data mining tasks include classification and clustering analysis of objects into categories, discovery of associations and correlations among data items, characterization and summarization of subsets of objects, finding sequential patterns and similarities in ordered data, etc.

**Software engineering** Software engineering is an engineering discipline that applies a systematic approach to produce reliable and efficient software. Software development process consists of different phases, including requirements, analysis, design, specification, implementation, testing, deployment, and maintenance. Object-oriented methodology to software engineering is based on several techniques and principles such as inheritance, modularity, polymorphism and encapsulation.

### Definition of the Subject

Granular computing (abbreviated, GrC) is a general computing paradigm that effectively deals with elements and granules that are generated from elements. Granules are vaguely viewed as generalized subsets of the universe of discourse and may be in different forms such as classes, clusters, groups, and intervals. The essence of GrC is to build an efficient computational model for complex applications with huge amounts of data, information and knowledge. Though the term has been proposed recently, the basic ideas and principles of granular computing have been explored in a variety of fields under the different names such as divide-and-conquer, and have been extensively applied in many problem solving disciplines such as politics, sociology, economics, computational intelligence with different forms either consciously or unconsciously by human beings.

Research on GrC will not only reveal the computational structures of complex problems and systems, build the mathematical foundations of complex problem solving, and formalize the common problem solving strategy, but also help practitioners explore novel specific problem solving strategies in special application domains. Additionally, by consciously using GrC strategies and heuristics, human being has a better chance to find a good solution to a complex problem.

### Introduction

The terminology of granular computing (GrC) was first proposed by Professor T. Y. Lin in 1996 as a label of family of theories, methodologies, and techniques in computational intelligence that make use of granules, although this strategy can be dated back to the ancient time. Its basic ideas and principles have been studied in various application domains. Especially in the form of partitions, the theory has been accumulated for hundred of years in mathematics.

Before the term, granular computing, was invented, some results and thoughts on granular computing had been achieved. The explicit study of granular computing is originated from Zadeh. In 1979, Zadeh [25] introduced the notion of information granulation and suggested that fuzzy set theory might find potential applications in this regard.

Though the granular computing is focusing on the non-partition theories, nevertheless, the partition case was the main source of inspiration, especially the partition theory in computer sciences. In 1982, Pawlak [21] proposed Rough Set Theory (RST) to deal with inexact information. It is an uncertainty theory using a very special form of granules, called equivalence classes. It is primarily due to the rough set theory (partition theory) that causes researchers to realize the importance of generalized granular computing notion.

In 1985, Hobbes [7] presented a theory of granularity as the base of knowledge representation, abstraction, heuristic search, and reasoning. In his theory the problem world is represented as various grains and only interesting ones are abstracted to learn concepts. The conceptualization of the world can be performed at different granularities and switched between granularities. Even though his discussions mainly concentrated on the partition cases, his model is more general than rough sets, and includes reflexive and symmetric binary relations.

In 1988, from the approximation retrieval, Lin introduced the notion of the neighborhood systems (NS) as a model of uncertainty in database systems [10,11], where a neighborhood is a unit of uncertainty. Lin studied the mathematics foundation of the neighborhood system that originates from the perspective of topology. A topological neighborhood system attaches to every point a collection of subsets that satisfy a set of axioms, called the axioms of topology. Lin removed the axioms and extended the theory to a quite general notion [12,13,14], where a neighborhood system corresponds to a set of binary relationships.

In 1992, Giunchiglia and Walsh [4] presented a theory of abstraction to improve the conceptualization of

granularities, where granules are abstracted in hierarchical levels so that relevant elements/attributes can be extracted but irrelevant details are ignored.

In 1997, Zadeh [26,27] refined his information granulation, presented granular mathematics, and applied to word computation. About the same time, Lin used the term “granular computing” to label this growing research field (1996). Since then, granular computing has received more and more attentions and much research has been conducted in various aspects of this area, and has begun to play important roles in various fields, including machine learning and data mining, bioinformatics, e-Business, network security, high-performance computing and wireless mobile computing, information hiding, and social works [3,17,18,22,23,27]. International conferences on Granular Computing sponsored by IEEE Computational Intelligent Society have been annually held since 2005 [8,28]. The essence of these models and applications has been addressed by researchers to build efficient computational algorithms for handling huge amounts of data, information and knowledge. The objectives of these computation models are computer-centered and mainly concern the efficiency, effectiveness, and robustness of using granules such as classes, clusters, subsets, groups and intervals in problem solving. In recent years, some researchers have investigated the granular computing paradigm from perspectives of philosophy, cognitive science, and human thinking as well as the general strategies of interactions between granules and operations on granule coverings.

Generally speaking, granular computing is a twofold process: Granulation and computation, where the former transforms the problem domain to one with granules, whereas the latter computes these granules to solve the problem. Granulation is a very natural concept and appears almost everywhere in different names, such as chunking, clustering, partitioning, division, decomposition, just to name a few. According to Zadeh [25] and Lin [13],

“information granulation is a collection of granules, with a granule being a clump of objects (points) which are drawn towards an object. In other words, each object is associated a family of clumps.”

Up to now, modeling and applying general GrC is not as successful as desired, although some special models have been innovated with great success. So it is immature at this point to speculate what the general definition, principle and methodology of granular computing should be. Therefore, we will look at some concrete theories that may be informally called granular computing. The discussion

will start from the neighborhood system and extend to the fuzzy set system, the rough set theory, and the quotient space. Then, the perspectives of granular computing in data mining and software engineering will be introduced, including granularity and granulation, granular relationships, and computation with granules.

### **Neighborhood System:**

#### **A Mathematical Structure of Granular Computing**

According to Lin [12], granular computing can be mathematically structured as a neighborhood system, which is abstracted from the geometric notion of “near” or “negligible distance”. Roughly, a neighborhood system assigns each object in the universe a (possibly empty, finite, or infinite) family of non-empty subsets, called neighborhoods, to represent the semantics of “near”. Neighborhoods play the most fundamental role in granular computing and can be considered as “granules”. A neighborhood is precisely a topological term of “granules”, and a neighborhood system is vaguely a topological space.

#### **Neighborhood System**

Fundamental notions on neighborhood systems were presented by Lin [12,13,14]. Assume  $U$  is the universe of discourse and  $p$  is an object in  $U$ .

1. An object  $x$  is a neighbor of  $p$  if  $x$  is “near” to  $p$ . The notion “near” is a subjective judgment and can be semantically interpreted in different ways, for example, if they satisfy some relationship or condition. For a specific case, let’s consider a binary relationship  $R$  in  $U$ .  $x$  is “near” to  $p$ , if  $xRp$ .
2. A neighborhood of  $p$ , denoted by  $N(p)$ , in  $U$  is a non-empty subset of  $U$ , which may or may not contain  $p$ . A neighborhood system of  $p$ , denoted as  $NS(p)$ , in  $U$  is a family of neighborhoods of  $p$ . If  $p$  has no neighborhoods, then  $NS(p)$  is an empty family.
3. A neighborhood system of  $U$ , denoted by  $NS(U)$ , is the collection of  $NS(p)$  for  $\forall p \in U$ . Formally,  $NS(U) = \{NS(p) | \forall p \in U\}$ .
4. Consider two neighborhood systems of  $U$ , say  $NS_1(U)$  and  $NS_2(U)$ .  $NS_1(U)$  is said a refinement of  $NS_2(U)$ , if for any neighborhood  $N_1$  in  $NS_1(U)$ , there exists a neighborhood  $N_2$  in  $NS_2(U)$  such that  $N_1 \subseteq N_2$ . Similarly,  $NS_2(U)$  is said a weak coarsening of  $NS_1(U)$ . (see [13] for strong refinement and coarsening).

From above, one can see, a neighborhood system is induced from the fundamental neighbors of objects, which are imprecisely defined as “near” to. If the “near” is explicitly defined according to a single binary relation, the result

neighborhood system is called a binary neighborhood system.

A binary neighborhood system BNS is defined as  $BNS = \langle U, B, V \rangle$ , where  $U$  and  $V$  are two universes, and  $B$  is a mapping function from  $U$  to  $V$ .

$$B: U \rightarrow V .$$

For  $\forall v \in V$ , let  $B_v$  denote the subset of  $U$ , that are source of  $v$  under the map  $B$ , that is,

$$B_v = \{u | u \in U \text{ and } B(u) = v\} .$$

Each  $B_v$  is called a basic neighborhood or a binary neighborhood.

One can infer that each neighborhood system  $BNS = \langle U, B, V \rangle$  defines a binary relation  $R \subseteq U \times V: \forall v \in V$  and  $u \in U$ ,  $uRv$  iff  $u \in B_v \subseteq U$ . Thus the binary relation  $R$  is fully determined by  $B$ . Conversely  $B$  is fully determined by a binary relation  $R$ .

If  $B$  is a binary relation  $R \subseteq U \times V$ , then for each object  $v \in V$ , the binary neighborhood in terms of  $v$  will be

$$N_v = \{u | u \in U \text{ and } uRv\} .$$

This is a subset of  $U$ , whose all elements are related to  $v$  by  $R$ .

A binary relation, in the usual sense, is a very special form of a neighborhood system, where each object has either zero or one neighborhood.

If  $U = V$  and the mapping function  $B$  is considered as a binary relation, then we have a simple binary neighborhood system, denoted as  $BNS = \langle U, R \rangle$ .

For simplicity, we call this binary neighborhood system with a binary relation as binary relation neighborhood system (BRNS).

Let's consider two binary relation neighborhood systems BRNS1 and BRNS2 built on the same universes, where  $BRNS1 = \langle U, R1, V \rangle$  and  $BRNS2 = \langle U, R2, V \rangle$ . BRNS1 is said a weak refinement of BRNS2, if  $R2 \subseteq R1$ , and a strong refinement if every neighborhood of BNS1 is a union of neighborhoods of BNS2. On the other hand, BRNS1 is said a weak/strong coarsening of BRNS2, if  $R1 \subseteq R2$ .

## Approximation

Assume  $BNS = \langle U, B, V \rangle$  is a binary neighborhood system and  $X \subseteq U$ .  $X$  can be approximated by neighborhoods of BNS.

Following topological spaces, the interior approximation, interior for short, of  $X$  is defined as either

$$\text{Interior}(X) = \bigcup_{x \in X} \{N(x) | N(x) \subseteq X\} ,$$

or

$$\text{Interior}(X) = \{x | x \in U \text{ and } \exists N(x) \subseteq X\} .$$

Similarly, the exterior approximation, exterior for short, of  $X$  can be defined as either

$$\text{Exterior}(X) = \bigcup_{x \in X} \{N(x) | N(x) \cap X \neq \emptyset\} ,$$

or

$$\text{Exterior}(X) = \{x | x \in U \text{ and } \forall N(x) \cap X \neq \emptyset\} .$$

The interior of a subset  $X$  of the universe can be approximated by the union of all neighborhoods of the universe that are completely contained in the subset, or by a subset of the universe that consists of all objects that have at least one neighborhood contained in  $X$ ; while the exterior of a subset  $X$  can be approximated by the union of all neighborhoods that have non-empty intersection with  $X$ , or by a subset of the universe that consists of all objects whose all neighborhoods have non-empty intersection with  $X$ .

A subset  $X$  of  $U$  is open if  $\forall p \in X$ , there is a neighborhood  $N(p) \subseteq X$ .  $X$  is closed, if its complement is open. A neighborhood system of  $p$ ,  $NS(p)$ , is open, if all neighborhoods of  $p$  are open. A neighborhood system of  $U$ ,  $NS(U)$ , is open, if  $\forall p \in U$ ,  $NS(p)$  is open.

If the universe  $U$  is a topological space and a neighborhood system of  $U$ ,  $NS(U)$ , is open, then  $NS(U)$  also defines the same topological space on  $U$ . In such a case, both  $NS(U)$  and the collection of open sets are called topology.

An object  $p$  is a limit point of a set  $E$ , if every neighborhood of  $p$  contains a point of  $E$  other than  $p$ . The set of all limit points of  $E$  is called derived set.  $E$  together with its derived set is a closed set.

One can easily verify that,  $NS(U)$  is discrete if every  $NS(p)$  is singleton and  $NS(U)$  is indiscrete if  $NS(U)$  is a singleton  $\{U\}$ .

A covering of  $U$  is an open neighborhood system  $NS(U)$ .

A partition of  $U$  is a covering of  $U$ , where  $\forall p, q \in U$ , and  $p \neq q$ , either  $NS(p) = NS(q)$ , or  $NS(p) \cap NS(q) = \emptyset$ .

## Rough Set System: An Equivalence Neighborhood System

Let's consider a special case of the binary neighborhood system  $BNS = \langle U, R \rangle$ , where  $R$  is an equivalence binary relation on  $U \times U$ . For an element  $u \in U$ , by definition, the binary neighborhood of  $u$  is defined as

$$N_u = \{x | x \in U, \text{ and } xRu\} .$$

Since  $R$  is an equivalence relation,  $N_u$  is the equivalence class  $[u]$  induced by  $u$  with respect to  $R$ . Thus, the neigh-

borhood system (the collection of all neighborhoods) will be the family of equivalence classes induced by  $R$ .

### Partition

Rough set theory was proposed by Pawlak in 1982 to deal with inexact information by using rough sets to approximate a crisp set [21]. By investigating the granularity of knowledge from the point of view of the rough set theory, one can view the rough set theory as a granular computing model based on partitions [22]. Basically, suppose  $U$  is a finite and nonempty universe. Let  $R \subseteq U \times U$  be an equivalence relation on  $U$ . The pair  $\langle U, R \rangle$  is called an approximation space. The relation  $R$  is a special BNS; it can be conveniently represented by a mapping  $B_R$  from  $U$  to the power set of  $U$ , and defined as

$$B_R : U \rightarrow 2^U,$$

$$B_R(x) = [x]_R = \{y \in U | xRy\}, \quad \forall x \in U.$$

The subset  $[x]_R$  is the equivalence class containing  $x$ , and is called an elementary subset. The family of all equivalent classes, denoted by  $U/R = \{[x]_R | x \in U\}$ , defines a partition of the universe  $U$ , namely, a family of pairwise disjoint subsets whose union covers the whole universe.

From the perspective of the rough set theory, each equivalence class is considered as a whole granule instead of many individuals. With these elementary subsets, any subset of  $U$  can be approximated.

### Approximation

Let  $X$  be a subset of  $U$  and  $R$  be an equivalence relation on  $U$ . The lower approximation of  $X$  based on  $R$ , denoted by  $\text{Low}_R(X)$ , is defined as

$$\text{Low}_R(X) = \cup\{Y \in U/R | Y \subseteq X\},$$

which contains all elementary subsets of  $U$  that are completely included in  $X$ .

The upper approximation of  $X$  based on  $R$ , denoted by  $\text{Upp}_R(X)$ , is defined as

$$\text{Upp}_R(X) = \cup\{Y \in U/R | Y \cap X \neq \emptyset\},$$

which contains all elementary subsets of  $U$  that have non-empty intersection with  $X$ .

In the partition model, these forms of approximations are equivalent to the interior and exterior defined in the BNS, respectively. However, they are not appropriate to be generalized to the NS. For example, in the topological space,  $\text{Upp}_R(X)$  is always the whole space, no matter what you choose.

### Information Systems

The rough set theory has found its applications in data mining, especially in classification and decision-making problem domains, called information systems.

An information system IS is defined as:  $IS = \langle U, C, D, \{V_a\}_{a \in C \cup D}, f \rangle$ , where  $U = \{u_1, u_2, \dots, u_n\}$  is a non-empty set of objects (tuples), called data set or decision table,  $C$  is a non-empty set of condition attributes, and  $D$  is a non-empty set of decision attributes and  $C \cap D = \emptyset$ .  $V_a$  is the domain of attribute  $a$  with at least two elements.  $f$  is a function:  $U \times (C \cup D) \rightarrow V = \cup_{a \in C \cup D} V_a$ , which maps each pair of object and attribute to an attribute value.

Let  $A \subseteq C \cup D$ , and  $t, s \in U$ . A binary relation  $R_A$ , called an *indiscernibility* relation, is defined as follows:  $R_A = \{(t, s) \in U \times U | \forall a \in A, t[a] = s[a]\}$ , where  $t[a]$  indicates the value of attribute  $a \in A$  of object  $t$ . The *indiscernibility* relation, denoted by  $\text{IND}(A)$ , is an equivalence relation on  $U$ . With this equivalence relation  $R_A$ , one can construct a binary relation neighborhood system  $\langle U, \text{IND}(A) \rangle$ .

Let  $X$  be a subset of  $U$  in an information system and represent a concept.  $X$  can be lower-approximated and upper-approximated by finding its lower approximation and upper approximation using elementary subsets of  $U/\text{IND}(A)$ . The lower approximation of  $X$  contains all objects in  $U$  which are *definitely* included in  $X$ , while the upper approximation of  $X$  contains all objects in  $U$  which are *potentially* included in  $X$ . On the other hand, one can see that the complement of the upper approximation of  $X$  contains all objects in  $U$  which are definitely excluded by  $X$ . As a concept,  $X$  has its lower approximation as the positive region, whereas the complement of its upper approximation as its negative region. From the perspective of machine learning or classification, the positive region contains the positive examples of  $X$ , while the negative region contains the negative examples of  $X$ . Between the positive region and the negative region is called the boundary region of  $X$ .

One main application of concept approximation using granules in the rough set theory is to reduce the size of a large decision table by finding an attribute reduct. A condition attribute  $a \in C$  is a core attribute of  $C$  in  $U$  with respect to  $D$  if

$$\forall X \in U/\text{IND}(D), \quad \text{Low}_{\text{IND}(C)}(X) \neq \text{Low}_{\text{IND}(C - \{a\})}(X).$$

A reduct of the condition attribute set is a minimum subset of it that has the same classification capability as itself. Thus a reduct of the condition attribute set can be used to represent the entire condition attribute set.

Formally, a subset  $A$  of  $C$ ,  $A \subseteq C$ , is defined as a reduct of  $C$  in  $U$  with respect to  $D$  if

$$\forall X \in U/\text{IND}(D), \text{ Low}_{\text{IND}(A)}(X) = \text{Low}_{\text{IND}(C)}(X), \\ \text{and } \forall B \subset R, \text{ Low}_{\text{IND}(B)}(X) \neq \text{Low}_{\text{IND}(C)}(X).$$

A condition attribute  $a \in C$  is said to be a reduct attribute if  $\exists B \subseteq C$ ,  $B$  is a reduct of  $C$  and  $a \in B$ . A reduct  $R$  of  $C$  is called a minimum reduct of  $C$  if  $\forall Q \subset R$ ,  $Q$  is not a reduct of  $C$ .

Assume  $P \subseteq C \cup D$  and  $Q \subseteq C \cup D$ , the positive region of  $Q$  with respect to  $P$ , denoted  $\text{POS}_P(Q)$ , is defined as

$$\text{POS}_P(Q) = \bigcup_{X \in U/\text{IND}(Q)} \text{Low}_{\text{IND}(P)}(X),$$

which contains all objects in  $U$  that can be classified using the information contained in  $P$ . With this definition, the degree of dependency of  $Q$  from  $P$ , denoted  $\gamma_P(Q)$ , is defined as  $\gamma_P(Q) = \text{POS}_P(Q)/|U|$ , where  $|X|$  denotes the cardinality of the set  $X$ .

The degree of attribute dependency provides a measure how an attribute subset is dependent on another attribute subset.  $\gamma_P(Q) = 1$  means that  $Q$  totally depends on  $P$ ,  $\gamma_P(Q) = 0$  indicates that  $Q$  is totally independent from  $P$ , while  $0 < \gamma_P(Q) < 1$  denotes a partial dependency of  $Q$  from  $P$ . Particularly, assume  $P \subset C$  and  $Q = D$ , then  $\gamma_P(D)$  can be used to measure the dependency of the decision attributes on a subset of conditional attributes. The task of the rough set attribute reduction is to find a subset of the conditional attributes that functions as the original conditional attribute set without loss of classification capability. This subset of the conditional attribute set is called *reduct*. One can prove that  $R \subseteq C$  is a *reduct* of  $C$ , if and only if  $\text{POS}_R(D) = \text{POS}_C(D)$ , or equivalently,  $\gamma_R(D) = \gamma_C(D)$ .

Thus, an attribute reduct is such a subset of condition attributes that the decision attribute has the same dependency degree on it as that on the entire set of condition attributes, and no attribute can be eliminated from it without affecting the dependency degree. Given an information system, however, for a given decision table, there may exist more than one reduct. Each reduct can be used as an alternative group of attributes to represent the original information system. It has been proved that every reduct must contain all the core attributes. However, finding all reducts of the condition attribute set is unfortunately *NP-hard*.

## Quotient Space

There multiple reducts of condition attributes in an information system may exist. Each reduct can be used to define

an equivalence relation. Thus it can be applied to construct a quotient space. From the perspectives of mathematics and topology, a quotient space is a space where equivalent points are glued together, and therefore a new space can be constructed.

Assume the universe of discourse  $U$  is a topological space,  $R$  is an equivalence relation on  $U$ .  $U/R$ , the set of all equivalence classes of  $R$ , is called a quotient space, where the topology on  $U/R$  is defined as below:

A subset  $X \subseteq U/R$  is open if and only if their union is open .

Topological properties of a quotient space as well as relationships between quotient spaces have been studied to identify and select attributes to form high-quality reducts. The quotient space theory has also been explored to solve information hiding, network security, and other problems [29].

## Fuzzy Set System: A Fuzzy Neighborhood System

In this section, another special case of binary neighborhood system is considered. Assume a binary neighborhood system  $\text{BNS} = \langle U, B, V \rangle$ , where  $U$  is a universe of discourse (the data space),  $V$  is a concept space, and  $B$  is a mapping function from  $U \times V$  to the unit interval. Formally,

$$B: U \times V \rightarrow [0, 1].$$

Since each point in the concept space is actually a concept, and can be assigned a name, Lin [14] suggested that this neighborhood system be defined as a 4-tuple system  $\langle U, B, V, C \rangle$ , where  $U$ ,  $B$ , and  $V$  are the same as before, while  $C$  is the concept name space. This system is called a fuzzy granular structure, and each concept in the concept space is a fundamental fuzzy clump or fuzzy neighborhood. In the simple case,  $B$  is a fuzzy binary relation and thus the 4-tuple system above is a single level granulation, called a fuzzy binary granular structure.

## Information Granularity

Information granularity with Fuzzy set was proposed by L. A. Zadeh in 1979 [25] to provide a basis for the construction of more general theories in which the evidence is allowed to be fuzzy in nature. In the information granularity, the data granules are viewed as a proposition in the general form of

$$g: "X \text{ is } G" \text{ is } \lambda,$$

where  $X$  is a variable taking values in a universe of discourse  $U$ ,  $G$  is a fuzzy set of the universe and is characterized by its membership function  $\mu_G$ , and  $\lambda$  is a fuzzy probability characterized by a possibility distribution over an unit interval.

A typical example of such a proposition is given in the universe  $U$  of real numbers, as below:

$g$ : "X is small" is likely ,

where the fuzzy subset  $G$  is "small" and the fuzzy probability  $\lambda$  is a fuzzy subset of the unit interval.

Each fuzzy subset  $G$  of the universe can be generally understood as a concept, for example, "small", "very large", "old", "young", etc. All concepts are put together to form a concept space, denoted by  $C$ ,

$$C = \{G_1, G_2, \dots, G_N\} .$$

Each proposition  $g$  as above is considered as an evidence, and all evidences comprise a collection of propositions,

$$V = \{g_1, g_2, \dots, g_N\} , \text{ where } g_i: "X \text{ is } G_i" \text{ is } \lambda_i , \\ i = 1, 2, \dots, N .$$

The proceeding information granularity can be characterized as a neighborhood system. Consider a mapping function  $B$  from  $U \times V$  to  $[0,1]$ , defined as:

$$B: U \times V \rightarrow [0, 1] ,$$

$$B(x, g) = \mu_\lambda(\mu_{G(g)}(x)) , \quad \forall x \in U \text{ and } \forall g \in V .$$

where  $\mu_{G(g)}$  is the membership function of the concept  $G$  in the proposition  $g$ . Thus the 4-tuple  $\langle U, B, V, C \rangle$  forms a neighborhood system in the form of fuzzy binary granular structure.

A special case of fuzzy granules is called non-probability-qualified granule and in the form of "X is  $G$ ". In this case, the corresponding fuzzy neighborhood system can be re-written as a 4-tuple system  $FNS = \langle U, B, V, C \rangle$ , where  $U$  is the universe of discourse (data space),  $V$  is a family of fuzzy subsets of  $U$ ,  $C$  is the name space whose names correspond to the fuzzy subsets in  $V$ , and  $B$  is a mapping function:

$$B: U \times V \rightarrow [0, 1] ,$$

$$B(x, v) = \mu_v(x) , \quad \forall x \in U \text{ and } \forall v \in V .$$

Consider the following example:  $U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ ,  $V = \{V_1, V_2, V_3, V_4\}$ , where (using discrete fuzzy

subset notation)

$$V_1 = 1.0/0 + 0.67/1 + 0.33/2 + 0/3 + 0/4 + 0/5 + 0/6 \\ + 0/7 + 0/8 + 0/9 ,$$

$$V_2 = 0/0 + 0.33/1 + 0.67/2 + 1.0/3 + 0.67/4 + 0.33/5 \\ + 0/6 + 0/7 + 0/8 + 0/9 ,$$

$$V_3 = 0/0 + 0/1 + 0/2 + 0/3 + 0.33/4 + 0.67/5 + 1.0/6 \\ + 0.67/7 + 0.33/8 + 0/9 ,$$

$$V_4 = 0/0 + 0/1 + 0/2 + 0/3 + 0/4 + 0/5 + 0/6 + 0.33/7 \\ + 0.67/8 + 1.0/9 ,$$

and the collection of words  $C = \{\text{Low}, \text{Median-Low}, \text{Median-High}, \text{High}\}$  to name fuzzy subsets  $V_1, V_2, V_3$ , and  $V_4$ .

Conditional granules can be represented as fuzzy rules such as "If  $X = u$  Then  $Y$  is  $G$ ". These kinds of granules can also be represented as fuzzy set membership functions or expressions of these functions [14].

## Approximation

In a fuzzy neighborhood system  $FBS = \langle U, B, V, C \rangle$ , the concept space  $V$  (a family of fuzzy subsets of  $U$ ) is also called a fuzzy covering of  $U$ , and each  $v$  in  $V$  is called a cover of  $U$ . For each  $u \in U$ ,  $v \in V$  is called a fundamental fuzzy neighborhood of  $u$ , if  $B(u, v) = \mu_v(u) > 0$ . Since each object  $u$  in the data space  $U$  may have multiple fundamental fuzzy neighborhoods,  $u$  can be represented by a vector of these fundamental fuzzy neighborhoods, called multi-valued representation for multi-level granulation. All fundamental fuzzy neighborhoods of  $u$ , put together, is denoted by  $FNS(u)$ .

Given a fuzzy neighborhood system  $FBS = \langle U, B, V, C \rangle$ , a subset  $X$  of  $U$  can be approximated by a membership function. The minimum approximation  $\min(X)$  of  $X$  in  $FBS$  can be defined as the following membership function:

$$\mu_X(x) = \min\{\mu_v(x) | v \in FNS(x)\} \text{ if } x \in X ,$$

and 0 otherwise .

Similarly, the maximum approximation  $\max(X)$  of  $X$  in  $FBS$  can be defined by the following membership function:

$$\mu_X(x) = \max\{\mu_v(x) | v \in FNS(x)\} \text{ if } x \in X ,$$

and 0 otherwise .

One can see from above that the minimum (maximum) approximation of a subset of  $U$  is a fuzzy subset of  $U$  that minimizes (maximizes) the membership degrees of its elements to their fuzzy fundamental neighborhoods.

Computing with words is to make inferences from words to words and is a very important topic of information granularity [27]. Fundamental neighborhoods can also be linearly combined to approximate formal words (concept names) [14]. Basically, let  $C = \{C_1, C_2, \dots, C_m\}$  be the concept space of a fuzzy neighborhood system, and  $V = \{V_1, V_2, \dots, V_m\}$  be the corresponding fuzzy subsets of the universe  $U$ . Assume  $r_1, r_2, \dots, r_m$  are real numbers. Consider the following linear expression:

$$r_1 C_1 + r_2 C_2 + \dots + r_m C_m = \sum_{i=1, \dots, m} r_i C_i.$$

Mathematically, the collection of all such expressions forms an abstract vector space, and each vector in the space is called a formal word, corresponding to a new concept. This new concept can be approximated by a fuzzy subset of  $U$ , of which the membership function is defined in terms of the membership functions  $\mu_{Vi}, i = 1, 2, \dots, m$ , for example, the minimum approximation or the maximum approximation discussed previously. The membership function of the new concept can also be defined as a linear combination of membership functions of its all fundamental fuzzy neighborhoods, i.e.

$$\mu_{\text{New Concept}}(x) = \sum_{i=1, \dots, m} r_i \mu_{Vi}(x).$$

In above example, a concept Median can be defined in the middle between the concepts Median-Low and Median-High, that is,

$$\begin{aligned} \text{Median} &= 0 * \text{Low} + 1/2 * \text{Median-Low} \\ &\quad + 1/2 * \text{Median-High} + 0 * \text{High}. \end{aligned}$$

This concept (word) can be approximated by the fuzzy subset defined below:

$$\mu_{\text{Median}} = 1/2 * \mu_{V2} + 1/2 * \mu_{V3},$$

and the fuzzy subset is:

$$\begin{aligned} 0/0 + 0.16/1 + 0.33/2 + 0.5/3 + 0.5/4 + 0.5/5 \\ + 0.5/6 + 0.33/7 + 0.16/8 + 0/9. \end{aligned}$$

### Granular Computing in Data Mining

Granular computing, as a strategy of problem solving and human thinking, has been extensively applied in knowledge discovery and data mining. For example, machine learning is to interpolate data and integrate concepts based on background knowledge that is viewed as granules. In

neural network computing, activation functions can also be regarded as granules, and thus the propagation and processing of activation functions can be viewed as a form of computations with granules. A gene is a granule, genetic algorithms process and transform genes which are actually granules, and therefore genetic algorithms can also be considered as one form of granular computing. In this section classification and clustering analysis and especially association mining are reviewed from the perspective of granular computing.

### Granular Computing in Classification and Clustering

Classification problem is, given a training set of examples with class labels, to construct a classifier that is able to assign a class label to a new example without class label that is not in the training set [6]. All examples are described with a given set of attributes. On the other hand, clustering is partitioning a data set by the similarity into clusters. Classes in classification and clusters in clustering are granules, subsets of examples, and the granulation is to partition the set of examples into granules. Computation with classes is to induce classification rules to characterize each class and predict the classes of future examples [9]. Similarly, computation with clusters is to summarize or characterize each cluster and discriminate between clusters.

The classification can be based on the refinement and coarsening relationships between granules [24]. A granule  $g_1$  is defined as a refinement of another granule  $g_2$ , or equivalently,  $g_2$  is a coarsening of  $g_1$ , if every sub-granule or object of  $g_1$  is contained in some sub-granules of  $g_2$ . Partitions and coverings are two simple and commonly used granulations of a universe. A partitioning of a universe is a collection of its non-empty and pair-wise disjoint subsets, whose union is the universe. On the other hand, a covering of a universe is a collection of its non-empty subsets that cover the whole universe. A partitioning is a special case of covering.

On the other hand, the clustering can be built upon the similarity relationship [24]. Similarity between granules is a basic and fundamental granular relationship and a key to forming an intrarelationship of a granule. Furthermore, it can be used to measure closeness or nearness amongst granules. Various measures can be exploited to calculate the similarity between two granules, which can be defined as the average distance between individuals in the two granules, and the distance between individuals can be measured as, for example, the Euclidean distance between points, the Levenshtein edit distance between strings or texts, the amino acid distance in biology, and the Mahalanobis squared distance between statistics data items.

### Granular Computing in Association Mining

Discovering association rules are another important area of data mining. The basic task of mining association rules is to find frequent itemsets which are granules, and then extract associations or correlations between items in the frequent itemsets, which is the computation with granules.

An association rule is an implication of the form  $A \rightarrow B$ , where  $A$  and  $B$  are subsets of regarded attributes, and  $A \cap B = \emptyset$ . For different types of attribute values the implication of the association form may have different meanings. The problem of mining association rules can be interpreted from the perspective of granular computing.

Mining association rules in large databases was first presented by Agrawal, Imielinski, and Swami in 1993 [2] to solve the basket data problem, in which, given a large database of customer transactions (basket data) that consist of items purchased by customers, significant associations between items are pursued. For example, “most often transactions that purchase bread and butter also purchase milk”, and “a customer purchasing tea is likely to also purchase coffee”. These significant associations are described as a set of association rules and quantitatively measured with support and confidence.

Formally, let  $I = \{I_1, I_2, \dots, I_m\}$  be a set of items, each  $I_k$  being an item, where  $m$  is the number of items. Let  $D$  be a database of transactions, where each transaction corresponds to a tuple in  $D$ , denoted by  $T$ . A transaction  $T$  is a set of items represented as a binary vector, with  $T[k] = 1$  if  $T$  contains the item  $I_k$ , and  $T[k] = 0$  otherwise, for  $1 \leq k \leq m$ . Assume  $X \subseteq I$  is a subset of items in  $I$ . A transaction  $T$  satisfies  $X$  if for all items  $I_k \in X$ ,  $T[k] = 1$ .

In the basket data problem, an association rule is of the form  $X \rightarrow Y$ , where  $X, Y \subset I$ , and  $X \cap Y = \emptyset$ . The rule  $X \rightarrow Y$  holds in  $D$  with confidence  $c$  if at least  $c\%$  transactions in  $D$  that contain  $X$  also contain  $Y$ , and support  $s$  if at least  $s\%$  transactions in  $D$  contain  $X \cup Y$ .

The task of mining association rules in  $D$  is to generate above implications with support and confidence greater than or equal to the user-specified support threshold  $\delta$  and confidence threshold  $\sigma$  respectively. Confidence is a measure of the rule strength, while support corresponds to statistic significance.

The problem of mining association rules can be decomposed into two subproblems:

1. Discovering all frequent itemsets: A subset  $X$  of  $I$  is a frequent itemset if its support is at least  $\delta$ ;
2. Generating association rules: For each frequent itemset  $X$ , construct implications of the form  $X - Y \rightarrow Y$ , for all  $Y \subset X$ . If the confidence of such a form is at least  $\sigma$ , it is considered as an association rule.

Currently many approaches to solving the basket data problem have been developed and efficient algorithms have been implemented. The classical solution exploits level-wise search to generate all candidate frequent itemsets and then prunes them [1]. The level-wise methods are based on the observation that if an itemset is frequent then all of its subsets are frequent. Thus, the basic idea is to evaluate 1-itemsets first, find frequent singleton itemsets, and then evaluate their supersets 2-itemsets, and so forth. Such evaluations are repeated until no frequent itemset is found or the size of the itemsets approaches some threshold.

Louie and Lin reformulate the association relations into bit data model [19] and Lin also reformulates these relations into granular model [15] which has been extended to a theory of attributes (features) [16].

From the perspective of granular computing, association rule mining can be formalized as a binary neighborhood system  $BNS = \langle I, B, D \rangle$ , where the universe of discourse is the set  $I$  of all items, transaction database  $D$  (the collection  $D$  of tuples) is a subset of the power set of  $I$ , that is,  $D \subset 2^I$ , and the binary relation  $B$  can be understood as: For all items  $i$  and  $j$ ,  $iBj$ , if and only if there exists an tuple in  $D$  that contains both  $i$  and  $j$ .  $D$  forms a covering of  $I$ , but not a partition of  $I$ .

An itemset  $X \subset I$  can be approximated by a subset  $C(X)$  of  $D$ , where  $C$  is defined as

$$C(X) = \{Y | Y \in D, X \subseteq Y\}, \text{ for all } X \subseteq I.$$

$X$  is a frequent itemset, if the cardinality of  $C(X)$ , denoted by  $|C(X)|$ , is greater than or equal to the support threshold  $\delta$ . A frequent itemset is called granule. The level-wise search process can be applied to find all granules.

The computation with granules in generating association rules is mainly within granules. For each subset  $Y$  of such a granule (a frequent itemset)  $X$ , one can find the approximation  $C(X)$  and  $C(Y)$  of  $X$  and  $Y$  respectively. The confidence of implication  $X - Y \rightarrow Y$  can be calculated as  $|C(X)|/|C(Y)|$ .

### Granular Computing in Software Engineering

In software engineering, the strategies of granular computing have been broadly used in all phases. The Granulate-and-conquer is a softer version of classical divide-and-conquer strategy that employs recursive algorithms to solve problems. A very common technique used in the classical “non-partitioning” recursive call is dynamic programming. Functional decomposition is to partition user requirement into granules (functions). Structured programming is to organize the computer programs as a collection of modules (procedures, functions, routines). The

intrarelationships among these granules are based on their ingredients such as input parameters, output results, executable statements, whereas the interrelationships involve the module interfaces and procedure calls. In object-oriented programming, granules are classes, intrarelationships are the interactions between components of classes such as instance fields, methods and constructors in the Java language; and the interrelationships are class inheritance, aggregation, association, delegation, dependency, etc.

In this section, the different phases of modern software engineering process are investigated from the granular computing perspective, based on object-oriented methodology [20], including requirement analysis, system analysis, design, and implementation.

## User Requirements

In the object-oriented methodology, user requirements analysis involves two main aspects: Identifying business actors and use cases. Both aspects can be conducted with granular computing paradigm. However, only identifying use cases will be investigated in this subsection.

Basically, each use case is a snippet of the business and may involve two-way communication between actors and the system. If the business requirement is considered as the problem domain, then the use cases will be the granules. Although there's no a set of rule for deciding how to granulate the business into use cases and most analysts partition the business process into use cases based on common sense, business logic, and their experience. The basic ingredients of the business requirement analysis from the granular computing perspective can be summarized as follows.

**Granules:** Use cases.

**Granulation method:** Top-down decomposition for complete covering and bottom-up combination to form a hierarchy of use cases.

**Granular relationships:** Some relationships between use cases can be recognized. For example, inheritance is a specialization and generalization relationship in which a large case is decomposed into small cases or a set of small cases are combined to constitute a large case. Inheritance is often referred as an is-a relationship: A granule  $g_1$  is a special case of another granule  $g_2$ . The is-a relationship plays a very important role in the object-oriented methodology and concept analysis, since " $g_1$  is-a  $g_2$ " reveals that  $g_1$  inherits all features and functions of  $g_2$  [24]. Other relationships between use cases include: Inclusion where the source case has some of its steps provided by the target case, and ex-

tension where the source case adds steps to the target case [20].

**Granular computation:** Identifying the process of use cases, including input data, output data and business logic, as well as the communication with actors.

## System Analysis

The basic goal of system analysis is to find candidate classes that describe the objects that might be relevant to the system, relationships between the classes, as well as attributes for the classes. With granular computing terminology, classes are granules that all together should cover the system requirement and satisfy the needs of use cases that have been identified before. The ingredients of system analysis from the granular computing perspective are summarized as follows:

**Granules:** Classes.

**Granulation method** is to identify classes. Candidate classes are often indicated by nouns in the use cases except those that represent the system, actors, boundaries, and trivial types. Two basic rules should be followed to identify classes: High cohesion inside classes and low coupling between classes. Class identification might be top-down and/or bottom-up, but the general complete coverage is necessary.

**Granular relationships** correspond to class relationships, which can include the following types [20]: Inheritance (is-a relationship) where a subclass inherits all of the attributes and behavior of its superclass(es); association, where objects of one class are associated with objects of another class; aggregation (strong association), where an instance of one class is made up of instances of another class; composition (strong aggregation) where the composed object can't be shared by other objects and dies with its composer; and others as well.

**Computation with granules** is two-fold. First, determine the internal structure and behaviors of objects of classes, including the instance fields and method/constructor prototypes; and second, interface the connections between classes. Associating classes often affects the design of internal structures of classes.

## System Design

Basically the software system design is to decompose a system into physical and logical components, and determine the technologies to be used to implement the system. Traditional system design focuses on the system functional partitioning, while modern software system design is based on object-oriented technology. In this subsection,

object-oriented system architecture design and technology are discussed from the perspective of granular computing.

System design involves multiple steps, including system topology, software partitioning, concurrency and security policies as well as communications between components. The following discussion will concentrate in the topology design and system partitioning.

The first task of current object-oriented system design is normally to determine the topology of a networked system. One popular system topology is the three-tier or multi-tier architecture to separate user interfaces, program logic and data in the system. In the three-tier architecture, the client tier presents the user interface to the user so that he/she can enter data and view results; the middle tier – also known as the business logic tier or server tier – runs multi-thread program code using large processors and lots of memory; and the data tier stores the data and provides safe concurrent access to it, typically with the help of a database management system. Thus, the system is granulated into three components. Besides the design of these three tiers, the protocols between tiers are also important.

To granulate the software, thus, object-oriented system design partitions the system components into layers: *User Interface, Control, Network, Server, Business, Persistence, and Database*.

## System Implementation

Software system implementation can also be viewed from the perspective of granular computing. The typical program structure in object-oriented software system is actually a granular structure. Consider a Java-based software system, which can be characterized as follows:

```

Software system
  Packages
    Classes
      Instance fields
      Constructors
      Methods
    Interfaces
    Method prototypes
  
```

In this structure, a software system is designed as a set of packages, which may be installed in different physical devices. Each package contains a set of interfaces whose subgranules are method prototypes and a set of classes which is partitioned into three types of components: instance fields, constructors, and methods.

During the design of classes, algorithms are one of designer's main concerns. Granular computing is also an

important strategy for this purpose. Let's consider the divide-and-conquer technique that is frequently used in algorithm design. The general paradigm of divide-and-conquer [5] is

**Divide:** Divide the problem  $S$  in two or more disjoint subsets  $S_1, S_2, \dots$

**Recur:** Solve the subproblems recursively.

**Conquer:** Combine the solutions to  $S_1, S_2, \dots$ , into a solution to  $S$ .

The base cases for the recursion are subproblems of constant size.

In the above paradigm, each subproblem is a granule, which can be granulated further into smaller subproblems. Granulation process usually follows the top-down method and the obtained granules should completely cover the parent problem. The granularity constitutes a hierarchy, where the top is the problem originally given, while at the bottom are all base cases. The computing with granules consists of not only dividing a non-base problem or solving a base problem but also combining solutions to subproblems to form a solution to the parent problem.

Another important general algorithm design paradigm is dynamic programming, which is actually a special case of divide-and-conquer, and thus a granular computing strategy. The main difference between them is: divide-and-conquer solves each subproblems individually, while dynamic programming stores solutions to all subproblems so that when a subproblem is reencountered, the solution can be obtained directly without resolving it.

## Future Directions

In this paper, basic principles and models of granular computing were briefly summarized. The ingredients of granular computing were investigated in terms of granulation methods and criteria, granular relationships, and computations with granules as well. Data mining technologies, especially association and correlation mining, and object-oriented technologies were reviewed from the perspective of granular computing. Granular computing methodology has attracted researchers and practitioners from various fields, but the unified architecture and generally accepted framework of granular computing have not been established. It is still an open question what granular computing is and how granular computing is performed. To answer such kind of questions, the future research on granular computing will be focusing on the following directions.

**Foundations of granular computing:** The mathematical foundation of granular computing will be one of important research topics, which will develop the gran-

ular structures of granules, granular relationships between granules, and computation with granules from the point of view of mathematics. Neighborhood systems, fuzzy set theory, rough set theory, and quotient space theory have initialized this research direction, but from different concerns and computing tasks. Granulation is based on partitioning and covering in terms of the nature of applications. Although it is difficult to propose a general process for granulation and in most applications granulation is performed by human experts, the basic principles and rules will be built for practitioners to follow. For example, granulation should meet the universal approximation properties. The infinite input space is granulated into finite granules, but finite granules should be able to approximate the universe of discourse in some ways.

**Framework of granular computing:** The hierarchical structure of granular computing will be established to represent the granules, relationships between granules, and computation with granules from the point of view of machine-centered computations. This research is originated from human thinking and problem solving and will be developed from philosophy, psychology, and behavior sciences. The general framework of granular computing will concentrate in the granulation and granules transformation.

**Information integration:** Basically, granular computing granulates complex problems into subproblems, conquer subproblems, and integrate solutions to subproblems to obtain the solution to the original complex problem. Granulate-and-conquer is the natural extension of the divide-and-conquer strategy, and addresses the integration of the subsolutions on granules into the total solutions on the original problem. For general granular computing, it is possible that two distinct problems be granulated into the same collection of granules. Different integrations of the same subsolutions may lead to the different results.

**Applications of granular computing:** Granular computing is strongly domain dependent. For example, in fuzzy logic control, the unknown control functions can be expressed by the membership functions of fuzzy granules; in structural programming, the programmers partition the functions according to their experience and background knowledge; while in rough set theory, the indiscernibility equivalence relation is provided in priori by experts and unknown concepts are approximated by equivalence classes. These examples demonstrate that granulation is not algorithmic and strongly domain dependent. Similarly, computation with granules is dependent on the domain

knowledge and the computation task. Extracting the common properties and features of granular computing in various application domains will be emphasized to enhance the general framework of granular computing and contribute to the foundation of granular computing.

## Bibliography

1. Agrawal R, Srikant R (1994) Fast algorithms for mining association rules. In: Proceedings of the 20th International Conference on Very Large Data Bases, Santiago, Chile, pp 487–499
2. Agrawal R, Imielinski T, Swami A (1993) Mining association rules between sets of items in large databases. In: Proceedings of the ACM SIGMOD Conference, Washington DC, 1993
3. Bargiela A, Pedrycz W (2002) Granular computing: An introduction. Kluwer Academic, Boston
4. Giunchiglia F, Walsh T (1992) A theory of abstraction. *Artif Intell* 56:323–390
5. Goodrich M, Tamassia R (2004) Algorithm design: Foundations, analysis, and internet examples, 2nd edn. Wiley, New Jersey
6. Gordon AD (1981) Classification. Chapman Hall, London
7. Hobbs J (1985) Granularity. In: Proc. of International Joint Conference on Artificial Intelligence, pp 432–435
8. Hu X, Liu Q, Skowron A, Lin TY, Yager RR, Zhang B (2005) Proc. of IEEE International Conference on Granular Computing. IEEE Press, Los Alamitos
9. James M (1985) Classification algorithms. William Collins, Glasgow
10. Lin TY (1988) Neighborhood systems and relational database. Abstract. In: Proceedings of CSC '88, pp 725
11. Lin TY (1989) Neighborhood systems and approximation in database and knowledge base systems. In: Proceedings of the Fourth International Symposium on Methodologies of Intelligent Systems, pp 75–86
12. Lin TY (1997) Neighborhood systems – a qualitative theory for fuzzy and rough sets. In: Wang P (ed) Advances in machine intelligence and soft computing, vol IV. Duke University, North Carolina, pp 132–155
13. Lin TY (1998) Granular computing on binary relations I: Data mining and neighborhood systems. In: Skowron A, Polkowski L (eds) Rough sets in knowledge discovery. Physica, New York, pp 107–121
14. Lin TY (1999) Granular computing: Fuzzy logic and rough sets. In: Zadeh LA, Kacprzyk J (eds) Computing with words in information/intelligent systems. Physica, New York, pp 183–200
15. Lin TY (2000) Data mining and machine oriented modeling: A granular computing approach. *Appl Intell* 13(2):113–124
16. Lin TY (2002) Attribute (feature) completion – the theory of attributes from data mining prospect. In: Proceedings of International Conference on Data Mining, pp 282–289
17. Lin TY (2003) Granular computing. In: Lecture notes in computer science, vol 2639. Springer, Berlin, pp 16–24
18. Lin TY, Yao YY, Zadeh LA (eds) (2002) Data mining, rough sets and granular computing. Physica, Heidelberg
19. Louie E, Lin TY (2000) Finding association rules using fast bit computation: Machine-oriented modeling. In: Proceedings of International Symposium on Methodologies for Information Systems, pp 486–494

20. O'Docherty M (2005) Object-oriented analysis and design. Wiley, New Jersey
21. Pawlak Z (1982) Rough sets. *Int J Comput Inf Sci* 11:341–356
22. Pawlak Z (1998) Granularity of knowledge, indiscernibility and rough sets. In: Proceedings of IEEE International Conference on Fuzzy Systems, pp 106–110
23. Skowron A, Stepaniuk J (2001) Information granules: Towards foundations of granular computing. *Int J Intell Syst* 16:57–85
24. Yao JT (2006) Information granulation and granular relationships. In: Proc. of IEEE GrC, pp 326–329
25. Zadeh LA (1979) Fuzzy sets and information granularity. In: Gupta M, Ragade R, Yager R (eds) Advances in fuzzy set theory and applications. North-Holland, Amsterdam, pp 3–18
26. Zadeh LA (1997) Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets Syst* 19:111–127
27. Zadeh LA (1998) Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems. *Soft Comput* 2:23–25
28. Zhang Y, Lin TY (2006) Proc. of IEEE International Conference on Granular Computing. IEEE Press, Los Alamitos
29. Zhang L, Zhang B (2003) The quotient space theory of problem solving. In: Lecture Notes in Computer Science, vol 2639. Springer, Berlin, pp 11–15

## Granular Computing System Vulnerabilities: Exploring the Dark Side of Social Networking Communities

STEVE WEBB<sup>1</sup>, JAMES CAVERLEE<sup>2</sup>, CALTON PU<sup>1</sup>

<sup>1</sup> College of Computing, Georgia Institute of Technology, Atlanta, USA

<sup>2</sup> Department of Computer Science, Texas A&M University, College Station, USA

### Article Outline

- Glossary
- Definition of the Subject
- Introduction
- Background on Social Networking Communities
- Traditional Attacks Targeting
  - Social Networking Communities
- New Attacks Against Social Networking Communities
- Future Directions
- Bibliography

### Glossary

**Social networking community** A community of individuals (users) that are connected based on mutually shared activities, beliefs, goals, or relationships.

**Profile** A user-controlled web page that contains a picture of a community user along with various pieces of personal information for that user. This is often the online representation of an individual in a social networking community.

**Friend request** A communication mechanism that allows community users to initiate connections between their profiles and the profiles of their friends.

### Definition of the Subject

Online social networking communities are connecting hundreds of millions of individuals across the globe and facilitating new modes of interaction. Due to their immense popularity, an important question is whether these communities are safe for their users. In this paper, we address this safety question and show that social networking communities are susceptible to numerous attacks. Specifically, we identify two attack classes: traditional attacks that have been adapted to these communities (e.g., malware propagation, spam, and phishing) and new attacks that have emerged through malicious social networking profiles (e.g., rogue advertising profiles and impersonating profiles). Concretely, we describe examples of these attack types that are observable in MySpace, which is currently the most popular social networking community.

### Introduction

Over the past few years, social networking communities have experienced unprecedented growth. Communities such as MySpace and Facebook are connecting people in a variety of new and exciting ways, and as a result, individuals are attaching an increasing amount of value to their online personas. Unfortunately, the rising importance and prominence of these communities have also made them prime targets for attack by malicious entities.

We observe two distinct attack classes that threaten social networking communities and the privacy of their users. First, traditional attacks that have plagued Internet users for many years (e.g., malware propagation, spam, and phishing) have been adapted to take advantage of the unique properties of online communities. These traditional attacks are thriving in their new environment, and the severity of these attacks promises to grow as attackers become more sophisticated. The second attack class consists of new attacks that have emerged from the very fabric of these communities. One prominent example is the use of deceptive profiles (e.g., rogue advertising profiles and impersonating profiles) that are becoming more widespread, difficult to detect, and extremely costly to legitimate community members.

In this paper, we provide detailed descriptions for each of these attack classes, and we show that the continued success of social networking communities is contingent upon their ability to mitigate the risks associated with these attacks. For concreteness, we describe attacks that are observable in MySpace, which is the most popular social networking community in terms of unique visitors (more than 65 million in February 2008) [3], total traffic (4.29% of all US Internet visits in February 2008) [6], and user base (more than 110 million active accounts as of February 2008) [14]. Our observations show the practical importance of these attacks and their impact on millions of users. Additionally, since other social networking communities are both functionally and structurally similar to MySpace, the attacks we describe can be easily adapted to most (if not all) social networking communities.

The rest of the paper is organized as follows. Sect. “Background on Social Networking Communities” provides background information about social networking communities. In Sect. “Traditional Attacks Targeting Social Networking Communities”, we describe traditional attacks that have adapted to target social networking environments, including malware propagation, spam, and phishing. In Sect. “New Attacks Against Social Networking Communities”, we present new attacks that specifically

target social networking communities by utilizing deceptive profiles such as rogue advertising profiles and impersonating profiles. Sect. “Future Directions” concludes the paper.

## Background on Social Networking Communities

Social networking communities provide an online platform for people to manage existing relationships, form new ones, and engage in a variety of social interactions. Typically, a user’s online presence in these communities is represented by profile, which is a user-controlled web page that contains a picture of the user along with various pieces of personal information. Profiles connect to other profiles through explicitly declared friend relationships and numerous messaging mechanisms.

An example of a MySpace profile is shown in Fig. 1a. Some of a profile’s personal information is mandatory (e.g., the user’s name, age, gender, location, etc.), and some of it is optional (e.g., the user’s interests, relationship status, occupation, etc.). To facilitate self-expression, each profile also has free text “About me” and “Who I’d like to meet” sections, and users are allowed to embed various objects in their profiles such as pictures, audio clips, and videos. In addition to personal information and embedded



**Example Profile**

This is an example MySpace headline.  
Male  
25 years old  
ATLANTA, GEORGIA  
United States

Last Login: 11/2/2006

[View My Pics](#) | [Videos](#)

**Contacting Example Profile**

- [Send Message](#)
- [Forward to Friend](#)
- [Add to Friends](#)
- [Add to Favorites](#)
- [Instant Message](#)
- [Block User](#)
- [Add to Group](#)
- [Park User](#)

**MySpace URL:** [http://www.myspace.com/example\\_myspace\\_profile](http://www.myspace.com/example_myspace_profile)

**Example Profile's Interests**

General	General interests go here.
Music	Favorite musicians go here.
Movies	Favorite movies go here.
Television	Favorite television shows go here.
Books	Favorite books go here.
Hobbies	Hobbies go here.

**Example Profile's Details**

Status:	Single
Orientation:	Straight
Ethnicity:	White / Caucasian
Zodiac Sign:	Gemini
Smoke / Drink:	No / No
Education:	Grad / professional school
Occupation:	Ph.D. student



**Example Profile**

This is an example MySpace headline.  
Male  
25 years old  
ATLANTA, GEORGIA  
United States

Last Login: 11/2/2006

**Contacting Example Profile**

- [Send Message](#)
- [Forward to Friend](#)
- [Add to Friends](#)
- [Instant Message](#)
- [Block User](#)
- [Add to Group](#)
- [Park User](#)

**MySpace URL:** [http://www.myspace.com/example\\_myspace\\_profile](http://www.myspace.com/example_myspace_profile)

a
Public Profile
b
Private Profile

**Granular Computing System Vulnerabilities: Exploring the Dark Side of Social Networking Communities, Figure 1**  
Example MySpace profiles. In a, the profile is publicly accessible. In b, the profile is private

content, a user's profile also contains a list of links to the profiles of that user's friends. These friend links are bidirectional because a link is established only after both parties acknowledge the friendship. To initiate a friendship, a user sends a friend request to another user. If the other user accepts this request, the friendship is established, and a friend link is added to both users' profiles.

Aside from friend requests, MySpace provides a number of other communication facilities that enable users to communicate with each other within the community. These facilities include messaging, bulletin, commenting, blogging, and instant messaging (IM) systems. The messaging system allows users to exchange intra-community email messages with any other user (i.e., both friends and strangers). The bulletin system is essentially an exclusive bulletin board that only a user's friends can view, enabling users to communicate with all of their friends at once. Users can post comments on their friends' profiles using the commenting system, and the blogging system allows users to maintain blogs on their profiles, which other users can read and comment on. Finally, the IM system provides users with a mechanism to send intra-community instant messages to any other user.

Due to the wealth of private information that is accessible on user profiles and the various means of communication that are available, MySpace provides mechanisms to protect the privacy of its users. First and foremost, users have the ability to choose between making their profiles publicly viewable (the default option) or private. If a user's profile is designated as private, only the user's friends are allowed to view the profile's detailed personal information (e.g., the user's interests, blog entries, comments, etc.). However, as Fig. 1b shows, a private profile still reveals the user's name, picture, headline, gender, age, location, and last login date. MySpace also provides a few finer-grained privacy mechanisms. Users can control who is allowed to IM them (everyone, only friends, or no one) and who is allowed to leave blog comments (everyone or only friends). Users can also maintain a block list to prevent specific users from contacting them at all. Unfortunately, as we will see in the following sections, these privacy mechanisms have been unable to prevent a number of attacks.

### Traditional Attacks Targeting Social Networking Communities

Since social networking communities include communication facilities that are fundamentally similar to traditional means (i.e., email, blogs, instant messaging, etc.), many of the attacks that are effective against those tradi-

tional communication media have been adapted to exploit social network communications. Due to the massive size of many social networking communities, their tightly connected nature, and their relatively naïve user bases, these communities are target rich environments for attackers. In this section, we describe three of the adapted attacks that have been observed in MySpace: malware propagation, spam, and phishing.

### Malware Propagation

Malware creators aim to spread their malicious content to as many victims as possible. Since MySpace is the most popular community on the web, it has become a prime target for malware propagation. In fact, over the past couple of years, MySpace was attacked by at least one instance of each of the following malware categories: worms, spyware, and adware. For the remainder of this section, we will detail the most interesting occurrences of these attacks, and we will explain the threats they pose to social networking communities.

**Worms** The most successful example of rapid worm propagation in a social networking community occurred in MySpace on October 4, 2005. The worm was called the "Samy worm," and it generated more than a million friend requests for its creator (Samy) over the course of a single day. The basic operation of this worm was quite simple but extremely clever [11]. First, Samy wrote the worm using Javascript, and then, he embedded it in his MySpace profile. MySpace disallows users from adding scripts to their profiles by filtering scripting tags and removing specific strings (e.g., "javascript"). To evade these filters, Samy exploited the behavior of popular web browsers such as Internet Explorer. Specifically, he hid the code inside a Cascading Style Sheet (CSS) tag and obfuscated the strings that MySpace would filter (e.g., "javascript" became "java\nscript"). Thus, even though MySpace had security mechanisms in place, the lax security of certain web browsers allowed the obfuscated code to execute successfully.

When a MySpace user accessed Samy's profile, the embedded Javascript code sent Samy a friend request on that user's behalf. The code also embedded itself in the user's profile (in the same manner that it was embedded in Samy's profile). Consequently, when other MySpace users visited the newly infected profile, the code would send Samy friend requests from those users and propagate itself to their profiles. As Samy described it, "If 5 people viewed my profile, that's 5 new friends. If 5 people viewed each of their profiles, that's 25 more new friends."

Fortunately, this worm was relatively harmless for its infected users because it only affected their MySpace profiles and not their actual machines. However, the same cannot be said for MySpace. Due to the worm's viral growth pattern, the MySpace administrators were forced to temporarily shut down the site to stop the worm's propagation and remove the worm's code from the infected users' profiles. The service outage was relatively brief, but this incident clearly illustrates the potential damage that social networking worms are capable of inflicting. Specifically, this worm teaches two very important lessons. First, the success of the worm's obfuscated code clearly highlights the importance of web site security as well as web browser security. A delicate balance exists between functionality and security, and this balance must be respected when designing and developing online communities. Second, the worm's propagation speed showcases how quickly the entire community could become infected. In this case, more than a million users were affected in less than 24 hours.

A much more dangerous social networking worm appeared on MySpace in the middle of July 2006. Similar to the Samy worm, this worm propagated itself through the profiles of unsuspecting MySpace users. However, unlike the Samy worm, this worm's code was hidden inside a malformed Shockwave Flash (.swf) file that directly exploited a vulnerability on users' machines. Additionally, instead of generating innocuous friend requests, this worm actually redirected users' web browsers to a politically charged blog posting.

The manner in which this worm spread is as follows. First, the worm's creator generated a malformed .swf file and embedded it in a MySpace profile. This .swf file exploited a critical vulnerability in Macromedia Flash Player v8.0.24.0 and earlier versions, which allows an attacker to execute code on an affected machine. When a MySpace user with a vulnerable Macromedia Flash Player accessed the profile, the code in the .swf file automatically redirected the user's browser to a blog posting that contained political propaganda. Finally, the code propagated itself by embedding a copy of the .swf file in the infected user's profile.

This worm was also relatively harmless; however, it was far more troubling than the Samy worm because it actually exploited a vulnerability on users' computers, which allowed it to execute code. Fortunately, the worm's creator was only interested in spreading political ideals because nothing prevented the worm from installing any number of malicious utilities on its victims' machines. For example, the worm could have installed spyware or adware, and it could have even zombified the infected machine (i.e., compromise the computer and enlist it in a bot-

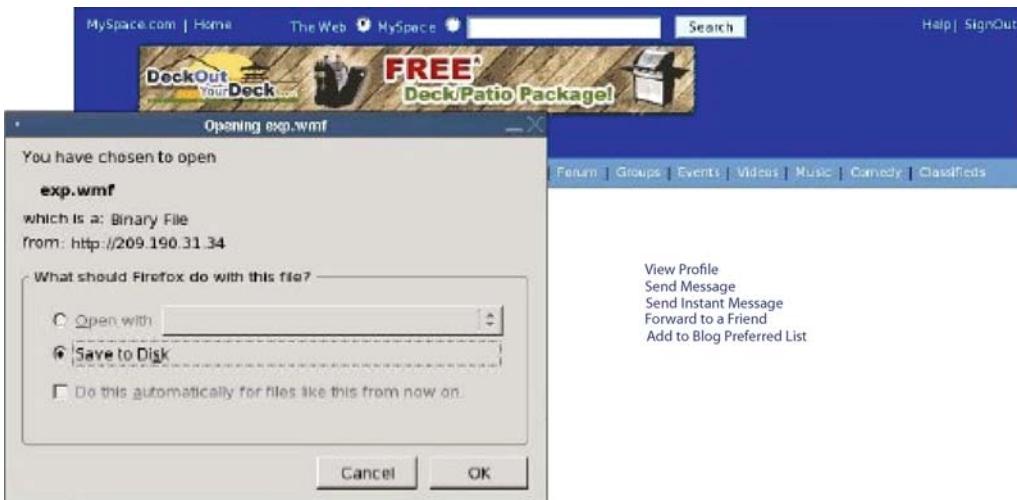
net [4]). Combine these frightening, yet completely realistic, scenarios with the viral propagation patterns of these worms, and it becomes immediately obvious how devastating worms could be in a social networking environment and why they must be prevented.

**Spyware** Although spyware has yet to be distributed within the payload of a social networking worm, it has already made a few appearances in social networking communities. The most prominent example of spyware appearing in MySpace occurred towards the beginning of July 2006. At that time, an advertisement for deckoutyourdeck.com, which contained a malformed Windows Metafile (.wmf) image, was inserted into one of the ad networks that MySpace uses. This malformed .wmf image exploited a critical vulnerability in the Graphics Rendering Engine of Windows, which allows remote code execution.

MySpace displays an ad banner, which is randomly selected from MySpace's supplying ad networks, at the top of every profile. When a user accessed a profile that displayed the deckoutyourdeck.com ad, the user was prompted to download the embedded .wmf image. Figure 2 shows a screenshot of this prompted download. If the user was running an unpatched version of Windows, the .wmf file installed an assortment of programs, including known spyware utilities such as PurityScan [9]. These spyware utilities pose a serious threat to the user because they track the user's web browsing behaviors, and they install various other third-party applications without the user's consent.

Unfortunately, despite the fact that Microsoft released a patch for this vulnerability more than six months before the ad appeared on MySpace, over a million users were affected. Thus, this incident reveals two very important points. First, many users have a false sense of security in these communities. One of the most basic secure browsing principles is never to download questionable content, yet more than a million users gladly accepted this suspicious (and completely unsolicited) download request. Second, an alarming number of users are not vigilant about protecting themselves against security threats. This incident proves that at least a million users neglected to install security patches for more than six months. Therefore, social networking communities must assume that their users are completely vulnerable and take every precaution necessary to protect them from malicious content.

**Adware** An interesting instance of adware appearing in MySpace occurred around the same time as the previous spyware example. A security researcher was browsing MySpace profiles and found two that were named after a known adware company called "Zango" (formerly



**Granular Computing System Vulnerabilities: Exploring the Dark Side of Social Networking Communities, Figure 2**  
Embedded .wmf image within a deckoutyourdeck.com advertisement

known as “180solutions”). Both profiles were created to deceive users into downloading and installing the Zango Search Assistant and Toolbar (two known adware programs) [2]. The first profile claimed that the programs could “protect kids from predators,” and the second profile was even more deceptive.

When a user accessed the second Zango profile, a popup with a license agreement immediately launched, asking the user to accept the license in order to play a video file. This license popup is shown in Fig. 3. If the user accepted the license by clicking the “Play Now” button, a video file began to play, but secretly, the Zango Search Assistant and Toolbar were also installed on the user’s system. As Fig. 3 illustrates, this popup contained a number of deceptive elements. First, the popup did not explicitly indicate that an installation was taking place. The text in the top-left corner of the popup mentions the Zango Search Assistant and Toolbar, but it does not tell the user that they will be installed. Additionally, none of the popup’s buttons mention an installation. Instead, they are entitled “Play Now” and “Play,” which implies that the only action will be the play-back of a video file. Finally, the most prominent features of the popup are the video’s preview window in the top-right corner and the “Play Now” button. The actual license agreement and its preselected checkbox are positioned at the bottom, where they are likely to be overlooked. Consequently, many users played the video before they noticed the license agreement, and the adware was successfully installed on their machines.

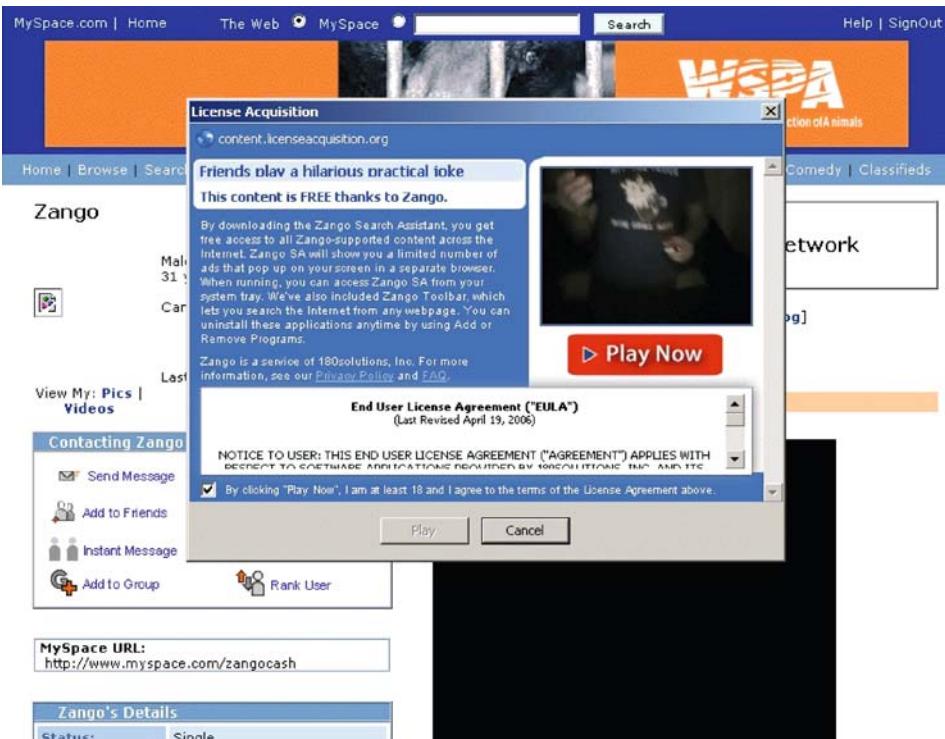
According to Zango, both profiles were created by a Zango developer that was acting against the company’s

policy of not targeting MySpace. As a result, the deceptive video clip was removed, and the company released a public apology. However, this occurrence clearly illustrates the potential for abuse by adware companies in social networking communities. Additionally, the success of this deception further illustrates the naïvety of users and the need for these communities to provide effective security mechanisms that protect against malicious content.

## Spam

In addition to malware propagation, another type of attack on traditional communication media (e.g., email, blogs, instant messaging, etc.) is spam. Since MySpace provides similar communication facilities, they are also susceptible to spamming abuse. In fact, almost all spamming activities that occur outside the community can be recreated inside the community. To aggravate the problem, spammers can also use the profile information posted by users to target them outside the social networking community. Although MySpace’s Terms of Use Agreement prohibits users from providing contact information such as email addresses and URLs on their profiles, many users still do so at their own peril. Consequently, spammers can use that information to spam those users using traditional techniques.

Spamming relies on the open nature of communications; thus, the vulnerabilities of MySpace’s communication facilities are proportionate to the openness of their access. On the conservative side, the commenting and bulletin systems are the most spam resistant because they can only be used by a user’s friends. The blogging, IM, and



**Granular Computing System Vulnerabilities: Exploring the Dark Side of Social Networking Communities, Figure 3**  
**Deceptive Zango popup license agreement**

friend request systems also provide the option of disallowing non-friends from using them to contact a user, but this protection is not enabled by default. Thus, these three systems are more susceptible to spamming because any user can use them to contact other users that have not enabled this protection. The messaging system is the most vulnerable to abuse because it does not provide an effective spam prevention mechanism. It allows users to report a message as being spam, but those users have no assurances that immediate action will be taken. Therefore, users may receive a number of spam messages before MySpace administrators eliminate the offending spammer from the system. MySpace also provides a general block listing mechanism that allows users to completely prevent all communication from specific users. However, this feature is ineffective due to the ease with which users can create new MySpace profiles (i.e., if spammers have been blocked, they can create new profiles and continue their spamming activities undeterred).

Spammers abuse these communication systems for the same reason they abuse traditional communication facilities: promotion. Spammers want to expose their products, web sites, and viewpoints to as many individuals as

possible, and spamming provides them with a technique to accomplish that goal. For social networking communities, this spamming activity represents a huge problem for a couple of reasons. First, spam content wastes a significant amount of resources, including storage space, bandwidth, and users' time. This last wasted resource leads us to the second major consequence of social network spam. When users waste time dealing with spam, it has a negative effect on their overall experience with these communities because it prevents them from participating in their desired activities. For example, when users are forced to sift through an inbox full of annoying spam messages, they are unable to send and receive messages. Similarly, when users are burdened with removing an endless stream of spam comments on their blogs, they are unable to post new content. Consequently, if users become overwhelmed with spam, they will have no incentive to continue interacting with these communities, and the communities will be forced to shut down.

In addition to spamming the social networking community, spammers are also able to use the information on user profiles to more effectively spam users outside the community. By mining email addresses, IM screen names,

and other contact information from these profiles, spammers can spam users using traditional techniques (e.g., email spam, spim, blog comment spam, etc.). Additionally, spammers can use the personal information on these profiles to construct user-specific spam content that is more relevant to the user, making it more likely to be read. For example, if a user's profile contains a great deal of sports-related information, a spammer can leverage this information to create a sports-oriented spam message for that user. Since the message's content matches the user's interests, the user is much more likely to read it, and as a result, the spam's sales pitch is more likely to be successful. To make their messages even more deceptive, spammers can also masquerade as users' friends. A user's profile contains a list of that user's friends; thus, spammers can use this information to make their messages appear as if they were sent by those friends. Since a user is much more likely to read and trust a message from a friend, the spammer has a higher probability of success with these disguised messages [7].

## Phishing

In a traditional phishing attack, a phisher seeks to obtain a targeted user's sensitive information through means of deception. Historically, phishers have been interested in credit card information, banking information, and login information for various web sites (e.g., eBay, PayPal, etc.). As social networking communities have become more popular, complex networks of friends have been established within them. Consequently, social networking communities have become a prime target for phishers because they are portals to a large number of potential victims.

A MySpace phishing attack utilizes the same deceptive techniques that are employed in traditional phishing attacks. First, a user is presented with a seemingly legitimate URL (called a *phishing URL*) that appears to be affiliated with MySpace. This phishing URL is typically propagated in two distinct manners: the communication systems provided by MySpace and traditional communication channels (e.g., email, blogs, instant messaging, etc.). For example, in May 2006, a phisher used the MySpace spamming techniques described above to send a phishing message to various MySpace users. This message had "CHECK OUT these old school pictures ..." as its subject and a phishing URL in its body [15]. Upon accessing one of these phishing URLs, the user is directed to a fraudulent web page that appears identical to the authentic MySpace login page. When the user enters the necessary MySpace login information, the fraudulent page stores that information and uses it to redirect the user to the authentic MySpace com-

munity. Thus, the user is completely unaware of the attack, while the phisher successfully obtains the user's sensitive login information.

Phishing attacks represent a serious threat to MySpace users for three reasons. First, victimized users often lose control of their profiles. In many cases, phishers will immediately change the login information of profiles they have compromised, locking victims out of their own profiles. Since users spend a great deal of time and energy building new friendships, writing blogs, and customizing their profiles, it is somewhat traumatic when they lose control of their creations. Even in cases where a compromised profile's login information remains the same, the phisher's activities severely damage that profile's credibility. For example, if a phisher compromises a user's profile and begins spamming that user's friends, those friends will eventually distrust the compromised profile and remove it from their lists of friends. The second reason these attacks are dangerous is due to the viral propagation patterns that are possible in social networking communities. As shown with the malware examples above, one compromised user can quickly escalate into a network full of compromised users. Specifically, once a single user's login information is compromised, the phisher can use that user's profile to propagate the attack to the user's friends. In our spam discussion, we mentioned a few MySpace communication systems that are somewhat spam resistant. However, the spam resistance of those systems assumes that a user's friends can be trusted (i.e., they are not spammers, phishers, etc.). Thus, if a user's profile becomes compromised, that user's friends are immediately at risk because the phisher can contact them under the guise of a profile they trust. As a result, a compromised user's friends are highly likely to become phishing victims (i.e., access the phishing URL), and by the time they realize they should distrust the compromised profile, it will be too late. The final threat posed by MySpace phishing attacks relies on the knowledge that many computer users reuse the same login information at many different sites [10]. Thus, if a user's MySpace login information is compromised, that user's login information at those other sites is automatically compromised as well.

In addition to launching phishing attacks that specifically target social networking communities, phishers can also use the private information found in these communities to make their traditional phishing attacks more effective. As previously mentioned, many MySpace users include various pieces of personal information on their profiles (e.g., location, interests, occupation, etc.). Phishers can easily use this personal information to construct user-specific messages that a potential victim would be much

more likely to read and trust. For example, a phisher could send the potential victim a fraudulent eBay email that contains auction information for products the victim would be interested in buying. Each of these products could be selected using the user's interests and associated with one or more phishing URLs. Since this email message contains user-specific content, the victim is more likely to read it and access one of its phishing URLs. As a result, this user-specific attack has a higher probability of success than a generic phishing attack.

### New Attacks Against Social Networking Communities

Initially, social networking communities were composed of ordinary people – typically kids and young adults that created profiles to interact with their friends. However, as these communities began to expand, new parties became interested because the communities evolved into portals for connecting and interacting with these ordinary people. The first wave of new participants were entertainers (e.g., musicians, comedians, and actors) that created profiles to communicate with their current fans and reach out to potential new ones. Not long after that, companies began partnering with communities to create advertising profiles (i.e., profiles that advertise the company and its products). Eventually, other public figures such as politicians became involved, creating profiles to promote their campaigns, solicit volunteers, and request donations. Unfortunately, the explosive growth of these communities and the conflicting interests of their various participants have generated a range of new attacks. In this section, we describe attacks that utilize two types of malicious social networking profiles: rogue advertising profiles and impersonating profiles.

#### Rogue Advertising Profiles

According to analysts at eMarketer, companies currently spend around \$280 million on social network advertising in the US, and by 2010, that figure is expected to grow to \$1.9 billion [8]. Thus, it is not surprising that MySpace has partnerships with numerous companies, which allow those companies to advertise legitimately in the social networking community. For example, both Burger King and Wendy's have embraced the potential of social network advertising by creating advertising profiles on MySpace. Burger King uses the King, a marketing character that appears in the company's commercials, to promote its MySpace profile at <http://www.myspace.com/burgerking>. Similarly, Wendy's uses a small hamburger patty named Smart to promote its MySpace profile at <http://www.myspace.com/wendysquare>.

[myspace.com/wendysquare](http://www.myspace.com/wendysquare). However, this new advertising craze is not exclusive to fast food restaurants. MySpace profiles also exist for television shows (e.g., "It's Always Sunny in Philadelphia"), commercial products (e.g., Herbal Essences shampoo), and movies (e.g., "Talladega Nights").

Although MySpace's Terms of Use Agreement strictly prohibits commercial use of the community without prior approval, a number of companies violate this policy by creating advertising profiles without MySpace's consent. We refer to these profiles as *rogue advertising profiles*, and they appear in many forms with varying degrees of deception. The most innocuous group of rogue advertising profiles is created by small web site operators that are merely trying to generate web traffic for their sites. These profiles are relatively harmless because they simply include URLs for these small sites, and they are openly affiliated with a company (i.e., they do not attempt to deceive users into believing the profiles belong to an ordinary individual). However, these profiles are still unacceptable because they pollute the community with unwanted and unauthorized advertising.

A much more deceptive group of rogue advertising profiles is generated by nefarious companies such as gambling and pornographic web sites. Unlike legitimate advertising profiles, which are clearly identifiable as a marketing device, these deceptive rogue advertising profiles appear as though they are maintained by an individual (and not a company). As a result, naïve users can easily mistake these rogue advertising profiles as ordinary MySpace profiles that were created by ordinary MySpace users. The deceptive profile construction process proceeds as follows. First, a fraudulent MySpace profile is created using a picture of an attractive female. This profile also contains a provocative description in its "About me" section that visitors assume was written by the pictured female. Conveniently, this description usually includes at least one reference to the URL of a nefarious web site. If the profile is particularly deceptive, the description includes an instant messenger screen name instead of a URL. The inclusion of a screen name is especially manipulative because most users assume that a conversation over IM can only be accomplished by a real person. However, the screen names found on these rogue advertising profiles are almost always attached to an IM bot (i.e., a computer program that emulates a real conversation) that attempts to direct its victims to the URL of a nefarious web site. Once the profile is completed, its creator sends friend requests to males near a specific geographic location, which the profile is also associated with (i.e., the pictured female claims to live there). When the males receive these requests, a num-

**luvmonica**

Female  
20 years old  
LONG BEACH, California  
United States

Last Login: 8/4/2006

[View My Pics | Videos](#)

Contacting luvmonica

- [Send Message](#)
- [Forward to Friend](#)
- [Add to Friends](#)
- [Add to Favorites](#)
- [Instant Message](#)
- [Block User](#)
- [Add to Group](#)
- [Rank User](#)

**luvmonica's Details**

Status:	Single
Zodiac Sign:	Taurus
Children:	I don't want kids

luvmonica is in your extended network

**luvmonica's Latest Blog Entry [Subscribe to this Blog]**  
[\[View All Blog Entries\]](#)

**luvmonica's Blurb:**

About me:  
M.S.N - GLOSSYBUNZPHOTMAIL.COM - ADD ME IM USUALLY ALWAYS ON AND I HAVE A WEBCAM (DISREGARD THE MAN IN MY PIC, IT'S OLD) - I JUST MOVED TO A NEW TOWN AND IM TRYING TO MEET SOME NEW PEOPLE OR SOMETHING. SO I'M GOING TO MEET SOMEONE ELSE. I LOVE MEN MY AIM IS SMILEYDIMA BUT I PREFER MSN PLEASE... I CANNOT GO ON WEBCAM ON... I may not be perfect but some parts of me are very good. I like to go to the beach, go to the mall, eat pizza, eat PEAS FAV FOOD - ITALIAN FAV PLACE - HAWAII FAV CLOTHING LINE -

**Censored**

a

**flirtymonica**

Female  
20 years old  
SAN JOSE, California  
United States

Last Login: 8/5/2006

[View My Pics | Videos](#)

Contacting flirtymonica

- [Send Message](#)
- [Forward to Friend](#)
- [Add to Friends](#)
- [Add to Favorites](#)
- [Instant Message](#)
- [Block User](#)
- [Add to Group](#)
- [Rank User](#)

**flirtymonica's Details**

Status:	Single
Zodiac Sign:	Virgo
Children:	I don't want kids

flirtymonica is in your extended network

**flirtymonica's Latest Blog Entry [Subscribe to this Blog]**  
[\[View All Blog Entries\]](#)

**flirtymonica's Blurb:**

About me:  
MS.N - GLOSSYBUNZPHOTMAIL.COM - ADD ME IM USUALLY ALWAYS ON AND I HAVE A WEBCAM (DISREGARD THE MAN IN MY PIC, IT'S OLD) - I JUST MOVED TO A NEW TOWN AND IM TRYING TO MEET SOME NEW PEOPLE OR SOMETHING. SO I'M GOING TO MEET SOMEONE ELSE. I LOVE MEN MY AIM IS SMILEYDIMA BUT I PREFER MSN PLEASE... I CANNOT GO ON WEBCAM ON... I may not be perfect but some parts of me are very good. I like to go to the beach, go to the mall, eat pizza, eat PEAS FAV FOOD - ITALIAN FAV PLACE - HAWAII FAV CLOTHING LINE -

**Censored**

b

**kinkymonica**

Female  
20 years old  
SAN JOSE, California  
United States

Last Login: 8/13/2006

[View My Pics | Videos](#)

Contacting kinkymonica

- [Send Message](#)
- [Forward to Friend](#)
- [Add to Friends](#)
- [Add to Favorites](#)
- [Instant Message](#)
- [Block User](#)
- [Add to Group](#)
- [Rank User](#)

**kinkymonica's Details**

Status:	Single
Zodiac Sign:	Libra
Children:	I don't want kids

kinkymonica is in your extended network

**kinkymonica's Latest Blog Entry [Subscribe to this Blog]**  
[\[View All Blog Entries\]](#)

**kinkymonica's Blurb:**

About me:  
M.S.N = KISSYFLIRTPHOTMAIL.COM - IF YOU DONT HAVE MSN - then add me on aim at KINKYSHAKES but I REALLY PREFER MSN as I cannot go on aim on my laptop. I am on IM now though..... I'm the type of girl you'll meet once in a lifetime. If not then never, or not self centered. I like going to the beach. I really like trying new things and seeing new places. I dont like being ppl & im not into drama either. I like guys who are down to earth, intelligent, confident and have sexy laughs. I like to hang out with my close ones, go all out and being in the moment. Were all living a short life so might as well represent who you are. I Love skating and I like my dogs. I also like to go to the beach. I like sometimes shots of Vodka. Straight I'm also athletic and enjoy to bike

**Censored**

c

**kinkymonica**

Female  
20 years old  
OKLAHOMA CITY, Oklahoma  
United States

Last Login: 8/13/2006

[View My Pics | Videos](#)

Contacting kinkymonica

- [Send Message](#)
- [Forward to Friend](#)
- [Add to Friends](#)
- [Add to Favorites](#)
- [Instant Message](#)
- [Block User](#)
- [Add to Group](#)
- [Rank User](#)

**kinkymonica's Details**

Status:	Single
Zodiac Sign:	Sagittarius
Children:	I don't want kids

kinkymonica is in your extended network

**kinkymonica's Latest Blog Entry [Subscribe to this Blog]**  
[\[View All Blog Entries\]](#)

**kinkymonica's Blurb:**

About me:  
MS.N = KISSYFLIRTPHOTMAIL.COM - IF YOU DONT HAVE MSN - then add me on aim at KINKYSHAKES but I REALLY PREFER MSN as I cannot go on aim on my laptop. I am on IM now though..... I'm the type of girl you'll meet once in a lifetime. If not then never, or not self centered. I like going to the beach. I really like trying new things and seeing new places. I dont like being ppl & im not into drama either. I like guys who are down to earth, intelligent, confident and have sexy laughs. I like to hang out with my close ones, go all out and being in the moment. Were all living a short life so might as well represent who you are. I Love skating and I like my dogs. I also like to go to the beach. I like sometimes shots of Vodka. Straight I'm also athletic and enjoy to bike

**Censored**

d

**Granular Computing System Vulnerabilities: Exploring the Dark Side of Social Networking Communities, Figure 4**  
Pornographic rogue advertising profiles

ber of them accept and visit the profile, making the deception a success.

Figure 4 shows four examples of deceptive rogue advertising profiles, which we slightly modified to hide objectionable content. By analyzing these examples, a number of interesting observations emerge. First, the profiles are extremely similar. Figures 4a and 4b share the exact same “About me” text, and Figs. 4c and 4d share the same “About me” text and profile picture. Thus, these two pairs of profiles were probably created by the same individuals. The next observation addresses the content of each profile’s “About me” text. All four of these profiles are particularly deceptive because they all provide instant messenger screen names to induce users to contact IM bots. Finally, all of the profiles claim to be located in different cities. Despite their uncanny similarities, the profiles in Figs. 4a and 4b are supposedly located in different parts of California, and the profiles in Figs. 4c and 4d are in completely different parts of the country. These location differences exist because each profile is meant to target males in

a different geographic area. This also explains why the profiles are allowed to be so similar. As long as the profiles in a given location are unique, users are less likely to become suspicious that the profiles are fraudulent.

Deceptive rogue advertising profiles represent a serious security threat because they directly manipulate the behavior of users. Since the profiles appear to be maintained by ordinary (albeit very attractive) people, other users mistakenly trust their content. Additionally, users often believe they have made a connection with the people pictured on these profiles because the users received friend requests from these profiles. Consequently, users access the URLs that are on the profiles (or in IM conversations) and become victims to the content found on the corresponding web pages. In most cases, these pages contain pornographic or gambling-related material, but nothing prevents malicious individuals from embedding malware on the pages. Thus, to protect its users, social networking communities must develop techniques to identify and eliminate these profiles.

**Rupert Murdoch**  
  
 "The Dirty Digger"  
 Male  
 75 years old  
 New York, New York  
 United States  
[View My Pics | Videos](#)  
**Contacting Rupert Murdoch**  
  
**MySpace URL:** [http://www.myspace.com/murdoch\\_rupert](http://www.myspace.com/murdoch_rupert)  
  
**Rupert Murdoch's Interests**  

General	Global domination and the conjunction of church and state. Censorship of private or liberal interest groups. Censorship of MySpace pages that do not agree with my views. Dictatorships are fun... as long as I'm in charge.
Music	Very monotonous music to my ears.
Movies	Fox Movies
Television	20th Century Fox 20th Century Fox Espanol 20th Century Fox Home Entertainment 20th Century Fox International 20th Century Fox Television Blue Sky Studios Fox Animation Pictures Fox Studios Australia Fox Studios Italia Fox Television Studio Fox Broadcasting Company Fox Sports Australia Fox Television Stations FOXTEL STAR FOX Movie Channel Fox News Channel Fox College Sports Fox Sports Fox Sports Australia Fox Sports Espanol Fox Sports Net

a

**everywhere**  
**Rupert Murdoch is in your extended network**

**Rupert Murdoch's Latest Blog Entry [Subscribe to this Blog]**  
 The Typical E-Mails I Get.... and my Typical Response: [\(view more\)](#)  
 The things I do... [\(view more\)](#)  
 I propose Internment Camps in Britain [\(view more\)](#)  
 I own, Fox News... is 'Outfoxed' right or wrong? [\(view more\)](#)  
 How do you feel about me buying MySpace? [\(view more\)](#)  
[\[View All Blog Entries\]](#)

**Rupert Murdoch's Blurb:**  
**About me:**  
*Roger Ten Years of Propaganda, Fox News!*  
 ... width="425" height="250">



I just bought MySpace.com, soon I will own the rest of the internet.  
 I aspire to become the most powerful man in the world. I promise the effects and connotation between right wing conservatism and government. I encourage my television and radio stations to become Republican soapboxes.

There are many important benefits to fascism.

"I'm a catalyst for change. You can't be an outsider and be successful over 30 years without leaving a certain amount of scar tissue around the place."

"The world is changing very fast. Big will not beat small anymore. It will be the fast beating the slow."

"In motivating people, you've got to engage their minds and their hearts. I motivate people. I hope, by example - and perhaps by excitement, by having productive ideas to make others feel involved."

"For better or for worse, our company (The News Corporation Ltd.) is a reflection of my thinking, my character, my values."

**Rupert Murdoch**  
  
 "World's Most Powerful Man"  
 Male  
 75 years old  
 Melbourne, Victoria  
 Australia  
 Last Login: 10/10/2006  
[View My Pics | Videos](#)

**Contacting Rupert Murdoch**  


**MySpace URL:** <http://www.myspace.com/lordrupertmurdoch>

**Rupert Murdoch's Interests**

General	money world domination Aussie Rules Football Cricket film opera politics hunting crocodiles
Television	The O'Reilly Factor The Apprentice Laguna Beach Golden Girls
Heroes	Moses, money and Mussolini
Groups:	Napoleon Dynamite, Koala Bears Are Just Like Us, The World's Hottest, You're Now Owned By FOX and RUPERT MURDOCH'S UNITED, MySpace Republicans

[View All Rupert Murdoch's Groups](#)

**Rupert Murdoch's Details**

Status:	Married
Hobbies:	Friends
Orientation:	Straight
Hometown:	Melbourne, Australia
Body type:	5' 10" / Average

[View All Rupert Murdoch's Details](#)

**Rupert Murdoch is in your extended network**

**Rupert Murdoch's Latest Blog Entry [Subscribe to this Blog]**  
[\[View All Blog Entries\]](#)

**Rupert Murdoch's Blurb:**  
**About me:**  
 YES, I AM THE SAME RUPERT MURDOCH THAT OWNS MYSPACE.COM AND FOX

**DISCLAIMER:**  
 Unlike the other Rupert Murdochs on this site, I am THE REAL one. Those other impostors will be hearing from my lawyers soon!!!

I might have been born in 1931, but I'm the biggest bloke around. It's true, I own MySpace.com. Therefore, I own you.

My acquisition of this sophisticated web network of friends is just another step in my quest to dominate the world. **AUSTRALIA WINNSHIN** with my ownership of the most powerful broadcasting corporations in the world (such as Fox News), and now with myspace.com in my hands, I am ready to impose Australian supremacy upon the world that so long ignored and oppressed our lovely kangaroos inhabiting continent.

Yes, I'm Aussie. Born and raised in Melbourne to Sir Kitch Murdoch and my mother. Her name? It's not important! Oh yes, Dame Elisabeth Murdoch! She was a good mother! Yes, yes!

I am also the majority shareholder and managing director of News Corporation, the world's largest and most influential media corporation. So there you go. I have lots of money and I can go on many vacations. Neener Neener Neener!

**THINGS THAT THIS WORLD'S MOST POWERFUL MAN ENJOYS:**  
 vegemite  
 koalas  
 aussie family - quite lovely people  
 capital punishment  
 surfing  
 surfing  
 surfing on the beach  
 PIRINGU'S WORLD!  
 Asian women  
 Freedom and democracy to people's of Iraq!  
 China  
 get together's at bill's (O'Reilly's)  
 myspace.com  
 family is quite important to me...because they'll take over once my time expire!

**KURIOUS THING THAT ANNOYS ME:**  
 The EU  
 The USA  
 Chris Patten!  
 Pro choice  
 Women's rights in general  
 American's mad obsession with Aussie accents...I always get asked to say something in "Australian". Get it right, we all speak English.

**Yay!**  
**What I'd like to meet:**  
 Jesus  
 Moses  
 God  
 and ask them for more money  
 Ashton Simpson  
 Angelina Jolie  
 Brad Pitt

**Granular Computing System Vulnerabilities: Exploring the Dark Side of Social Networking Communities, Figure 5**  
 Example impersonating profiles for Rupert Murdoch

## Impersonating Profiles

Due to the enormous popularity of social networking communities, they have emerged as forums for individuals to voice their opinions about various topics as well as other people. In MySpace, some users are more tactful and express their views in blog postings on their profiles or in comments on their friends' profiles. However, other users have taken their crusade to an entirely new level, creating new MySpace profiles that directly target specific ideals, companies, and even people. We refer to these profiles as *impersonating profiles* because they impersonate their targets to convey their message.

MySpace contains profiles that impersonate actors (e.g., Tom Hanks), athletes (e.g., Michael Jordan), technologists (e.g., Bill Gates), and politicians (e.g., George W. Bush). Even ancient philosophers, such as Socrates and Aristotle, have profiles dedicated to them. In fact, most of these individuals are impersonated by multiple distinct profiles. In some cases, these profiles are meant to be an homage to the individuals in question. However, more often than not, the impersonating profiles are meant to be slanderous to the targets of the impersonation. For example, after News Corporation purchased MySpace in July 2005, impersonating profiles for Rupert Murdoch (News Corporation's CEO) began to appear, making claims such as, "I just bought MySpace.com, soon I will own the rest of the internet" and "Dictatorships are fun ... as long as I'm in charge." Two examples of these profiles are shown in Figs. 5a and 5b. To make matters worse, impersonators are able to use MySpace's communication facilities to contact other users under the guise of these profiles. Thus, these counterfeit Rupert Murdoch profiles can express any number of scandalous opinions, and the negative backlash will be directed at the real Rupert Murdoch.

From the impersonators' standpoint, these profiles appear to be amusing satire. However, from the victims' standpoint, these profiles represent a form of identity theft as well as a public relations nightmare. The general public is unable to verify the authenticity of these profiles; thus, any slanderous comments found on the profiles (or received under the guise of the profiles) will be incorrectly

attributed to the real victim. This is a serious concern for social networking systems, as it can lead to a loss of trust in the system and its ability to protect user privacy and security. It is important for social networking systems to implement strong authentication and verification mechanisms to prevent such impersonating profiles from appearing on the platform.

associated with the victims of the impersonation. Using the example above, nothing prevents users from believing that Rupert Murdoch actually created one of those profiles or that he actually shares the viewpoints it contains. Admittedly, these viewpoints are somewhat absurd, but it is easy to envision subtler comments that can severely damage an individual's reputation. Since the success of most public figures is completely contingent upon the strength of their fan bases, an impersonating profile represents a direct threat to its victim's credibility and livelihood.

Sadly, this growing epidemic is not limited to high profile individuals. Private citizens such as teachers, principals, and police officers have also been victimized by these impersonating profiles. In fact, over the past couple of years, the news has been filled with stories about these types of attacks. For example, in December 2005, a 16-year-old boy posted an impersonating profile of a local police officer, which contained various derogatory statements about the police officer's appearance, intelligence, and sexual orientation [13]. Then, in April 2006, an eighth grader created impersonating profiles of his English teacher that contained racist remarks and falsely represented the teacher as a pornographer and child molester [12]. In September 2006, a high school assistant principal sued two students for creating an impersonating profile that falsely identified her as a lesbian [1]. These examples clearly illustrate the magnitude of this problem, and unfortunately, they represent a small sample of a growing list of incidents.

In addition to the risks already mentioned, impersonating profiles can also be used to amplify the severity of the other attacks we have already discussed. Specifically, a spammer could create an impersonating profile for a MySpace user and use the profile to spam that user's friends. Malware creators and phishers could also use this approach to spread their malicious content to the user's friends. Since these friends would be under the impression that they were communicating with the authentic profile, they would be much more likely to trust the communication, and as a result, the attacks would be far more successful.

## Future Directions

In only a few years, social networking communities have made staggering strides in popularity (MySpace welcomed more than 65 million unique visitors in February 2008 [3]) and importance (YouTube was acquired by Google for \$1.65 billion in October 2006 [5]). Aside from their social and economic impact, one of the most important ques-

tions is whether social networking environments are safe for their users. In this paper, we have analyzed this safety question and described several security and privacy threats that have translated into real attacks in MySpace. Some of these attacks have been adapted from other Internet environments, including variants of malware propagation, spam, and phishing. Other attacks are new and unique to social networking environments, including the creation of rogue advertising profiles and impersonating profiles. From our real world observations, it is clear that additional efforts should be made to protect social networking users. We hope that researchers will identify and counteract emerging variations of these known attacks and provide measures for safeguarding the future of these valuable communities.

## Bibliography

- Associated Press (2006) Official Sues Students Over MySpace Page. <http://www.sfgate.com/cgi-bin/article.cgi?file=/news/archive/2006/09/22/national/a092749D95.DTL>
- Boyd C (2006) Teenagers used to push Zango on Myspace. <http://www.vitalsecurity.org/2006/07/teenagers-used-to-push-zango-on.html>
- Freiert M (2008) February Top Social Networks - Make way for the new guys. <http://blog.compete.com/2008/03/07/top-social-networks-traffic-feb-2008>
- Geer D (2005) Malicious bots threaten network security. IEEE Computer 38(1):18–20
- Google Press Center (2006) Google To Acquire YouTube for \$1.65 Billion in Stock. [http://www.google.com/press/presrel/google\\_youtube.html](http://www.google.com/press/presrel/google_youtube.html)
- Hitwise (2008) Hitwise US - Top 20 Websites - February, 2008. <http://www.hitwise.com/datacenter/rankings.php>
- Jagatic T et al (2007) Social Phishing. Communications of the ACM 50(10):94–100
- King R (2006) Marketing to Kids Where They Live. [http://www.businessweek.com/technology/content/sep2006/tc20060908\\_974400.htm?campaign\\_id=bier\\_tcs.g3a.091106a](http://www.businessweek.com/technology/content/sep2006/tc20060908_974400.htm?campaign_id=bier_tcs.g3a.091106a)
- Krebs B (2006) Hacked Ad Seen on MySpace Served Spyware to a Million. [http://blog.washingtonpost.com/securityfix/2006/07/mySPACE\\_ad\\_served\\_adware\\_to\\_mo.html](http://blog.washingtonpost.com/securityfix/2006/07/mySPACE_ad_served_adware_to_mo.html)
- Ross B et al (2005) Stronger password authentication using browser extensions. In: Proceedings of the 14th Usenix Security Symposium, Baltimore, 31 Jul – 5 Aug 2005, pp 17–32
- Samy (2005) Technical explanation of the MySpace worm. <http://namb.la/popular/tech.html>
- Sanchez M (2006) Pranksters posting fake profiles on MySpace. <http://www.dfw.com/mld/dfw/news/local/15255785.htm?template=contentModules/printstory.jsp>
- Seibel J (2006) Boy charged in creating fake Myspace profile. <http://www.jsonline.com/story/index.aspx?id=413620>
- Swartz J (2008) Social-networking sites going global. [http://www.usatoday.com/money/industries/technology/2008-02-10-social-networking-global\\_N.htm](http://www.usatoday.com/money/industries/technology/2008-02-10-social-networking-global_N.htm)
- Technocrat (2006) Myspace Phishing Attacks on the Rise. <http://djtechnocrat.blogspot.com/2006/05/myspace-phishing-attacks-on-rise.html>

## Granular Flows

JAMES W. DUFY

Department of Physics, University of Florida,  
Gainesville, USA

### Article Outline

- [Glossary](#)
- [Definition of the Subject](#)
- [Introduction](#)
- [Granular Fluid and Its Statistical Mechanics](#)
- [Macroscopic Balance Equations](#)
- [“Normal” States and Hydrodynamics](#)
- [Navier–Stokes Approximation](#)
- [Future Directions](#)
- [Acknowledgments](#)
- [Appendix](#)
- [Bibliography](#)

### Glossary

**Granular matter** a system comprised of a large number of grains, or particles, of macroscopic size. Examples include powders, sand, seeds, and the surface of Mars.

**Granular fluid** an activated (driven) state of granular matter such that the grains move (flow) and collide frequently.

**Statistical mechanics** a field of physics that addresses systems with many degrees of freedom based on the fundamental microscopic laws to describe derived macroscopic properties.

**Macrostate** a statistical description of a system with many degrees of freedom in terms of limited information about that system.

**Hydrodynamic fields** the local densities of mass, energy, and momentum defined at each point in the system of interest.

**Normal state** a macrostate whose time evolution is described entirely through that of the average hydrodynamic fields.

**Balance equations** exact equations for the time derivatives of the hydrodynamic fields in terms of associated fluxes and sources.

**Constitutive equations** expressions for the fluxes and sources of the balance equations as functionals of the hydrodynamic fields.

**Hydrodynamics** a macroscopic description of the system in terms of a closed, deterministic set of equations for the average hydrodynamic fields, resulting from the exact balance equations with approximate constitutive equations.

**Navier–Stokes hydrodynamics** local, first order in time, partial differential equations for states with small spatial gradients in the hydrodynamic fields (constitutive equations calculated to first order in the spatial gradients).

### Definition of the Subject

The terminology granular matter refers to systems with a large number of hard objects (grains) of mesoscopic size ranging from millimeters to meters. Geological examples include desert sand and the rocks of a landslide. But the scope of such systems is much broader, including powders and snow, edible products such as seeds and salt, medical products like pills, and extraterrestrial systems such as the surface regolith of Mars and the rings of Saturn. The importance of a fundamental understanding for granular matter properties can hardly be overestimated. Practical issues of current concern range from disaster mitigation of avalanches and explosions of grain silos to immense economic consequences within the pharmaceutical industry. In addition, they are of academic and conceptual importance as well as examples of systems far from equilibrium.

Under many conditions of interest, granular matter flows like a normal fluid [1]. In the latter case such flows are accurately described by the equations of hydrodynamics. Attention is focused here on the possibility for a corresponding hydrodynamic description of granular flows. The tools of nonequilibrium statistical mechanics [3], developed over the past fifty years for fluids composed of atoms and molecules [4,5], are applied here to a system of grains for a fundamental approach to both qualitative questions and practical quantitative predictions. Applications of basic atomic physics principles to granular fluids have accelerated during the past decade, starting with an emphasis on molecular dynamics (MD) simulations [6] and kinetic theory [7,8], and more recently with the theoretical methods of the type described here [9,10,11,12,13,14].

### Introduction

To start with the familiar, consider a jar of vitamin pills, mustard seeds, or peanuts. Remove the lid and pour them into a bowl, observing that the “flow”, or their collective motion, has some similarity to that of a normal fluid such as water. The collective motion in both cases is the consequence of collisions among their constituents, grains or atoms, and their large number. It is tempting to make the correspondence of grains to atoms in considering the similarities of flows in these two types of fluids. The objective here is to explore in formal detail the extent to

which that correspondence is conceptually and quantitatively justified. An important prerequisite is the integrity of the grains during their motion. Each grain is comprised of a large number of atoms or molecules. Integrity refers to their retention of mass and shape following interactions with other grains or with their environment. As such, the grains behave as “particles” whose detailed internal structure is not essential to their description, which is captured instead by a few parameters describing their shape, mass, and collisional properties with other grains. However, an important consequence of this underlying molecular structure is a redistribution of translational kinetic energy of the grains and internal energy of the constituent molecules. At the mesoscopic level this appears as an energy loss on collisions between pairs of grains. This is a central feature of granular fluids differentiating them from atomic fluids: the inelasticity of granular pair collisions.

Granular matter occurs in two classes of states, compact and activated [15]. In the first case, the grains form a static packed configuration within the container due to the effects of gravity on their relatively large mass and their inelastic collisions. Any initial motion is quickly dissipated and their kinetic energy becomes negligible relative to the gravitational potential energy. Important questions arise about the possible and probable packing configurations that determine the stresses within the system and the distribution of forces on the container. For example, chains of particles in contact can occur as arches to support matter above them while reducing their force on the matter below. There is an intense interest in the study of such states, generically referred to as contact mechanics.

Activated states refer to continuously driven systems, or gravity free conditions. For example, a container of grains in a compact configuration can be shaken to impose kinetic energy and motion among the grains. Similarly, unrestrained systems in a gravitational field will flow towards lower potential energy (e.g., hopper flow or flow down an incline). Initial activation in space laboratory experiments provides another example (self-sustained fluidization). For the flows considered here as candidates for a hydrodynamic description continual collisions are essential. This means that within each small cell, still containing many particles, the particles are moving randomly relative to the collective motion of that cell. Thus, ballistic motion or beams with all particles moving independently in the same direction are excluded.

Both compact and activated grains may occur immersed in a continuum such as water or air that may have a strong or weak effect on their collective properties. For compact systems water may provide a lubrication effect

that affects the dominant class of configurations. For activated systems it can provide an additional dissipative drag between collisions among the grains. When the medium plays an important role the systems is said to be wet. In the opposite limit it is said to be dry. Finally, it is possible for compact and activated components of a system to coexist as heterogeneous states. Here, only the simplest case of dry systems in fully activated flows are considered. These are referred to in the following as granular fluids.

Advances in the study of granular fluids have arisen from many communities, including chemical engineering, materials sciences, and physics. The additional academic and conceptual importance of granular matter as practical systems for exploring the relevance of many-body fluid methods is primarily for the physics community. Granular matter, viewed as a system of particles with inelastic interactions, provides new opportunities to test the qualitative and quantitative limits of many-body methods developed over the past century for atomic and molecular systems. This is the field of non-equilibrium statistical mechanics [3,4,5]. Granular matter provides a new testing ground for a reconsideration of the most fundamental concepts and tools [1], with the potential for enhanced understanding of their place in atomic and molecular systems as well.

Statistical mechanics addresses the difficult many-body problem of extracting macroscopic properties of experimental interest from the very large number of constituent particles. The results express these properties in terms of the fundamental “microscopic” features of these particles, such as mass, shape, degree of inelasticity, and collisional properties. In this way the properties of the vitamin pills, mustard seeds, and peanuts are distinguished at a fundamental level. Also, conceptual issues such as the limitations of a macroscopic description are exposed through the insistence on their logical evolution from the fundamental microdynamics. In the next section, the granular fluid is described as a system of particles interacting via pairwise additive, nonconservative forces. The microscopic dynamics of these particles leads to balance equations for the mass density, energy density, and momentum density. Their averages define the “hydrodynamic fields” which are candidates for a macroscopic, continuum mechanics description. These exact equations are described in Sect. “[Macroscopic Balance Equations](#)” and the need for “constitutive equations” to provide a closure is described. The origin of constitutive equations, and consequently the origin of hydrodynamics, is associated with the concept of “normal states” [16] in Sect. “[Normal States and Hydrodynamics](#)”. The normal state for the case of small spatial deviations from homogeneity is constructed

formally in Sect. “[Navier–Stokes Approximation](#)”, resulting in the constitutive equations for Navier–Stokes hydrodynamics [17]. This derivation also provides insight into the context in which such a description should hold, and differences from the corresponding results for a normal fluid are noted. Empirical evidence [18,19], simulations [20,21,22], and corresponding results from kinetic theory [23,24,25,26] support the applicability of this hydrodynamic description under appropriate conditions. Finally, the results are summarized in Sect. “[Future Directions](#)” and some comments on the outlook for future developments are offered.

The presentation here is focused on recent work of the author and his collaborators for application of statistical mechanics to explore hydrodynamics for a granular gas. Consequently, the references quoted are heavily weighted toward those developmental studies. Apologies are offered for the exclusion of the vast and important complementary literature on simulations, kinetic theory, and experiments also bearing on this topic. Many of these can be found in the list of Books and Reviews given here.

## Granular Fluid and Its Statistical Mechanics

### Nonequilibrium Statistical Mechanics

Consider a system of  $N \gg 1$  identical grains (hereafter referred to as particles) in a volume  $V$ , whose initial positions  $\{\mathbf{q}_i\}$  and velocities  $\{\mathbf{v}_i\}$ ,  $1 \leq i \leq N$ , are specified. The positions and velocities define a point in a  $6N$  dimensional space denoted by  $\Gamma = \{\mathbf{q}_i, \mathbf{v}_i\}$ , defining the microstate of the system. A macrostate is defined by a probability density  $\rho(\Gamma)$  in this space, representing statistical rather than precise knowledge of the system. The field of statistical mechanics addresses properties of macrostates, based on the recognition that for very large  $N$  the details of microstates are neither experimentally accessible nor practically calculable. Properties of interest are represented by functions  $A(\Gamma)$ , and their values for a macrostate  $\rho(\Gamma)$  are determined from the expectations

$$\langle A; \rho \rangle \equiv \int d\Gamma \rho(\Gamma) A(\Gamma). \quad (1)$$

In this section, a brief overview of the essential ingredients of nonequilibrium statistical mechanics is given, broadened from its usual form [3] to include granular matter.

The dynamics of macrostates is determined from the underlying dynamics of the microstates. The initial point  $\Gamma$  changes in time since the particles have velocities and move to new positions. They move in straight lines until one or more come within the force field of other particles, at which point their velocities change as well as their

positions. The forces are taken to be pairwise additive, such that the total force on particle  $i$  is  $\mathbf{F}_i = \sum_j \mathbf{F}_{ij}$ , where  $\mathbf{F}_{ij}$  is the force on particle  $i$  due to particle  $j$ . This does not mean that the interactions are pairwise sequential; three or more particles can interact simultaneously. The pair forces are restricted by Newton’s third law,  $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$ , with conservation of momentum. Otherwise quite general forces can be considered to represent the shape of the particles and their degree of inelasticity. It is assumed here that the force range vanishes outside a distance  $\sigma/2$  from the center of each particle so that  $\sigma$  characterizes the size of the particles. Furthermore, the particles are taken to be strongly repulsive so that their mean maximum overlap  $d$  on collision is small compared to their size,  $d/\sigma < 1$ . However, their size can be large or small compared to the mean distance between particles  $(V/N)^{1/3}$ , depending on whether the density of the system is small or large, respectively. Most importantly for the purposes here, these forces do not conserve energy. This property captures the feature of real grains that center of mass kinetic energy is lost as they distort during pair collisions. Further details of the force law are not required at this point.

The dynamics consists of straight line motion along the direction of the velocity at time  $t$  (free streaming), until the force range of any pair of particles, say  $i, j$ , overlaps. The relative velocity  $\mathbf{g}_{ij} = \mathbf{v}_i - \mathbf{v}_j$  of that pair changes according to Newton’s second law for the chosen force law  $\mathbf{F}_{ij}$ . Subsequently, all particles continue to stream freely until another pair has a force range of overlap, and the collisional change is repeated for that pair. In this way a trajectory  $\Gamma_t \equiv \{\mathbf{q}_1(t), \dots, \mathbf{q}_N(t), \mathbf{v}_1(t), \dots, \mathbf{v}_N(t)\}$  is generated for  $t > 0$ . This trajectory is unique and invertible. The statistical mechanics for a fluid of inelastic particles [9,10,11,12] is comprised of the dynamics just described, a macrostate specified in terms of a probability density  $\rho(\Gamma)$ , and a set of observables generically denoted by  $A(\Gamma)$ . The expectation value for an observable at time  $t > 0$  for a state  $\rho(\Gamma)$  given at  $t = 0$  is defined by

$$\langle A(t); 0 \rangle \equiv \int d\Gamma \rho(\Gamma) A(\Gamma_t) \equiv \int d\Gamma \rho(\Gamma) e^{tL} A(\Gamma) \quad (2)$$

where  $A(t) = A(\Gamma_t)$ , and  $\Gamma_t \equiv \{\mathbf{q}_1(t), \dots, \mathbf{q}_N(t), \mathbf{v}_1(t), \dots, \mathbf{v}_N(t)\}$  is the phase point evolved to time  $t$  from  $\Gamma = \Gamma_{t=0}$ . The dynamics can be represented in terms of a generator  $L$  defined by the second equality of (2). There are two components to the generator, corresponding to the two steps of free streaming and velocity changes due to interactions

$$L = \sum_{i=1}^N \mathbf{v}_i \cdot \nabla_i + \frac{1}{2m} \sum_{i=1}^N \sum_{j \neq i}^N \mathbf{F}_{ij} \cdot (\nabla_{\mathbf{v}_i} - \nabla_{\mathbf{v}_j}). \quad (3)$$

An alternative equivalent representation of the dynamics is obtained by transferring the dynamics from the observable  $A(\Gamma)$  to the state  $\rho(\Gamma)$  by the definition

$$\begin{aligned} \int d\Gamma \rho(\Gamma) e^{tL} A(\Gamma) &\equiv \int d\Gamma \left( e^{-t\bar{L}} \rho(\Gamma) \right) A(\Gamma) \\ &\equiv \int d\Gamma \rho(\Gamma, t) A(\Gamma). \end{aligned} \quad (4)$$

The representation in terms of a dynamical state  $\rho(\Gamma, t)$  is referred to as Liouville dynamics. Its generator  $\bar{L}$  is the formal adjoint of  $L$  which is found to be

$$\bar{L} = L + \frac{1}{2m} \sum_{i=1}^N \sum_{j \neq i}^N (\nabla_{v_i} - \nabla_{v_j}) \cdot \mathbf{F}_{ij}. \quad (5)$$

The difference between  $L$  and  $\bar{L}$  arises because the forces are non-conservative and therefore depend on the relative velocities of each pair as well as their positions. Time correlation functions for two observables  $A$  and  $B$  are defined in a similar way

$$\begin{aligned} \langle A(t)B; 0 \rangle &\equiv \int d\Gamma (e^{tL} A(\Gamma)) \rho(\Gamma) B(\Gamma) \\ &= \int d\Gamma A(\Gamma) \left( e^{-t\bar{L}} \rho(\Gamma) \right) (e^{-tL} B(\Gamma)). \end{aligned} \quad (6)$$

or

$$\langle A(t)B; 0 \rangle \equiv \langle AB(-t); t \rangle. \quad (7)$$

In summary, averages like  $\langle A(t); 0 \rangle$  and correlation functions  $\langle A(t)B; 0 \rangle$  are the central properties of interest for a macroscopic description of physical systems. The microscopic dynamics can be represented in terms of the observables  $A(\Gamma, t)$  or the states  $\rho(\Gamma, t)$  which are determined from specified initial values and the equations

$$(\partial_t - L) A(\Gamma, t) = 0, \quad (\partial_t + \bar{L}) \rho(\Gamma, t) = 0. \quad (8)$$

In the following most of the analysis is done in terms of the states, and the associated equation of motion is known as the Liouville equation.

### **Liouville Equation and Cooling**

For an isolated system, the total energy decreases monotonically due to the loss of energy on each pair collision. This is reflected in a decrease of the average kinetic energy of the particles between collisions and hence is referred to as collisional “cooling”. The energy per particle at time  $t$  and its loss are

$$\epsilon(t) \equiv N^{-1} \langle E; t \rangle, \quad \omega(t) \equiv -\partial_t \epsilon(t) = N^{-1} \langle LE; t \rangle. \quad (9)$$

This cooling effect is common to all solutions to the Liouville equation and it is useful to separate the dynamics into that due to this cooling and the residual time dependence

$$\rho(\Gamma, t) \equiv \rho(\Gamma, \epsilon(t), t). \quad (10)$$

The Liouville equation then can be written

$$\partial_t \rho(\Gamma, \epsilon, t) |_\epsilon + (-\omega(\epsilon, t) \partial_\epsilon + \bar{L}) \rho(\Gamma, \epsilon, t) = 0. \quad (11)$$

The time derivative is now taken at constant  $\epsilon$ . The notation  $\omega(\epsilon, t)$  reflects the fact that it is a linear functional of  $\rho(\Gamma, \epsilon, t)$ , from its definition (9). This is a useful form that isolates a primary effect of the nonconservative forces (cooling) from the residual dynamics that will be associated with relaxation of the spatial inhomogeneities of interest below. For notational simplicity (11) is written

$$(\partial_t + \bar{L}) \rho(\Gamma, \epsilon, t) = 0, \quad \bar{L} \equiv -\omega(\epsilon, t) \partial_\epsilon + \bar{L}. \quad (12)$$

The corresponding equation for observables is

$$(\partial_t - \mathcal{L}) A(\Gamma, \epsilon, t) = 0, \quad \mathcal{L} \equiv -\partial_\epsilon \omega(\epsilon, t) + L. \quad (13)$$

where it is understood that  $\partial_\epsilon$  operates on everything to its right.

### **Stationary Homogeneous State**

An isolated normal fluid supports an equilibrium state. This is a stationary solution to the Liouville equation with translational invariance, the Gibbs states. From the discussion above it is clear that isolated granular fluids have no truly stationary state due to cooling. However, there is a “universal” homogeneous state similar to the Gibbs state in the sense that a wide class of homogeneous initial states rapidly approach this state, on the time scale of a few collisions per particle. It is simple in the sense that all of its time dependence is that associated with cooling

$$\rho_0(\Gamma, t) = \rho_0(\{\mathbf{q}_{ij}, \mathbf{v}_i\}, \epsilon(t)). \quad (14)$$

Here,  $\mathbf{q}_{ij} = \mathbf{q}_i - \mathbf{q}_j$  so the solution also has translational invariance. In the representation (12) it is seen to be a stationary solution to the Liouville equation

$$\bar{L}\rho_0 = 0. \quad (15)$$

There is no longer any explicit time dependence since  $\omega(\epsilon, t) = \omega(\epsilon)$  and  $\bar{L} = -\omega(\epsilon) \partial_\epsilon + \bar{L}$  for this state. This solution is referred to as the homogeneous cooling state (HCS). Clearly, it is the close analogue of the Gibbs state for a normal fluid. It is an example of a “normal” state in the sense that all of its time dependence occurs through one of the hydrodynamic fields (the energy). This concept is sharpened below.

## Macroscopic Balance Equations

The origins of a macroscopic description for a fluid are the balance equations for the average mass density  $\langle m(\mathbf{r}) ; t \rangle$ , energy density  $\langle e(\mathbf{r}) ; t \rangle$ , and momentum density  $\langle \mathbf{g}(\mathbf{r}) ; t \rangle$ , where  $\mathbf{r}$  denotes an arbitrary field point within the system [16]. These will be referred to as the hydrodynamic fields since they are the ones that are expected to obey the hydrodynamic equations under appropriate conditions. The phase functions  $m(\Gamma, \mathbf{r})$ ,  $e(\Gamma, \mathbf{r})$ , and  $\mathbf{g}(\Gamma, \mathbf{r})$  are well known and their explicit forms will not be needed here. They will be denoted collectively by  $a_\alpha(\mathbf{r})$

$$\partial_t a_\alpha(\mathbf{r}, t) \leftrightarrow \{m(\mathbf{r}), e(\mathbf{r}), \mathbf{g}(\mathbf{r})\}. \quad (16)$$

It follows from (8) that they obey the microscopic balance equations [14]

$$\partial_t a_\alpha(\mathbf{r}, t) = La_\alpha(\mathbf{r}, t) = -\nabla \cdot \mathbf{b}_\alpha(\mathbf{r}, t) - \delta_{\alpha 2} w(\mathbf{r}, t). \quad (17)$$

To obtain this result, it has been recognized that the quantity  $-La_\alpha(\mathbf{r}, t)$  can be written as the sum of a divergence  $\nabla \cdot \mathbf{b}_\alpha(\mathbf{r}, t)$  plus a remainder  $w(\mathbf{r}, t)$  that cannot be so represented. For a normal fluid  $w(\mathbf{r}, t)$  vanishes and (17) become the local conservation laws for mass, energy, and momentum. The  $\mathbf{b}_\alpha(\mathbf{r}, t)$  are the corresponding fluxes. This clarifies why  $a_\alpha(\mathbf{r}, t)$  are selected for a macroscopic description. Their time dependence is determined by the scale of the spatial gradients of the fluxes, and averages of the latter become small as the system approaches homogeneity. Consequently, on long time scales the  $\langle a_\alpha(\mathbf{r}) ; t \rangle$  are the only surviving dynamical variables, and it is under these conditions that these fields obey hydrodynamic equations. The mass and momentum are conserved for a granular fluid as well, but there is a loss of energy  $w(\mathbf{r}, t)$  due to the non-conservative forces. It is no longer obvious that the energy is still one of the slow variables since its time scale is coupled to  $w(\mathbf{r}, t)$  which does not become small for nearly homogeneous states. Thus, an additional requirement for the existence of a macroscopic description in terms of  $\langle a_\alpha(\mathbf{r}) ; t \rangle$  is that the time scale of  $\langle e(\mathbf{r}) ; t \rangle / \langle w(\mathbf{r}) ; t \rangle$  must be larger than that for non-hydrodynamic properties. This issue is discussed further below.

The macroscopic balance equations follow from the averages of (17)

$$\partial_t \langle a_\alpha(\mathbf{r}) ; t \rangle + \nabla \cdot \langle \mathbf{b}_\alpha(\mathbf{r}) ; t \rangle = -\delta_{\alpha 2} \langle w(\mathbf{r}) ; t \rangle. \quad (18)$$

These equations are formally exact, but of little practical use as they do not form a closed (self-deterministic) set of equations for  $\langle a_\alpha(\mathbf{r}) ; t \rangle$ . Closure requires expressing the average flux  $\langle \mathbf{b}_\alpha(\mathbf{r}) ; t \rangle$  and energy loss  $\langle w(\mathbf{r}) ; t \rangle$

as functionals of the fields  $\langle a_\alpha(\mathbf{r}) ; t \rangle$ . Such relationships are called “constitutive equations”. The combination of the exact balance equations with some form of constitutive equations provides the most general definition of hydrodynamics.

Construction of the constitutive equations is simplified by extracting the effects of convection. The velocity  $\mathbf{U}(\mathbf{r}, t)$  of a cell at point  $\mathbf{r}$  is defined in terms of the average momentum

$$\langle \mathbf{g}(\mathbf{r}) ; t \rangle \equiv \langle m(\mathbf{r}) ; t \rangle \mathbf{U}(\mathbf{r}, t). \quad (19)$$

The fluxes are functions of the positions and velocities  $\mathbf{b}_\alpha(\mathbf{r}) = \mathbf{b}_\alpha(\mathbf{r}; \{\mathbf{q}_i, \mathbf{v}_i\}) = \mathbf{b}_\alpha(\mathbf{r}; \{\mathbf{q}_i, \mathbf{V}_i + \mathbf{U}(\mathbf{r})\})$ , where the velocity in the local rest frame has been introduced,  $\mathbf{V}_i = \mathbf{v}_i - \mathbf{U}(\mathbf{r}, t)$ . Then defining the microscopic flux in the rest frame by  $\mathbf{b}'_\alpha(\mathbf{r}) = \mathbf{b}_\alpha(\mathbf{r}; \{\mathbf{q}_i, \mathbf{V}_i\})$  it follows that the average flux has the form [3]

$$\begin{aligned} \langle \mathbf{b}_\alpha(\mathbf{r}) ; t \rangle &= \langle \mathbf{b}'_\alpha(\mathbf{r}) ; t \rangle + c_{\alpha\eta} \mathbf{U}(\mathbf{r}, t) \cdot \langle \mathbf{b}'_\eta(\mathbf{r}) ; t \rangle \\ &\quad + \mathbf{U}(\mathbf{r}, t) d_\alpha (\{\langle a_\nu(\mathbf{r}) ; t \rangle\}). \end{aligned} \quad (20)$$

The first term is the flux of mass, energy, and momentum in a fluid element at rest, and represents the dissipative processes. The second and third terms are proportional to the flow velocity  $\mathbf{U}(\mathbf{r}, t)$  and are associated with convection. The coefficients of these terms are explicit functions of the fields  $\langle a_\alpha(\mathbf{r}) ; t \rangle$  (as is  $\mathbf{U}(\mathbf{r}, t)$ ). For a normal fluid, neglect of the rest frame fluxes leads to the perfect fluid Euler hydrodynamic equations. Hence, determination of the constitutive equations is reduced to expressing the rest frame fluxes and energy loss as functionals of the fields.

## “Normal” States and Hydrodynamics

A hydrodynamic description is a closed set of equations for the hydrodynamic fields,  $\langle a_\alpha(\mathbf{r}) ; t \rangle$ . This follows from the exact macroscopic balance equations if the energy loss and fluxes can be represented as functionals of these fields

$$\begin{aligned} \langle w(\mathbf{r}) ; t \rangle &\rightarrow \omega(\mathbf{r} | \langle a_\alpha ; t \rangle), \\ \langle \mathbf{b}_\alpha(\mathbf{r}) ; t \rangle &\rightarrow \beta_\alpha(\mathbf{r} | \langle a_\alpha ; t \rangle). \end{aligned} \quad (21)$$

The arrow is used to indicate that such a functional representation need not be valid on all length and time scales, and any such restrictions constitute the domain of validity for hydrodynamics. The notation here and below is such that  $f(\mathbf{r}, t, \{\langle a_\alpha(\mathbf{r}) ; t \rangle\})$  denotes a function of  $\mathbf{r}$ ,  $t$  and of the fields  $\langle a_\alpha(\mathbf{r}) ; t \rangle$  at the point  $\mathbf{r}$ , while  $f(\mathbf{r}, t | \langle a_\alpha ; t \rangle)$  denotes a function of  $\mathbf{r}$ ,  $t$  and a functional of the  $\langle a_\alpha ; t \rangle$  at all space points. With such constitutive re-

lations the macroscopic balance Eqs. (18) become hydrodynamic equations

$$\partial_t \langle a_\alpha(\mathbf{r}); t \rangle + \nabla \cdot \boldsymbol{\beta}(\mathbf{r} | \langle a_\alpha; t \rangle) = -\delta_{\alpha 2} \omega(\mathbf{r} | \langle a_\alpha; t \rangle). \quad (22)$$

The average energy loss and fluxes are averages of specific functions of the particle positions and velocities, and hence are linear functionals of the solution to the Liouville equation. The existence of constitutive equations is therefore related to a special property of the solution which will be called “normal” (this terminology originates in a related context for derivation of hydrodynamics from the Boltzmann kinetic equation [3]). The class of “normal” distributions is defined by the functional forms

$$\rho_n(\Gamma, t) = \rho_n(\{\mathbf{q}_{ij}, \mathbf{v}_i\} | \langle a_\alpha; t \rangle). \quad (23)$$

All time dependence and all the breaking of translational invariance for normal states occurs only through the hydrodynamic fields. A familiar example of a normal distribution for real fluids is the *local* Gibbs distribution

$$\rho_{el}(\Gamma | \langle a_\alpha; t \rangle) = \exp \left\{ q - \int d\mathbf{r} y_\alpha(\mathbf{r}, t) a_\alpha(\mathbf{r}) \right\}. \quad (24)$$

Here  $q$  is a normalization constant, and  $y_\alpha(\mathbf{r}, t)$  are conjugate fields determined by the requirement that the averages of  $a_\alpha(\mathbf{r})$  give the specified values  $\langle a_\alpha(\mathbf{r}); t \rangle$ . In this way  $y_\alpha(\mathbf{r}, t)$  are functionals of the hydrodynamic fields and  $\rho_{el}(\Gamma | \langle a_\alpha; t \rangle)$  is normal. The importance of normal solutions is that they yield directly the desired functionals of (21)

$$\omega(\mathbf{r} | \langle a_\alpha; t \rangle) = \int d\Gamma \rho_n(\Gamma | \langle a_\alpha; t \rangle) w(\mathbf{r}) \quad (25)$$

$$\boldsymbol{\beta}_\alpha(\mathbf{r} | \langle a_\alpha; t \rangle) = \int d\Gamma \rho_n(\Gamma | \langle a_\alpha; t \rangle) \mathbf{b}_\alpha(\mathbf{r}). \quad (26)$$

The normal state in (25) and (26) must be a solution to the Liouville equation. In general, the time derivative in the Liouville equation can be separated into that which occurs through  $\langle a_\alpha; t \rangle$  plus the residual time dependence, generalizing (10)

$$\rho(\Gamma, t) = \rho(\Gamma, t | \langle a_\alpha; t \rangle). \quad (27)$$

The Liouville equation then becomes

$$\begin{aligned} & \partial_t \rho |_{\langle a_\alpha; t \rangle} - \int d\mathbf{r} \frac{\delta \rho}{\delta \langle a_\alpha(\mathbf{r}); t \rangle} \\ & \cdot \{ \nabla \cdot \langle \mathbf{b}_\alpha(\mathbf{r}); t \rangle + \delta_{\alpha 2} \langle w(\mathbf{r}); t \rangle \} + \bar{L}\rho = 0. \end{aligned} \quad (28)$$

A normal solution results when  $\partial_t \rho_n |_{\langle a_\alpha; t \rangle} \rightarrow 0$ . For specified fields, (28) becomes an equation for the  $\Gamma$  dependence of the normal phase space density as a functional of the fields. This dependence then allows determination of the normal forms in (25) and (26). Finally, with the form of the hydrodynamic equation determined at that point, their solution with suitable initial and boundary conditions provides the explicit forms for the fields, and completes the normal solution. The existence and determination of this solution is the central problem for establishing a hydrodynamic description for both normal and granular fluids.

The concept of a normal solution and its use in the macroscopic balance equations makes no special reference to whether the fluid is atomic or granular, and is not restricted to states near homogeneity. In this general context, hydrodynamics is not a simple set of local partial differential equations such as the familiar Navier–Stokes equations. The latter are a special case of this more general idea, and their inadequacy for some conditions should not be interpreted as the absence of a more complex hydrodynamic description.

In closing this Section a qualitative explanation of why a normal solution can be expected is given, by analogy with the similar expectation for atomic fluids. For a wide class of initial states there is a first stage of rapid velocity relaxation in each small region toward the universal homogeneous state (HCS or Gibbs, respectively). However, the hydrodynamic parameters of that universal state are specific to each region so it is only locally homogenous, as in (24) for the atomic fluid. Subsequently, these differences in the parameters of neighboring cells are decreased by the fluxes of mass, energy, and momentum across their boundaries. It is this second stage where a normal description in terms of the hydrodynamic fields can be expected, indicating also that the space and time scales for a hydrodynamic description should be large compared to those of the first stage. This basic conceptual picture is essentially the same for both atomic and granular fluids, and the rapid approach of the first stage is indeed observed in molecular dynamics simulation studies of both equilibrium and the HCS.

### Navier–Stokes Approximation

Equation (28) presents a formidable problem and further progress requires specialization to specific cases of interest. Perhaps the simplest of these are weakly inhomogeneous states. These are states for which all spatial gradients of first order are small and all higher order derivatives are negligible. Small gradients means that the relative change in the hydrodynamic fields over the largest micro-

scopic length scale  $\ell_0$  is small:  $\ell_0 \partial_r \ln \langle a_\alpha; t \rangle \ll 1$ . There are two characteristic length scales, the mean free path and the grain diameter. For a dilute gas the mean free path is largest, while for a dense fluid the grain size is largest. Under these conditions a solution to the Liouville equation can be sought as an expansion to leading order in these small gradients. This will be referred to as the Navier-Stokes approximation.

According to the discussion at the end of the last section, a normal solution is expected after the system has relaxed to its local HCS form, denoted by  $\rho_{0\ell}(\Gamma | \langle a_\alpha; t \rangle)$ , representing the fluid as having each cell in its own HCS. Define the deviations of the hydrodynamic fields from some common reference value by

$$\delta \langle a_\alpha; t \rangle = \langle a_\alpha; t \rangle - a_{0\alpha}, \quad (29)$$

where  $a_{0\alpha}$  is the same for all cells. Then, the local HCS must satisfy the conditions

$$\begin{aligned} \rho_{0\ell}(\Gamma | a_{0\alpha} + \delta \langle a_\alpha; t \rangle) |_{\delta \langle a_\alpha; t \rangle=0} \\ = \rho_0(\Gamma; a_{0\alpha}), \end{aligned} \quad (30)$$

$$\begin{aligned} \frac{\partial \rho_0}{\partial a_{0\alpha}} \\ = \int d\Gamma \frac{\delta \rho_{0\ell}(\Gamma | a_{0\alpha} + \delta \langle a_\alpha; t \rangle)}{\delta \langle a_\alpha(\mathbf{r}); t \rangle} |_{\delta \langle a_\alpha; t \rangle=0}, \dots \end{aligned} \quad (31)$$

i.e., the local HCS and all of its functional derivatives must agree with those of the HCS in the homogenous limit. Also, as a normal distribution its time dependence is through the exact hydrodynamic fields for the fluid state considered. This means the averages of the corresponding microscopic fields  $a_\alpha(\mathbf{r})$  for the local HCS and for the solution to the Liouville equation must be the same

$$\int d\Gamma (\rho - \rho_{0\ell}) a_\alpha(\mathbf{r}) = 0. \quad (32)$$

A more complete discussion of the construction of  $\rho_{0\ell}$  from knowledge of  $\rho_0$  is given elsewhere [14]. For the purposes here properties (30), (31), and (32) are sufficient.

The local HCS distribution,  $\rho_{0\ell}$ , is not a solution to the Liouville equation except in limit that all hydrodynamic fields become the same for each cell. Instead, it is a reference state approximating the actual solution after its first stage of velocity relaxation. To construct a solution  $\rho$  define its deviation from  $\rho_{0\ell}$  by

$$\rho(\Gamma, t | \langle a_\alpha; t \rangle) = \rho_{0\ell}(\Gamma, t | \langle a_\alpha; t \rangle) + \Delta(\Gamma, t | \langle a_\alpha; t \rangle). \quad (33)$$

The Liouville Equation (28) gives

$$\begin{aligned} \partial_t \Delta - \int d\mathbf{r}' \frac{\delta \Delta}{\delta \langle a_\alpha(\mathbf{r}'); t \rangle} \\ \cdot \{ \nabla \cdot \langle \mathbf{b}_\alpha(\mathbf{r}); t \rangle + \delta_{\alpha 2} \langle w(\mathbf{r}); t \rangle \} + \bar{L} \Delta \\ = \int d\mathbf{r}' \frac{\delta \rho_{0\ell}}{\delta \langle a_\alpha(\mathbf{r}'); t \rangle} \\ \cdot \{ \nabla \cdot \langle \mathbf{b}_\alpha(\mathbf{r}); t \rangle + \delta_{\alpha 2} \langle w(\mathbf{r}); t \rangle \} - \bar{L} \rho_{0\ell}. \end{aligned} \quad (34)$$

This equation is still exact, but if only small gradient states are considered it simplifies by retaining terms only of first order in the gradients. To be precise, the ultimate use of this solution is to calculate local properties of the form

$$\begin{aligned} A(\mathbf{r}, t | \{y_\alpha(t)\}) \\ = \int d\Gamma a(\Gamma, \mathbf{r}) \rho(\Gamma, t | \langle a_\alpha(\mathbf{r}); t \rangle + \delta \langle a_\alpha; t \rangle). \end{aligned} \quad (35)$$

Therefore, in the following analysis the gradient expansions are referred to the field point  $\mathbf{r}$  of interest,  $\langle a_\alpha; t \rangle = \langle a_\alpha(\mathbf{r}); t \rangle + \delta \langle a_\alpha; t \rangle$ , i.e. the common reference values in (29) are the exact field values at the chosen point,  $a_{0\alpha} = \langle a_\alpha(\mathbf{r}); t \rangle$ . The gradient expansion is carried out relative to these values. Of course the results will be general and applicable to any choice for  $\mathbf{r}$ .

The details of the gradient expansion are given in the Appendix. The solution to the Liouville equation to first order in the gradients is

$$\begin{aligned} \rho(\Gamma, t | \langle a_\alpha; t \rangle) = \rho_0(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) + (1 - \mathcal{P}) \\ \cdot \left( \mathbf{M}_\beta(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) + \int_0^t dt' \left( e^{-(I\bar{L} + K^T)t'} \right)_\beta \right. \\ \left. \cdot (1 - \mathcal{P}) \boldsymbol{\Upsilon}_v(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \right) \cdot \nabla \langle a_\beta(\mathbf{r}); t \rangle. \end{aligned} \quad (36)$$

with the definitions

$$\mathbf{M}_\beta = \int d\mathbf{r}' \left( \frac{\delta \rho_{0\ell}}{\delta \langle a_\alpha(\mathbf{r}'); t \rangle} \right)_{\delta \langle a_\alpha; t \rangle=0} \mathbf{r}', \quad (37)$$

$$\boldsymbol{\Upsilon}_\alpha = -(I\bar{L} + K^T)_{\alpha\beta} \mathbf{M}_\beta. \quad (38)$$

The generator for the dynamics  $I\bar{L} + K^T$  has a contribution from  $\bar{L}$  which is the same as in (12), with  $\omega$  evaluated for the HCS as a function of the exact hydrodynamic fields at the point  $\mathbf{r}$  and time  $t$

$$\bar{L} = -\omega_0(\langle a_\alpha(\mathbf{r}); t \rangle) \partial_{(e(\mathbf{r}); t)} + \bar{L}. \quad (39)$$

The second contribution to the generator of the dynamics is the transpose of the matrix  $K_{\alpha\beta}$

$$K_{\alpha\beta} = \delta_{\alpha 2} \frac{\partial \omega_0(\langle a_\alpha(\mathbf{r}); t \rangle)}{\partial \langle a_\beta(\mathbf{r}); t \rangle}. \quad (40)$$

Finally,  $\mathcal{P}$  is a projection operator

$$\begin{aligned} \mathcal{P}X &= \Psi_\beta \int d\Gamma A_\beta X, \\ A_\beta &= V^{-1} \int d\mathbf{r} a_\beta(\mathbf{r}), \\ \Psi_\beta &\equiv \frac{\partial \rho_0}{\partial \langle a_\beta(\mathbf{r}); t \rangle} \end{aligned} \quad (41)$$

The phase functions  $A_\beta$  and  $\Psi_\beta$  form a biorthogonal set in the sense

$$\int d\Gamma A_\alpha \Psi_\beta = \delta_{\alpha\beta}. \quad (42)$$

The  $A_\alpha$  are the usual global invariants of the Liouville operator  $\bar{L}$  for a normal fluid; it is shown in the Appendix that the  $\Psi_\beta$  are the invariants of the new generator for dynamics in a granular fluid

$$(I\bar{L}_T + K^T)_{v\beta} \Psi_\beta = 0. \quad (43)$$

Equation (36) is not quite the normal solution desired. All terms depend on time through  $\langle a_\beta(\mathbf{r}); t \rangle$  as required, except for the last term which has an additional explicit time dependence through the upper limit of the time integral. This time dependence becomes negligible if the integrand is effectively non-zero after some short time scale  $\tau$ . Then for  $t \gg \tau$  the time integral becomes independent of  $t$  and can be taken formally to infinity. Thus, a normal solution is attained for this time scale

$$\begin{aligned} \rho_n(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) &= \rho_0(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &+ (1 - \mathcal{P}) \left( \mathbf{M}_\beta(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \right. \\ &+ \lim_{t_0 \rightarrow \infty} \int_0^{t_0} dt' \left( e^{-(I\bar{L} + K^T)t'} \right)_{\beta v} \\ &\cdot (1 - \mathcal{P}) \mathbf{Y}_v(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \Big) \\ &\cdot \nabla \langle a_\beta(\mathbf{r}); t \rangle. \end{aligned} \quad (44)$$

It is expected that the integrand should have this property of a short time scale since the domain of operation for the generator of time dependence is functions with

translational invariance (as a consequence of the gradient expansion). Hence there are no explicit slow hydrodynamic modes of finite wavelength. Also, there is no contribution from the homogeneous hydrodynamics (that for the invariants) due to the orthogonal projection  $(1 - \mathcal{P})$ . The appearance of this projection is an essential self-consistency of the analysis, and occurs as well for normal fluids. The expression (44) is only formal and the actual limit should be taken in the weak sense only after (36) has been used to define average properties. A technical complication is the occurrence of periodic time dependence, the Poincaré recurrence time. This can be removed by considering the thermodynamic limit of  $V \rightarrow \infty, N \rightarrow \infty$  at constant  $N/V$ . Therefore, averages using the normal solution to the Liouville equation are understood as having the thermodynamic limit followed by the long time limit at constant  $\langle a_\alpha(\mathbf{r}); t \rangle$ .

An alternative equivalent form results from performing the integral in (44) using the explicit form (38) and the property (99) of the Appendix

$$\begin{aligned} \rho_n(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) &= \rho_0(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &+ \lim_{t_0 \rightarrow \infty} (1 - \mathcal{P}) \left( e^{-(I\bar{L} + K^T)t_0} \right)_{\beta v} \\ &\cdot \mathbf{M}_v(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \cdot \nabla \langle a_\beta(\mathbf{r}); t \rangle. \end{aligned} \quad (45)$$

The decay time for the integrand of (44) now becomes the time after which (45) reaches its normal form.

## Constitutive Equations

The exact macroscopic balance equations are given by (22), and the necessary constitutive equations are given by (26) and (27) as averages over the normal solution. These can be made more explicit now using the small gradient result (44). Since the latter is a local function of the fields, the constitutive equations also will be local. Furthermore, since all components of the gradients in (45) depend on the common value  $\langle \mathbf{g}(\mathbf{r}); t \rangle$ , this can be eliminated through a Galilean transformation so that all properties refer to a fluid element at rest. Of course, the gradients of  $\langle \mathbf{g}(\mathbf{r}); t \rangle$  in that fluid element are nonzero.

Consider first the energy loss function  $\omega$

$$\begin{aligned} \omega(\langle a_\alpha(\mathbf{r}); t \rangle) &= \int d\Gamma \rho_n(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) w(\mathbf{r}) \\ &= \int d\Gamma \rho_n(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \bar{w}. \end{aligned} \quad (46)$$

The coefficients of the gradient in the normal solution have translational invariance and the average is independent of  $\mathbf{r}$ , except through its parametrization by  $\langle a_\alpha(\mathbf{r}); t \rangle$ .

The second equality takes this into account by replacing  $w(\mathbf{r})$  by its average  $\bar{w}$

$$\bar{w} = V^{-1} \int d\Gamma w(\mathbf{r}). \quad (47)$$

Since  $\omega(\langle a_\alpha(\mathbf{r}); t \rangle)$  is a scalar, fluid symmetry restricts the contributions to first order in the gradients to

$$\begin{aligned} & \omega(\langle a_\alpha(\mathbf{r}); t \rangle) \\ &= \omega_0(\langle a_\alpha(\mathbf{r}); t \rangle) + \omega_1(\langle a_\alpha(\mathbf{r}); t \rangle) \nabla \cdot \mathbf{U}(\mathbf{r}, t). \end{aligned} \quad (48)$$

Here, the flow velocity  $\mathbf{U}(\mathbf{r}, t)$  of (19) has been used in place of the momentum density. The first term is the contribution from the HCS distribution

$$\omega_0(\langle a_\alpha(\mathbf{r}); t \rangle) = \int d\Gamma \rho_0(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \bar{w}. \quad (49)$$

The coefficient of  $\nabla \cdot \mathbf{U}(\mathbf{r}, t)$  is

$$\begin{aligned} \omega_1(\langle a_\alpha(\mathbf{r}); t \rangle) &= \lim_{t_0 \rightarrow \infty} C_\omega(t_0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &= C_\omega(0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &+ \lim_{t_0 \rightarrow \infty} \int_0^{t_0} \partial_{t'} C_\omega(t', \langle a_\alpha(\mathbf{r}); t \rangle) dt' \end{aligned} \quad (50)$$

with the correlation function defined by

$$C_\omega(t) = \int d\Gamma \bar{w} (1 - \mathcal{P}) e^{-\bar{\mathcal{I}}t} M_U, \quad (51)$$

$$M_U \equiv \frac{1}{3} \int d\mathbf{r}' \mathbf{r}' \cdot \left( \frac{\delta \rho_{0\ell}}{\delta \mathbf{U}(\mathbf{r}', t)} \right)_{\delta \langle a_\alpha; t \rangle = 0}. \quad (52)$$

The coefficient  $\omega_0$  defines an ‘‘equation of state’’ for the granular hydrodynamics, and gives the first non-trivial result of this analysis. It is similar to the pressure (given below) and is an inherent property of the local state of each cell, independent of the gradients between cells. In contrast,  $\omega_1$  is a true transport coefficient characterizing communication between cells. The first equality of (50) provides the Helfand form for this coefficient, while the second equality gives the equivalent Green–Kubo form. Each has its practical utility, depending on the method used for its approximate evaluation. Both forms have proven useful for normal fluids, and further discussion is provided below. Both  $\omega_0$  and the transport coefficient  $\omega_1$  vanish for normal fluids since they characterize collisional energy loss.

The fluxes  $\beta_\alpha$  of (26) can be determined in a similar way. As indicated in (20), only the rest frame flux  $\beta'_\alpha$  is required. Furthermore, since all components of the gradients in (45) depend on the common value  $\langle \mathbf{g}(\mathbf{r}); t \rangle$ , this

can be eliminated through a Galilean transformation so that all properties refer to a fluid element at rest. Of course, the gradients of  $\langle \mathbf{g}(\mathbf{r}); t \rangle$  in that fluid element are nonzero. The component  $\beta'_1$  is the rest frame mass flux which is expected to vanish in order to give the continuity equation. This follows from the fact that  $\mathbf{b}'_1(\mathbf{r})$  is the microscopic momentum density

$$\begin{aligned} \beta'_1 &= \int d\Gamma \mathbf{g}(\mathbf{r}) \rho_0(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &+ \lim_{t_0 \rightarrow \infty} \int d\Gamma \mathbf{g}(\mathbf{r}) (1 - \mathcal{P}) (\dots) = 0 \end{aligned} \quad (53)$$

The first term vanishes since  $\langle \mathbf{g}(\mathbf{r}); t \rangle = 0$  in the rest frame, and the second term vanishes since  $(1 - \mathcal{P})$  projects orthogonal to the mass, energy, and momentum. Thus, the expected continuity equation is verified.

The fluxes  $\beta'_2$  and  $\beta'_\alpha$  for  $\alpha = 3, 4, 5$  are the rest frame energy and momentum fluxes. The energy flux transforms like a vector and therefore fluid symmetry (translational and rotational invariance) requires that it can depend only on gradients of scalars

$$\begin{aligned} \beta'_2 &= -\lambda(\langle a_\alpha(\mathbf{r}); t \rangle) \nabla T(\mathbf{r}, t) \\ &- \mu(\langle a_\alpha(\mathbf{r}); t \rangle) \nabla \langle m(\mathbf{r}); t \rangle. \end{aligned} \quad (54)$$

To make the connection with Fourier’s law for an atomic fluid, a temperature  $T(\mathbf{r}, t)$  has been introduced through the definition

$$\langle e(\mathbf{r}); t \rangle \equiv e_0(\langle m(\mathbf{r}); t \rangle, T(\mathbf{r}, t)). \quad (55)$$

For an atomic fluid the function  $e_0(\langle m(\mathbf{r}); t \rangle, T(\mathbf{r}, t))$  is chosen to be the thermodynamic internal energy density. As there is no thermodynamics for a granular fluid this function is arbitrary and simply constitutes a change of variables from  $\langle e(\mathbf{r}); t \rangle, \langle m(\mathbf{r}); t \rangle$  to  $T(\mathbf{r}, t), \langle m(\mathbf{r}); t \rangle$ . In this form (54) is a generalization of Fourier’s law where  $\lambda$  is the thermal conductivity [27]. The contribution from the gradient of the mass density is new to granular fluids ( $\mu = 0$  for atomic fluids). These coefficients are given by

$$\begin{aligned} \lambda(\langle a_\alpha(\mathbf{r}); t \rangle) &= \lim_{t_0 \rightarrow \infty} C_\lambda(t_0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &= C_\lambda(0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &+ \lim_{t_0 \rightarrow \infty} \int_0^{t_0} \partial_{t'} C_\lambda(t', \langle a_\alpha(\mathbf{r}); t \rangle) dt' \end{aligned} \quad (56)$$

$$\begin{aligned} \mu(\langle a_\alpha(\mathbf{r}); t \rangle) &= \lim_{t_0 \rightarrow \infty} C_\mu(t_0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &= C_\mu(0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &+ \lim_{t_0 \rightarrow \infty} \int_0^{t_0} \partial_{t'} C_\mu(t', \langle a_\alpha(\mathbf{r}); t \rangle) dt' \end{aligned} \quad (57)$$

with the correlation functions

$$C_\lambda(t) = \frac{1}{3} \int d\Gamma \beta'_2 \cdot (1 - \mathcal{P}) e^{-(\bar{\mathcal{L}} + K_{22})t} \mathbf{M}_T , \quad (58)$$

$$\begin{aligned} C_\mu(t) = & \frac{1}{3} \int d\Gamma \beta'_2 \cdot (1 - \mathcal{P}) \left( e^{-\bar{\mathcal{L}}t} \left( \mathbf{M}_m + \frac{K_{21}}{K_{22}} \mathbf{M}_T \right) \right. \\ & + e^{-(\bar{\mathcal{L}} + K_{22})t} \\ & \cdot \left( \frac{\partial e_0}{\partial \langle m(\mathbf{r}); t \rangle} - \frac{K_{21}}{K_{22}} \right) \frac{\partial T}{\partial e_0} |_{\langle m(\mathbf{r}); t \rangle} \mathbf{M}_T \left. \right) , \end{aligned} \quad (59)$$

$$\mathbf{M}_T \equiv \int d\mathbf{r}' \mathbf{r}' \left( \frac{\delta \rho_0 \ell}{\delta T(\mathbf{r}', t)} \right)_{\delta \langle a_\alpha; t \rangle = 0} . \quad (60)$$

$$\mathbf{M}_m \equiv \int d\mathbf{r}' \mathbf{r}' \left( \frac{\delta \rho_0 \ell}{\delta \langle m(\mathbf{r}'); t \rangle} \right)_{\delta \langle a_\alpha; t \rangle = 0} . \quad (61)$$

Finally, the set of vectors  $\beta'_\alpha$  for  $\alpha = 3, 4, 5$  define the pressure tensor  $\beta'_\alpha \Leftrightarrow P_{ij}$ . Fluid symmetry then determines that it can couple only to the momentum gradients, or equivalently the flow velocity gradients, in the form

$$\begin{aligned} P_{ij} = & p(\langle a_\alpha(\mathbf{r}); t \rangle) \delta_{ij} - \eta(\langle a_\alpha(\mathbf{r}); t \rangle) \\ & \cdot \left( \partial_i U_j(\mathbf{r}, t) + \partial_j U_i(\mathbf{r}, t) - \frac{2}{3} \delta_{ij} \nabla \cdot \mathbf{U}(\mathbf{r}, t) \right) \\ & - \kappa(\langle a_\alpha(\mathbf{r}); t \rangle) \delta_{ij} \nabla \cdot \mathbf{U}(\mathbf{r}, t) . \end{aligned} \quad (62)$$

The scalar function  $p(\langle a_\alpha(\mathbf{r}); t \rangle)$  is the pressure, now identified as

$$p(\langle a_\alpha(\mathbf{r}); t \rangle) = \int d\Gamma \rho_0(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \beta'_{3x} . \quad (63)$$

The transport coefficients in (62) are the shear viscosity  $\eta(\langle a_\alpha(\mathbf{r}); t \rangle)$  and the bulk viscosity  $\kappa(\langle a_\alpha(\mathbf{r}); t \rangle)$  given by

$$\begin{aligned} \eta(\langle a_\alpha(\mathbf{r}); t \rangle) = & \lim_{t_0 \rightarrow \infty} C_\eta(t_0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ = & C_\eta(0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ & + \lim_{t_0 \rightarrow \infty} \int_0^{t_0} \partial_{t'} C_\eta(t', \langle a_\alpha(\mathbf{r}); t \rangle) dt' \end{aligned} \quad (64)$$

$$\begin{aligned} \kappa(\langle a_\alpha(\mathbf{r}); t \rangle) = & \lim_{t_0 \rightarrow \infty} C_\kappa(t_0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ = & C_\kappa(0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ & + \lim_{t_0 \rightarrow \infty} \int_0^{t_0} \partial_{t'} C_\kappa(t', \langle a_\alpha(\mathbf{r}); t \rangle) dt' \end{aligned} \quad (65)$$

with the correlation functions

$$C_\eta(t) = \int d\Gamma \beta'_{3y} \cdot (1 - \mathcal{P}) e^{-\bar{\mathcal{L}}t} M_\eta , \quad (66)$$

$$C_\kappa(t) = \int d\Gamma \beta'_{3x} \cdot (1 - \mathcal{P}) e^{-\bar{\mathcal{L}}t} M_\kappa , \quad (67)$$

$$M_\eta \equiv \int d\mathbf{r}' x' \left( \frac{\delta \rho_0 \ell}{\delta U_y(\mathbf{r}', t)} \right)_{\delta \langle a_\alpha; t \rangle = 0} . \quad (68)$$

$$M_\kappa \equiv \int d\mathbf{r}' y' \left( \frac{\delta \rho_0 \ell}{\delta U_y(\mathbf{r}', t)} \right)_{\delta \langle a_\alpha; t \rangle = 0} . \quad (69)$$

This completes the formal derivation of the constitutive equations leading to the nonlinear Navier–Stokes equations, including expressions for the cooling rate, energy flux, and pressure tensor including contributions up through first order in the gradients of the hydrodynamic fields. These expressions are functions of the hydrodynamic fields to be determined by their detailed many-body analysis of the correlation functions.

### Green–Kubo Expressions

To contrast the results here with those for an atomic fluid, it is instructive to focus on the Green–Kubo forms for the transport coefficients [3]. These are given by the second equalities of (50), (56), (57), (64), and (65); the first equalities are the corresponding Helfand forms [28]. For atomic fluids there is no counter part to  $\omega_1$  and  $\mu$ . However, there are Green–Kubo expressions for the thermal conductivity and the two viscosities. For the discussion here only the thermal conductivity is considered, whose Green–Kubo expression is

$$\begin{aligned} \lambda(\langle a_\alpha(\mathbf{r}); t \rangle) = & C_\lambda(0, \langle a_\alpha(\mathbf{r}); t \rangle) \\ & + \lim_{t_0 \rightarrow \infty} \int_0^{t_0} \partial_{t'} C_\lambda(t', \langle a_\alpha(\mathbf{r}); t \rangle) dt' \end{aligned} \quad (70)$$

$$\partial_t C_\lambda(t) = \frac{1}{3} \int d\Gamma \beta'_2 \cdot (1 - \mathcal{P}) e^{-(\bar{\mathcal{L}} + K_{22})t} \boldsymbol{\gamma}_\lambda , \quad (71)$$

$$\boldsymbol{\gamma}_e = -(\bar{\mathcal{L}} + K_{22}) \mathbf{M}_e . \quad (72)$$

In contrast the thermal conductivity for an atomic fluid is

$$\lambda(\langle a_\alpha(\mathbf{r}); t \rangle) \rightarrow \lim_{t_0 \rightarrow \infty} \int_0^{t_0} \partial_{t'} C_\lambda(t', \langle a_\alpha(\mathbf{r}); t \rangle) dt' \quad (73)$$

$$\partial_t C_\lambda(t) \rightarrow \frac{1}{3T^2} \int d\Gamma \beta'_2 \cdot (1 - \mathcal{P}) e^{-Lt} \beta'_2 \rho_e , \quad (74)$$

In this last expression  $\rho_e$  is the equilibrium Gibbs ensemble, and it is understood that  $\beta'_2$  is the microscopic expression for the energy flux for a dynamics with conservative forces, and  $L$  is the Liouville generator for the corresponding dynamics.

There are several similarities and differences between the granular and atomic fluid expressions [14]. The latter is the time integral of a energy flux – energy flux equilibrium time correlation function. The granular fluid is similar, with one of the fluxes the same but the other flux is generated from the local HCS state. Also, the generator for the dynamics in the granular case has two additional effects,  $L$  replaced by  $\bar{L} + K_{22}$ , to represents homogeneous cooling of the reference state and its homogeneous response to perturbations. The projection orthogonal to the invariants of each dynamics ( $1 - \mathcal{P}$ ) occurs in both cases as a necessary condition for the long time limit of the time integral, and the corresponding existence of the normal state. Finally, the contribution from  $C_\lambda(0, \langle a_\alpha(\mathbf{r}) ; t \rangle)$  vanishes for a normal fluid (except for singular forces) but is non-zero for the granular fluid due to the non-conservative forces.

### Navier–Stokes Hydrodynamic Equations

In closing this section it is appropriate to record the results of substituting the Navier–Stokes constitutive equations, valid to first order in the gradients, into the exact macroscopic balance equations. This defines the Navier–Stokes hydrodynamic equations for a granular fluid

$$D_t m + m \nabla_{\mathbf{r}} \cdot \mathbf{U} = 0 \quad (75)$$

$$\begin{aligned} D_t e_0 + \omega_0 + \left( p + \omega_1 + \left( \frac{2}{3} \eta - \kappa \right) \nabla \cdot \mathbf{U} \right) \nabla \cdot \mathbf{U} \\ - \eta (\partial_\alpha U_\beta + \partial_\beta U_\alpha) \partial_\alpha U_\beta - \nabla \cdot (\lambda \nabla T + \mu \nabla m) = 0, \end{aligned} \quad (76)$$

$$\begin{aligned} D_t U_\alpha + m^{-1} \partial_\alpha \left( p - \left( \frac{2}{3} \eta + \kappa \right) \nabla \cdot \mathbf{U} \right) \\ - m^{-1} \partial_\beta \eta (\partial_\alpha U_\beta + \partial_\beta U_\alpha) = 0. \end{aligned} \quad (77)$$

For simplicity of notation,  $m \equiv \langle m(\mathbf{r}) ; t \rangle$  in these equations. They are a set of five nonlinear partial differential equations for the variables  $m$ ,  $e_0$ , and  $\mathbf{U}$ . They are a closed set of equations since  $\omega_0$ ,  $p$ , and the transport coefficients  $\omega_1$ ,  $\lambda$ ,  $\mu$ ,  $\eta$ , and  $\kappa$  are defined as functions of these variables. These definitions for the constitutive equations are the primary accomplishment of the statistical mechanical basis for the hydrodynamic equations. The form of (64)–(65) could have been guessed from the outset based on

the macroscopic balance equations and fluid symmetry. The underlying basis in the microdynamics of the particles provides the necessary details for how the parameters of these equations must depend on the fields. Here only the formal definitions have been identified. It is only the first half of the problem of completing these equations, as the evaluation of these definitions poses a difficult many body problem. Still, without this first half, the starting point for that detailed analysis would not be possible. This the case for atomic fluids as well.

### Future Directions

The objective here has been to formulate the basis for a macroscopic description of granular fluids using the fundamental principles of nonequilibrium statistical mechanics. The analysis presented follows that for an atomic fluid. First, the exact macroscopic balance equations are identified. Next, their closure is linked to the concept of a normal state and corresponding normal solution to the Liouville equation. This defines the domain of hydrodynamics in its most general sense, both for atomic and granular fluids. The construction of a normal solution is quite difficult in general, but can be accomplished for states with small gradients relative to locally homogeneous conditions. This gives the Navier–Stokes approximation described here.

Navier–Stokes hydrodynamics is applicable for most common states of atomic fluids, while deviations occur primarily for more complex polymeric molecular fluids. The latter have rheological properties corresponding to larger gradients relative to additional microscopic length and time scales. The construction of normal states in these cases is more difficult and is still at the semi-phenomenological stage [29]. Granular fluids provide a new motivation for renewed efforts to describe these more complex normal states. The reason is that even structurally simple granular fluids composed of spherically symmetric particles can exhibit rheology and other phenomena beyond the Navier–Stokes domain of validity [30,31]. This is due to the cooling rate in the energy balance equation which provides a new internal time scale, that can set the size of hydrodynamic gradients beyond any control through boundary conditions. For example, new steady states are possible for granular fluids due to the balance of this internal cooling with external forcing. In many cases this implies that the hydrodynamic description required is beyond the Navier–Stokes domain. The understanding of constitutive equations in these cases is poor at this point. It is hoped that the formal structure described here will provide the appropriate basis for studies of these problems.

The context of hydrodynamics depends on the formation of a normal state from more complex conditions. Above this has been described qualitatively as a two stage process of rapid velocity relaxation in each cell to a state near the local HCS, followed by hydrodynamic relaxation through exchange of mass, energy, and momentum between the cells on a longer time scale. This separation of microscopic and hydrodynamic time scales is essential to the dominance of the hydrodynamic excitations over all others at large space and time scales. It is justified for atomic fluids since the hydrodynamic times are determined by the wavelength of the phenomena studied. As the system approaches homogeneity, these time scales become much larger than the microscopic excitations and hydrodynamics prevails at large times. However, there is an additional hydrodynamic time scale for granular fluids, the cooling rate, which is not set by the wavelength alone. It would seem that this additional time scale must be large as well, implying a weak cooling rate. This condition is too strong. What matters is the rate of the approach to the homogeneous state, not any dynamics of that final state. In the above derivation of hydrodynamics the final form for the solution to the Liouville equation, Eq. (44) or (45), has a dynamics generated by  $I\bar{L} + K^T$  rather than simply that for the trajectories  $\bar{L}$ . This is significant since the former has the additional compensation for the cooling and for the homogeneous perturbations of that cooling. Hence the approach to the time dependent normal state is determined only by the remaining non-hydrodynamic relaxation. The time scale for relaxation to the normal state is independent of the hydrodynamic time scales of that normal state. Quantitative verification of these concepts is another important future direction for research on a hydrodynamic description for granular fluids.

## Acknowledgments

The author is indebted to Professor J. Javier Brey and Dr. Aparna Baskaran of Syracuse University for their collaboration on closely related linear response methods for granular fluids.

## Appendix

**Gradient expansion** In this Appendix the Liouville equation in the form (34) is written to first order in the gradients and solved. Also the invariants of the associated dynamics are identified.

Consider first the right side of (34) which can be written equivalently as

$$\begin{aligned} & \int d\mathbf{r}' \frac{\delta\rho_{0\ell}}{\delta\langle a_\alpha(\mathbf{r}'); t \rangle} \{ \nabla \cdot \langle \mathbf{b}_\alpha(\mathbf{r}); t \rangle + \delta_{\alpha 2} \langle w(\mathbf{r}); t \rangle \} - \bar{L}\rho_{0\ell} \\ &= - \int d\mathbf{r}' \frac{\delta\rho_{0\ell}}{\delta\langle a_\alpha(\mathbf{r}'); t \rangle} \langle La_\alpha(\mathbf{r}'); t \rangle - \bar{L}\rho_{0\ell} \\ &= \int d\mathbf{r}' \frac{\delta\rho_{0\ell}}{\delta\langle a_\alpha(\mathbf{r}'); t \rangle} \int d\Gamma a_\alpha(\mathbf{r}) \bar{L}(\rho_{0\ell} + \Delta) - \bar{L}\rho_{0\ell}. \end{aligned} \quad (78)$$

The first equality follows from (17) and (21). The first two terms are determined by the local HCS which can be expanded to first order in the gradients

$$\begin{aligned} \rho_{0\ell} &= \rho_0 (\langle a_\alpha(\mathbf{r}); t \rangle) \\ &\quad + \int d\mathbf{r}' \left( \frac{\delta\rho_{0\ell}}{\delta\langle a_\alpha(\mathbf{r}'); t \rangle} \right)_{\delta\langle a_\alpha; t \rangle=0} \\ &\quad \cdot (\langle a_\alpha(\mathbf{r}'); t \rangle - \langle a_\alpha(\mathbf{r}); t \rangle) + \dots \\ &= \rho_0 (\langle a_\alpha(\mathbf{r}); t \rangle) \\ &\quad + \mathbf{m}_\beta(\mathbf{r}, \langle a_\alpha(\mathbf{r}); t \rangle) \cdot \nabla \langle a_\beta(\mathbf{r}); t \rangle + \dots \end{aligned} \quad (79)$$

The functional derivatives are

$$\begin{aligned} \left( \frac{\delta\rho_{0\ell}}{\delta\langle a_\alpha(\mathbf{r}'); t \rangle} \right)_{\delta\langle a_\alpha; t \rangle=0} &= \delta(\mathbf{r}' - \mathbf{r}) \left( \frac{\partial\rho_0(\langle a_\alpha(\mathbf{r}); t \rangle)}{\partial\langle a_\alpha(\mathbf{r}); t \rangle} \right. \\ &\quad \left. + \frac{\partial\mathbf{m}_\beta(\mathbf{r}, \langle a_\alpha(\mathbf{r}); t \rangle)}{\partial\langle a_\alpha(\mathbf{r}); t \rangle} \cdot \nabla \langle a_\beta(\mathbf{r}); t \rangle \right) \\ &+ \mathbf{m}_\alpha(\mathbf{r}, \langle a_\alpha(\mathbf{r}); t \rangle) \cdot \nabla \delta(\mathbf{r}' - \mathbf{r}) + \dots \end{aligned} \quad (80)$$

Here,

$$\begin{aligned} \mathbf{m}_\beta(\mathbf{r}, \langle a_\alpha(\mathbf{r}); t \rangle) \\ \equiv \int d\mathbf{r}' \left( \frac{\delta\rho_{0\ell}}{\delta\langle a_\beta(\mathbf{r}'); t \rangle} \right)_{\delta\langle a_\alpha; t \rangle=0} (\mathbf{r}' - \mathbf{r}), \end{aligned} \quad (81)$$

and  $\rho_0(\langle a_\alpha(\mathbf{r}); t \rangle)$  is the actual HCS with its global density, energy, and momentum evaluated at the common values  $\langle a_\alpha(\mathbf{r}); t \rangle$ . It follows from (32) that the averages of  $a_\alpha(\mathbf{r})$  for  $\rho$ ,  $\rho_{0\ell}$ , and  $\rho_0$  are all the same. This in turn gives

$$\int d\Gamma a_\alpha(\mathbf{r}) \mathbf{m}_\beta = 0 = \int d\Gamma a_\alpha(\mathbf{r}) \Delta. \quad (82)$$

With these results and the fact that  $\Delta$  is of first order in the gradients, (78) to first order in the gradients becomes

$$\begin{aligned} & \int d\mathbf{r}' \frac{\delta\rho_{0\ell}}{\delta\langle a_\alpha(\mathbf{r}'); t \rangle} \{ \nabla \cdot \langle \mathbf{b}_\alpha(\mathbf{r}); t \rangle + \delta_{\alpha 2} \langle w(\mathbf{r}); t \rangle \} - \bar{L}\rho_{0\ell} \\ & \rightarrow \bar{L}\rho_0 - (1 - \mathcal{P})(I\bar{L} + K^T)_{\alpha\beta} \mathbf{m}_\beta \cdot \nabla \langle a_\alpha; t \rangle + \mathcal{P}\bar{L}\Delta. \end{aligned} \quad (83)$$

The matrix  $K^T$  is the transpose of  $K$

$$K_{\alpha\beta} = \delta_{\alpha 2} \frac{\partial \omega(\langle a_\alpha(\mathbf{r}); t \rangle)}{\partial \langle a_\beta(\mathbf{r}); t \rangle}, \quad (84)$$

and  $I$  is the unit matrix. The generator  $\bar{\mathcal{L}}$  is the same as that of (12) with  $\omega \rightarrow \omega_0(\langle a_\alpha(\mathbf{r}); t \rangle)$  for the HCS evaluated at the common values  $\langle a_\alpha(\mathbf{r}); t \rangle$

$$\bar{\mathcal{L}} = -\omega_0(\langle a_\alpha(\mathbf{r}); t \rangle) \partial_{\langle a_\alpha(\mathbf{r}); t \rangle} + \bar{L}. \quad (85)$$

Finally,  $\mathcal{P}$  is the projection operator

$$\mathcal{P}X = \frac{\partial \rho_0}{\partial \langle a_\alpha(\mathbf{r}); t \rangle} \int d\Gamma a_\alpha(\mathbf{r}) X. \quad (86)$$

The first term of (83) vanishes by definition of the HCS,  $\rho_0$ , confirming that the right side of the Liouville Eq. (34) is of first order in the gradients.

At this point, the Liouville Eq. (34) becomes

$$\begin{aligned} \partial_t \Delta - \int d\mathbf{r}' \frac{\delta \Delta}{\delta \langle a_2(\mathbf{r}'); t \rangle} \omega_0(\langle a_\alpha(\mathbf{r}'); t \rangle) + \mathcal{P} \bar{L} \Delta \\ = (1 - \mathcal{P}) \boldsymbol{\Upsilon}'_\alpha \cdot \nabla \langle a_\alpha; t \rangle, \end{aligned} \quad (87)$$

$$\boldsymbol{\Upsilon}'_\alpha \equiv -(I \bar{\mathcal{L}} + K^T)_{\alpha\beta} \mathbf{m}_\beta. \quad (88)$$

This equation is still exact up through contributions of first order in the gradients. It has solutions of the form

$$\Delta(\Gamma, t | \langle a_\alpha(\mathbf{r}); t \rangle) = \mathbf{G}_v(\Gamma, t, \langle a_\alpha(\mathbf{r}); t \rangle) \cdot \nabla \langle a_v(\mathbf{r}); t \rangle, \quad (89)$$

Substitution into (87) gives the corresponding equation for  $\mathbf{G}_v$

$$\partial_t \mathbf{G}_v + (1 - \mathcal{P})(I \bar{\mathcal{L}} + K^T)_{\nu\beta} \mathbf{G}_\beta = (1 - \mathcal{P}) \boldsymbol{\Upsilon}'_v, \quad (90)$$

with the solution

$$\begin{aligned} \mathbf{G}_v(\Gamma, t, \langle a_\alpha(\mathbf{r}); t \rangle) \\ = \int_0^t dt' \left( e^{-(1-\mathcal{P})(I \bar{\mathcal{L}} + K^T)t'} \right)_{\nu\beta} (1 - \mathcal{P}) \boldsymbol{\Upsilon}'_\beta. \end{aligned} \quad (91)$$

It is possible to add to (87) an arbitrary solution to the homogeneous equation corresponding to (90). As described in the text, this represents the dynamics of the first stage of rapid velocity relaxation to the local HCS. The interest here is in the second stage where possible formation of a normal solution occurs. Hence, it is simpler to choose that stage for initial conditions (initial local HCS).

Define the derivatives of the HCS by

$$\Psi_\beta(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \equiv \frac{\partial \rho_0}{\partial \langle a_\beta(\mathbf{r}); t \rangle}. \quad (92)$$

Then differentiate the equation for  $\rho_0$

$$\frac{\partial}{\partial \langle a_\beta(\mathbf{r}); t \rangle} \bar{\mathcal{L}} \rho_0 = 0, \quad (93)$$

to get

$$(I \bar{\mathcal{L}}_T + K^T)_{\nu\beta} \Psi_\beta = 0. \quad (94)$$

Since  $(I \bar{\mathcal{L}}_T + K^T)$  is the generator for the dynamics in (91) this shows that  $\Psi_\beta$  are the invariants of that dynamics.

The projection operator  $\mathcal{P}$  in (95) acts only on phase functions with translational invariance. In that case (86) simplifies to

$$\mathcal{P}X = \Psi_\beta \int d\Gamma A_\beta X, \quad A_\beta = V^{-1} \int d\mathbf{r} a_\alpha(\mathbf{r}). \quad (95)$$

The first equality of (82) becomes  $\mathcal{P} \mathbf{m}_\beta = 0$ . This in turn gives

$$\begin{aligned} \mathbf{m}_\beta &= (1 - \mathcal{P}) \mathbf{m}_\beta = (1 - \mathcal{P}) \mathbf{M}_\beta, \\ \mathbf{M}_\beta &\equiv \int d\mathbf{r}' \left( \frac{\delta \rho_0 \ell}{\delta \langle a_\alpha(\mathbf{r}'); t \rangle} \right)_{\delta \langle a_\alpha; t \rangle = 0} \mathbf{r}'. \end{aligned} \quad (96)$$

Then  $(1 - \mathcal{P}) \boldsymbol{\Upsilon}'_\alpha$  simplifies to

$$\begin{aligned} (1 - \mathcal{P}) \boldsymbol{\Upsilon}'_\alpha &= -(1 - \mathcal{P})(I \bar{\mathcal{L}} + K^T)_{\alpha\beta} (1 - \mathcal{P}) \mathbf{M}_\beta \\ &\equiv (1 - \mathcal{P}) \boldsymbol{\Upsilon}_\alpha \end{aligned} \quad (97)$$

with

$$\boldsymbol{\Upsilon}_\alpha = -(I \bar{\mathcal{L}} + K^T)_{\alpha\beta} \mathbf{M}_\beta. \quad (98)$$

Use has been made of the identity

$$(1 - \mathcal{P})(I \bar{\mathcal{L}} + K^T) \mathcal{P} = 0. \quad (99)$$

This same identity leads to a simplification of the dynamics in (91)

$$e^{-(1-\mathcal{P})(I \bar{\mathcal{L}} + K^T)t'} (1 - \mathcal{P}) = (1 - \mathcal{P}) e^{-(I \bar{\mathcal{L}} + K^T)t'} (1 - \mathcal{P}). \quad (100)$$

In summary, the solution to the Liouville equation to first order in the gradients is

$$\begin{aligned} \rho(\Gamma, t | \langle a_\alpha; t \rangle) &= \rho_0(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &+ (1 - \mathcal{P}) (\mathbf{M}_\beta(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle)) \\ &+ \int_0^t dt' \left( e^{-(I \bar{\mathcal{L}} + K^T)t'} \right)_{\nu\beta} \\ &\cdot (1 - \mathcal{P}) \boldsymbol{\Upsilon}_\beta(\Gamma, \langle a_\alpha(\mathbf{r}); t \rangle) \\ &\cdot \nabla \langle a_v(\mathbf{r}); t \rangle. \end{aligned} \quad (101)$$

## Bibliography

### Primary Literature

1. Kadanoff LP (1999) Built upon sand: Theoretical ideas inspired by granular flows. *Rev Mod Phys* 71:435–444
2. Haff PK (1983) Grain flow as a fluid mechanical phenomenon. *J Fluid Mech* 134:401–430
3. McLennan JA (1989) Introduction to Nonequilibrium Statistical Mechanics. Prentice-Hall, New Jersey
4. Hansen J-P, McDonald I (1986) Theory of Simple Liquids. Elsevier Press, London
5. Resibois P, De Leener M (1977) Classical Kinetic Theory of Fluids. Wiley, New York
6. See, for instance, Goldhirsch I, Tan ML, Zanetti G (1993) A molecular dynamical study of granular fluids: the unforced granular gas. *J Sci Comput* 8:1–40; McNamara S, Young WR (1996) Dynamics of a freely evolving, two-dimensional granular medium. *Phys Rev E* 53:5089–5100; Deltour P, Barrat JL (1997) Quantitative study of a freely cooling granular medium. *J Phys I* 7:137–151
7. Brilliantov N, Pöschel T (2004) Kinetic Theory of Granular Gases. Oxford, New York
8. Dufty JW (2001) Kinetic theory and hydrodynamics for a low density gas. *Adv Complex Syst* 4:397–407. cond-mat/0109215.201
9. Brey JJ, Dufty JW, Santos A (1997) Dissipative dynamics for hard spheres. *J Stat Phys* 87:1051–1066
10. Van Noije TPC, Ernst MH (2001) Kinetic theory of granular gases. In: Pöschel T, Luding S (eds) *Granular Gases*. Springer, New York
11. Dufty JW, Brey JJ, Lutsko J (2002) Diffusion in a granular fluid. I. Theory. *Phys Rev E* 65:051303; Lutsko J, Dufty JW, Brey JJ (2002) Diffusion in a granular fluid. II. Simulation. *Phys Rev E* 65:051305; Dufty JW, Garzó V (2001) Mobility and diffusion in granular fluids. *J Stat Phys* 105:723–744
12. Dufty JW (2000) Statistical mechanics, kinetic theory, and hydrodynamics for rapid granular flow. *J Phys: Condens Matter* 12:A47–A56
13. Dufty JW, Baskaran A, Brey JJ (2006) Linear response for a granular fluid. *JSTAT* L08002:1–8
14. Dufty JW, Baskaran A, Brey JJ (2007) Linear response and hydrodynamics for granular fluids. *Phys Rev E* 77, 031310; Baskaran A, Dufty JW, Brey JJ (2007) Transport coefficients for the hard sphere granular fluid. *Phys Rev E* 77, 031311
15. See for example articles (2002) In: Halsey T, Metha A (eds) *Challenges in Granular Physics*. World Scientific, Singapore
16. Dufty J, Brey JJ (2005) Origins of hydrodynamics for a granular gas. In: Pareschi L, Russo G, Toscani G, (eds) *Modelling and Numerics of Kinetic Dissipative Systems* Nova Science, New York, pp 17–30, cond-mat/0410133
17. Here the Navier–Stokes approximation is defined by calculating the cooling rate, energy flux, and momentum flux to first order in the gradients. However, the fluxes occur under a gradient in the macroscopic balance equations while the cooling rate does not. Hence the equations themselves do not have all terms to second order in the gradients (i.e., the additional terms of second order contributing to the cooling rate).
18. Huan C, Yang X, Candela D, Mair RW, and Walsworth RL (2004) NMR experiments on a three-dimensional vibrofluidized granular medium. *Phys Rev E* 69:041302
19. Bizon C, Shattuck MD, Swift JB, Swinney HL (1999) Transport coefficients for granular media from molecular dynamics simulations. *Phys Rev E* 60:4340–4351; Rericha EC, Bizon C, Shattuck MD, Swinney HL (2001) Shocks in supersonic sand. *Phys Rev Lett* 88:014302
20. See, for instance, Brey JJ, Ruiz-Montero MJ, Cubero D (1999) On the validity of linear hydrodynamics for low-density granular flows described by the Boltzmann equation. *Europhys Lett* 48:359–364; Brey JJ, Ruiz-Montero MJ, Cubero D, García-Rojo R (2000) Self-diffusion in freely evolving granular gases. *Phys Fluids* 12:876–883; Garzó V, Montanero JM (2002) Transport coefficients of a heated granular gas. *Physica A* 313:336–356; Montanero JM, Santos A, Garzó V (2005) DSMC evaluation of the Navier–Stokes shear viscosity of a granular fluid. In: Capitelli M (ed) *Rarefied Gas Dynamics 24* (AIP Conf Proc, vol 72), pp 797–802
21. Brey JJ, Ruiz-Montero MJ, Moreno F, García-Rojo R (2002) Transversal inhomogeneities in dilute vibrofluidized granular fluids. *Phys Rev E* 65:061302; Brey JJ, Ruiz-Montero MJ, Moreno F (2001) Hydrodynamics of an open vibrated granular system. *Phys Rev E* 63:061305
22. Brey JJ, Ruiz-Montero MJ, Maynar P, Garzia de Soria MI (2005) Hydrodynamic modes, Green–Kubo relations, and velocity correlations in dilute granular gases. *J Phys Cond Mat* 17:S2489–S2502
23. Brey JJ, Dufty JW, Kim CS, Santos A (1998) Hydrodynamics for granular flow at low density. *Phys Rev E* 58:4638–4653; Sela N, Goldhirsch I (1998) Hydrodynamic equations for rapid flows of smooth inelastic spheres, to Burnett order. *J Fluid Mech* 361:41–74
24. Dufty JW, Brey JJ (2002) Green–Kubo expressions for a low density granular gas. *J Stat Phys* 109:433–448. cond-mat/0201361
25. Dufty JW (2005) Some aspects of the Boltzmann equation for granular gase. In: Capitelli M (ed) *Rarefied Gas Dynamics* (AIP Conf Proc 762, New York), pp 789–796
26. Dufty JW, Brey JJ (2005) Hydrodynamic modes for granular gases. *Phys Rev E* 68:030302; Brey JJ, Dufty JW (2005) Hydrodynamic modes for a granular gas from kinetic theory. *Phys Rev E* 72:011303
27. Dufty JW (2007) Fourier’s law for a granular fluid. *J Phys Chem B* 111:15605–15612
28. Helfand E (1960) Transport coefficients from dissipation in a canonical ensemble. *Phys Rev* 119:1–9
29. Bird R, Armstrong R, Hassager O (1977) *Dynamics of Polymeric Liquids*. Wiley, New York
30. Santos A, Garzó V, Dufty JW (2004) Inherent rheology of a granular fluid in uniform shear flow. *Phys Rev E* 69:061303. cond-mat/0309320
31. Hrenya C (2007) (private communication, and to be published)

### Books and Reviews

- Campbell CS (1990) Rapid granular flows. *Ann Rev Fluid Mech* 22:57–92
- Mehta A (ed) (1993) *Granular Matter, An Interdisciplinary Approach*. Springer, New York
- Jaeger HM, Nagel SR, Behringer RP (1996) *Granular solids, liquids, and gases*. *Rev Mod Phys* 68:1259–1273
- Duran J (2000) *Sands, powders, and grains: an introduction to the physics of granular materials*. Springer, New York

- Pöschel T, Luding S (eds) (2001) Granular Gases. Springer, New York
- Halsey T, Metha A (eds) (2002) Challenges in Granular Physics. World Scientific, Singapore
- Campbell CS (2002) Granular shear flows in the elastic limit. *J Fluid Mech* 465:261–291
- Pöschel T, Brilliantov N (eds) (2003) Granular Gases Dynamics. Springer, New York
- Goldhirsch I (2003) Rapid granular flows. *Annu Rev Fluid Mech* 35:267–293
- Brilliantov N, Pöschel T (2004) Kinetic Theory of Granular Gases. Oxford, New York
- Hinrichsen H, Wolf D (eds) (2004) The Physics of Granular Media. Wiley-VCH, Berlin
- Coniglio A, Fierro A, Herrmann H, Nicodemi M (eds) (2004) Unifying Concepts in Granular Media and Glasses. Elsevier, Amsterdam
- Pöschel T, Schwager T (2005) Computational Granular Dynamics: Models and Algorithms. Springer, New York
- Duft JW (2007) Nonequilibrium statistical mechanics and hydrodynamics for a granular fluid. Six lectures at the Second Warsaw School on Statistical Physics. Kazimierz, Poland. arXiv:0707.3714

quences of formulae, sets of clauses, or other sets built out of symbols of the alphabet, or their subsets. The expressions of formal languages, even if created with a specific meaning in mind, do not carry themselves any meaning, they are just finite sequences of certain symbols. The meaning is being assigned to them by establishing a proper semantics.

**Semantics** Semantics for a given symbolic language  $\mathcal{L}$  assigns a specific interpretation in some domain to all symbols and expressions of the language. It also involves related ideas such as truth and model. They are called semantical concepts to distinguish them from the syntactical ones.

**Model** The word *model* is used in many situations and has many meanings but they all reflect some parts, if not all, of its following formal meaning. A structure  $M$ , called also an interpretation, is a model for a set  $\mathcal{E}_0 \subseteq \mathcal{E}$  of expressions of  $\mathcal{L}$  if and only if every expression  $E \in \mathcal{E}_0$  it true in  $M$ .

## Granular Model for Data Mining

ANITA WASILEWSKA<sup>1</sup>, ERNESTINA MENASALVAS<sup>2</sup>

<sup>1</sup> Computer Science Department, Stony Brook University, Stony Brook, USA

<sup>2</sup> Departamento de Lenguajes y Sistemas Informáticos, Facultad de Informática, Madrid, Spain

### Article Outline

Glossary

Definition of the Subject

Introduction

Granular Model: Syntax and Semantics for Data Mining

Semantic Model

Descriptive Model

Granular Model Revisited: Satisfaction and Truth

Future Directions

Bibliography

### Glossary

**Data mining** Data mining is a process that includes the following phases: creating the target data, data preprocessing, data mining proper, pattern evaluation, and knowledge presentation.

**Syntax** Syntax, or syntactical concepts, refer to simple relations among symbols and expressions of formal, symbolic languages. A symbolic language is a pair  $\mathcal{L} = (\mathcal{A}, \mathcal{E})$ , where  $\mathcal{A}$  is an alphabet and  $\mathcal{E}$  is the set of expressions of  $\mathcal{L}$ . The expressions are often the formulae of the language, but can also represent sets of se-

### Definition of the Subject

We present here three abstract models: descriptive, semantic, and granular. All of them are abstract structures that allow us to formalize some general properties of the data mining process and address the semantics-syntax duality inherent to any data mining process. A descriptive model formalizes the syntactical concepts and properties of the process. A semantic model formalizes its semantical properties. Finally, they form components of the granular model in which we establish a relationship between the descriptive and semantic models and provide a formal semantics for syntactical expressions of the language  $\mathcal{L}$  of the descriptive model, thus justifying a use of the word model.

Moreover, we provide, within the granular model, a formal definition of data mining as the process of information generalization. The notion of generalization is defined in terms of granularity of steps of the process. In the model the data is represented in a form of knowledge systems. Each knowledge system has a granularity associated with it and the process changes, or not, its granularity. The notion of granularity is crucial for defining some notions and components of the model, hence the granular model name.

### Introduction

Data mining, as defined in 1996 by Piatetsky-Shapiro [13] is a step (crucial, but a step nevertheless) in a KDD (knowledge discovery in databases) process. The Piatetsky-Shapiro's definition states that the KDD process consists of the following steps: developing an understanding of

the application domain; creating a target data set; choosing the data mining task i. e. deciding whether the goal of the KDD process is classification, regression, clustering, etc ...; choosing data preprocessing algorithms; choosing data mining algorithm(s); interpreting mined patterns; deciding if a re-iteration is needed; consolidating the discovered knowledge.

Since then the data mining (DM) term has evolved to become a name for all of the KDD process, or some parts of it, or even to be used as a name of an application of a data mining or learning algorithm.

In 1997 the Cross-Industry Standard Process for Data Mining (CRISP-DM) was proposed [15] to establish a standard for what they called, and others adopted, a data mining process. CRISP-DM standard was developed for business purposes and it included all of KDD process steps plus some extra steps such as a business understanding, business goal understanding followed by the KDD standard steps. Hence the KDD process became often a data mining process for industrial applications and was and is more and more often called just by the name of data mining.

To clarify these naming confusions we follow the standard terminology developed by data mining research in which we understand by data mining a KDD process in which its original data mining phase is now called *data mining proper phase*. For short we say that

*Data mining (DM) is a process that includes between the others the following phases: creating the target data, data preprocessing, data mining proper, pattern evaluation, and knowledge presentation.*

We formalize, within our granular model, the intuitive notions of the data mining preprocessing and data mining proper processes (Sect. “[Data Mining Process](#)”).

One of the main goals of data mining is to *provide comprehensible descriptions* of information extracted from the data bases. The descriptions come in different forms. In the case of classification problems it might be a set of characteristic or discriminant rules, it might be a decision tree or a neural network with fixed set of weights. In the case of association analysis it is a set of associations (frequent itemsets), or association rules with accuracy parameters. In the case of cluster analysis it is a set of clusters, each of which has its own description and a cluster name. In the case of approximate classification by the rough set analysis it is usually a set of discriminant or characteristic rules (with or without accuracy parameters) or a set of decision tables.

If the number of descriptions is relatively small and all descriptions are of a reasonable length say that we obtained a generalized knowledge from the initial target

database; that we mined more comprehensive, more general information. If an algorithm would reduce, for example, 100,000 records of the database of size 500 (number of attributes of the initial database) to 20 descriptions of size 10, we surely would say that we well generalized our information. Once it is done, a natural question of a *measure of quality* of such syntactic generalization and hence the validity of our method arises and this is being handled by our semantic model (Sect. “[Semantic Model](#)”). The format of the descriptions is defined in our descriptive model (Sect. “[Descriptive Model](#)”). The correctness and quality of syntactical descriptions is defined in our granular model by a satisfaction relation. The satisfaction relation (Sects. “[K- Satisfaction and K- Truth](#)”, “[Granular Model Revisited: Satisfaction and Truth](#)”) establishes a connection between syntactical and semantical approaches, and hence describe formally the syntax-semantic duality of the data mining process.

### **Granular Model: Syntax and Semantics for Data Mining**

We usually view data mining results and present them to the user in their descriptive, i. e. syntactic form as it is the most natural form of communication. But the data mining process is deeply semantical in its nature. We hence build our granular model on two levels: syntactic and semantic. The syntactic level is represented by a descriptive model, and semantic level is described by the semantic model. The semantics-syntax duality of data mining process is expressed in our granular model by the satisfiability relation.

Moreover, we use a granular model to provide a formal definition of data mining as the process of information generalization (Sect. “[Data Mining as Generalization](#)”). In the model the data preprocessing and data mining algorithms are defined as certain operators that act on data represented in a form of knowledge systems (Definition 9). Each knowledge system has a granularity associated with it (Definition 10) and the operators change, or not, its granularity. The notion of granularity is crucial for defining all notions and components of the model, hence the granular model name.

### **Granular Model Definition**

For a given data mining application one defines in detail all the components of the model (see Examples 7, 14). We provide here, as a part of the granular model definition, a general form of such definitions.

**Definition 1** A **granular model** is a system **GM** = (**SM**, **DM**,  $\models$ ) where:

- $SM$  is a semantic model;
- $DM$  is a descriptive model;
- $\models \subseteq \mathcal{P}(U) \times \mathcal{E}$  is called a satisfaction relation, where  $U$  is the universe of  $SM$  and  $\mathcal{E}$  is the set of descriptions defined by the  $DM$ .

The satisfaction relation  $\models$  (Definition 31) establishes the relationship between expressions of the semantic and descriptive models. It hence established formally the syntax-semantics duality of the data mining process. The models are  $SM$  and  $DM$  are defined in Definitions 13 and 21, respectively.

### Motivation and Examples

The semantic model is the most important component of our granular model. The intuitions behind the definition of the semantic model are as follows. When we perform the data mining procedures the first step in any of them is to drop the key attribute. This step allows us to introduce similarities in the database as records do not have anymore their unique identification. The input into the data mining process is hence always a data table obtained from the target data by removal of the key attribute. We call it a *target data table*.

As the next step we represent, following the rough set model, our target data table as Pawlak's information system [11] with the universe  $U$  by adding a new, non attribute column for the record names, i. e. objects of  $U$ . We take this set  $U$  as the universe of our model of  $SM$ .

Data mining, as it is commonly said, is a process of generalization. In order to model this process we have first to define what does it mean from semantical point of view that one stage of the process is more general than the other. The idea behind is very simple. It is the same as saying that  $(a + b)^2 = a^2 + 2ab + b^2$  is a more general formula than the formula  $(2 + 3)^2 = 2^2 + 2 \cdot 2 \cdot 3 + 3^2$ . This means that one description (formula) is more general than the other if it describes more objects. From a semantical point of view it means that data mining process consists of putting objects (records) in sets of objects. From a syntactical point of view the data mining process consists of building descriptions (in terms of attribute, values of attributes pairs) of these sets of objects, with some extra parameters, if needed.

To model a situation that allows us to talk about descriptions of sets of records we extend the notion of Pawlak's model of an information system to our notion of a *knowledge system* (Definition 9). The universe the knowledge system consists of some subsets of  $U$ , i. e. elements of  $\mathcal{P}(U)$ . For example a target data table (after preprocessing)

and the corresponding representation by Pawlak's information system, and a knowledge system with universe  $U$  of granularity one are as follows.

### Example 2

Target Data Table  $T_0$

$a_1$	$a_2$	$a_3$
small	small	medium
medium	small	medium
small	small	medium
big	small	small
medium	medium	big
small	small	medium
big	small	small
medium	medium	big
small	small	medium
big	small	medium
medium	medium	small
small	small	medium
big	small	big
medium	medium	small

Target Information System  $I_0$

$U$	$a_1$	$a_2$	$a_3$
$x_1$	small	small	medium
$x_2$	medium	small	medium
$x_3$	small	small	medium
$x_4$	big	small	small
$x_5$	medium	medium	big
$x_6$	small	small	medium
$x_7$	big	small	small
$x_8$	medium	medium	big
$x_9$	small	small	medium
$x_{10}$	big	small	medium
$x_{11}$	medium	medium	small
$x_{12}$	small	small	medium
$x_{13}$	big	small	big
$x_{14}$	medium	medium	small

A *knowledge system* corresponding to target table  $T_0$  is a system that all its objects are one element sets  $\{x\}$  for corresponding  $x \in U$  of  $T_0$ . We call such knowledge systems (Definition 11) a target knowledge system based on the target data table  $T_0$ . The knowledge systems with the property that all its objects are one element sets are called systems of *granularity one*.

### Example 3

Target Knowledge System  $K_0$

$\mathcal{P}^1(U)$	$a_1$	$a_2$	$a_3$
$\{x_1\}$	small	small	medium
$\{x_2\}$	medium	small	medium
$\{x_3\}$	small	small	medium
$\{x_4\}$	big	small	small
$\{x_5\}$	medium	medium	big
$\{x_6\}$	small	small	medium
$\{x_7\}$	big	small	small
$\{x_8\}$	medium	medium	big
$\{x_9\}$	small	small	medium
$\{x_{10}\}$	big	small	medium
$\{x_{11}\}$	medium	medium	small
$\{x_{12}\}$	small	small	medium
$\{x_{13}\}$	big	small	big
$\{x_{14}\}$	medium	medium	small

Assume now that we have applied some algorithm  $ALG_1$  and it has returned a following set  $\mathcal{D} = \{D_1, D_2, \dots, D_7\}$  of descriptions (in the granular model,  $D_i \in \mathcal{E}$  and  $\mathcal{E}$  is defined in the descriptive model). We write  $s$  for the attribute value *small*,  $m$  for *medium* and  $b$  for *big*.

$$\begin{aligned} D_1 &: (a_1 = s) \cap (a_2 = s) \cap (a_3 = m), \\ D_2 &: (a_1 = m) \cap (a_2 = s) \cap (a_3 = m), \\ D_3 &: (a_1 = m) \cap (a_2 = m) \cap (a_3 = b), \\ D_4 &: (a_1 = m) \cap (a_2 = m) \cap (a_3 = s), \\ D_5 &: (a_1 = b) \cap (a_2 = s) \cap (a_3 = s), \\ D_6 &: (a_1 = b) \cap (a_2 = s) \cap (a_3 = m), \\ D_7 &: (a_1 = b) \cap (a_2 = s) \cap (a_3 = b). \end{aligned}$$

Now, a natural question arises: how well this set of descriptions describes our original data i.e. how accurate is the algorithm  $ALG_1$  we have used to find them, how accurate is the *knowledge* we have thus obtained out of our data. To answer this, we use the target information system with the universe  $U$ , and for any  $D \in \mathcal{D}$  we examine a set  $S(D) = \{x \in U : D\}$  called a truth set for  $D$ . We define it formally in the Definition 27. Intuitively, the sets  $S(D) = \{x \in U : D\}$  contain all records (i.e. their identifiers) with the same description given in terms of attribute values of attribute pairs. The descriptions do not need to utilize all attributes of the target data, as it is often the case, and one of ultimate goals of data mining is to find descriptions with as few attributes as possible.

In association analysis the descriptions can represent the frequent itemsets. For example, for a frequent three itemset  $D = i_1 i_2 i_3$ , the truth set  $S(D)$  represents all transactions that contain items  $i_1, i_2, i_3$ .

Descriptions come in different forms, depending on the data mining goal and application. We define formally

a general form of descriptions as a part of the descriptive model (Definition 21).

For the data presented in the Examples 2, 3 and their descriptions  $D_i \in \mathcal{D}$  the sets  $S(D_i)$  are as follows.

$$\begin{aligned} S_1 &= S(D_1) = \{x \in U : D_1\} = \{x_1, x_3, x_6, x_9, x_{12}\}, \\ S_2 &= S(D_2) = \{x \in U : D_2\} = \{x_2\}, \\ S_3 &= S(D_3) = \{x \in U : D_3\} = \{x_5, x_8\}, \\ S_4 &= S(D_4) = \{x \in U : D_4\} = \{x_{11}, x_{14}\}, \\ S_5 &= S(D_5) = \{x \in U : D_5\} = \{x_4, x_7\}, \\ S_6 &= S(D_6) = \{x \in U : D_6\} = \{x_{10}\}, \\ S_7 &= S(D_7) = \{x \in U : D_7\} = \{x_{13}\}. \end{aligned}$$

We represent our results in a form of a *knowledge system* as follows. We write, as before,  $s$  for the attribute value *small*,  $m$  for *medium* and  $b$  for *big* and we put the names of proper subsets of  $U$  in the second representation of  $K_1$ .

#### Example 4

Resulting Knowledge System  $K_1$

$K_1(U)$	$a_1$	$a_2$	$a_3$
$\{x_1, x_3, x_6, x_9, x_{12}\}$	$s$	$s$	$m$
$\{x_2\}$	$m$	$s$	$m$
$\{x_5, x_8\}$	$m$	$m$	$b$
$\{x_{11}, x_{14}\}$	$m$	$m$	$s$
$\{x_4, x_7\}$	$b$	$s$	$s$
$\{x_{10}\}$	$b$	$s$	$s$
$\{x_{13}\}$	$b$	$s$	$b$

Resulting Knowledge System  $K_1$

$K_1(U)$	$a_1$	$a_2$	$a_3$
$S_1$	$s$	$s$	$m$
$S_2$	$m$	$s$	$m$
$S_3$	$m$	$m$	$b$
$S_4$	$m$	$m$	$s$
$S_5$	$b$	$s$	$s$
$S_6$	$b$	$s$	$s$
$S_7$	$b$	$s$	$b$

The representation of data mining results in a form of a knowledge system allows us to define how good is the *knowledge* obtained by a given algorithm. In our case the knowledge obtained describes 100% of our target data as

$$S_1 \cup S_2 \cup S_3 \cup \dots \cup S_7 = \{x_1, x_2, \dots, x_{14}\} = U.$$

Observe that the sets  $S_1, \dots, S_7$  are also disjoint and non-empty, i.e. they form a partition of the universe  $U$ . In such a case the knowledge system (and the algorithm) will be called *exact* (see Definition 12).

Moreover, we can see that the resulting system  $K_1$  is *more general* than the input data, as represented by the target system  $K_0$  because its *granularity* (i. e. maximum of cardinality of its granules, i. e. elements of its universe) is higher the granularity of  $K_0$ . This observation motivates the formal Definition 15. The granularity of our  $K_0$  is one, as is the granularity of all target knowledge systems.

The *granularity* of  $K_1$  is  $\max\{|S_1|, \dots, |S_7|\} = \max\{5, 1, 2, \dots\} = 5$ .

Let now assume that we have applied to our target data  $T$  (represented by  $K_0$ ) another algorithm  $ALG_2$  and it returned two descriptions  $D_1, D_2$  under a condition that we need only descriptions of length 2, such that the corresponding attribute-value pairs appear in the target data with frequency  $\geq 30\%$ .

*Example 5* Consider the following two descriptions  $D_1, D_2$ .

- $D_1 : (a_1 = s) \cap (a_2 = s)$ ,  $D_2 : (a_2 = s) \cap (a_3 = m)$ .
- Now we evaluate:  $S_1 = S(D_1) = \{x_1, x_3, x_6, x_9, x_{12}\}$ ,
- $S_2 = S(D_2) = \{x_1, x_2, x_3, x_6, x_9, x_{10}, x_{12}\}$ .

The description  $D_1$  fulfills the algorithm condition because it's of length 2 and  $\frac{|S_1|}{|U|} = \frac{5}{14} \sim 0.3508$ . The frequency% is greater than 30%.

The description  $D_2$  fulfills the algorithm condition because it's of length 2 and  $\frac{|S_2|}{|U|} = \frac{7}{14} = 0.5$ . The frequency% is hence greater than 30%.

Incorporating the parameters of # of attributes in the description and their frequency imposed by the  $ALG_2$  into our knowledge system we obtain the following table. The table is incomplete, as there are more descriptions fulfilling the algorithm conditions.

*Example 6*  
Knowledge System  $K_2$

$K_2(U)$	$a_1$	$a_2$	$a_3$	# of attr	frequency%
$S_1$	s	s	-	2	35
$S_2$	-	s	m	2	50

The sets  $S_1, S_2$  do not form a partition of the universe  $U$  as  $S_1 \cap S_2 \neq \emptyset$  and moreover,  $S_1 \cup S_2 \neq U$ . The algorithm  $ALG_2$  is hence *not exact*. It describes only 57% of the target data and what is described is described following the frequency conditions. Of course  $K_2$  is more general than  $K_0$  and the algorithm  $ALG_2$  generalized the target data, even if in an incomplete way. Now we form a new set of descriptions  $D_3, D_4$  associated with the resulting knowledge system  $K_2$ .

### Example 7

- $D_3 : (a_1 = s) \cap (a_2 = s) \cap (\# \text{of attr} = 2) \cap (\text{frequency\%} = 36)$ ,
- $D_4 : (a_2 = s) \cap (a_3 = m) \cap (\# \text{of attr} = 2) \cap (\text{frequency\%} = 50)$ .

We re-write  $D_3, D_4$  as follows.  $D_3 : D_1 \cap (\# \text{of attr} = 2) \cap (\text{frequency\%} = 36)$ ,  $D_4 : D_2 \cap (\# \text{of attr} = 2) \cap (\text{frequency\%} = 50)$ , where  $D_1, D_2$  are descriptions considered in Example 5.

We say that the set  $S_1 = S(D_1) = \{x \in U : D_1\}$  is a *truth set* for the description  $D_3$  under restrictions defined by the descriptions  $(\# \text{of attr} = 2) \cap (\text{frequency\%} = 36)$ . The set  $S_2 = S(D_2)$  is the truth set for  $D_4$  under restrictions  $(\# \text{of attr} = 2) \cap (\text{frequency\%} = 50)$ .

### Semantic Model

The definition of a semantic model presented here is a simpler and more comprehensible version of [17,18,21,22]. The initial investigations of the subject appeared also in [3,8,9,10]. The notion of a knowledge system is central to the definition and development of the semantic model, as we have seen in Sect. “Motivation and Examples”. We define it formally in the next section. We next use the knowledge system to formalize a notion of a granule and granularity (Definition 10) that is central to our granular model.

### Knowledge Systems and Granularity

Observe that the table depicting the knowledge system  $K_2$  (Example 6) is incomplete, i. e. not all attributes have assigned values. This fact in the formal definition will be represented by a *partial function*  $g$ . Also  $K_2$  contains some new *attributes describing properties of its granules*. In the formal definition we will call these new attributes the *knowledge attributes*.

The formal definitions of information system, knowledge and target knowledge systems, and their granularity and exactness are as follows.

**Definition 8** A Pawlak’s **information system** is a system  $I = (U, A, V_A, f)$ , where  $U \neq \emptyset$  is called a set of **objects**;  $A \neq \emptyset, V_A \neq \emptyset$  are called the set of **attributes** and **values of attributes**, respectively;  $f : U \times A \rightarrow V_A$  is called an **information function**.

**Definition 9** A **knowledge system** based on  $I = (U, A, V_A, f)$  is a system

$$K = (K(U), A, E, V_A, V_E, g)$$

where:

- The **universe**  $K(U)$  of  $K$  is a subset of the set  $\mathcal{P}(U)$  of all subsets of the universe  $U$  of  $I$ , i.e.  $K(U) \subseteq \mathcal{P}(U)$ ;
- $E$  is a finite set of **knowledge attributes** ( $k$ -attributes) such that  $A \cap E = \emptyset$ ;
- $V_E$  is a finite set of **values of k- attributes**;
- $g$  is a partial function called a **knowledge function** ( $k$ -function);
- $g : \mathcal{P} \times (A \cup E) \rightarrow (V_A \cup V_E)$  is such that:
  - (i)  $g | (\bigcup_{x \in U} \{x\} \times A) = f$ ;
  - (ii)  $\forall S \in \mathcal{P}, \forall a \in A((S, a) \in \text{dom}(g) \Rightarrow g(S, a) \in V_A)$ ;
  - (iii)  $\forall S \in \mathcal{P}, \forall e \in E((S, e) \in \text{dom}(g) \Rightarrow g(S, e) \in V_E)$ ;

We use the above notion of a knowledge system to define the granules of the universe and the granularity of the system, an hence later, the granularity of the data mining process.

**Definition 10** Any set  $S \in \mathcal{P}(U)$  i.e.  $S \subseteq U$  is called a **granule** of  $U$ . The cardinality  $|S|$  of  $S$  is called a **granularity** of  $S$ . The set

$$Gr_K = \{S \in \mathcal{P} : \exists b \in (E \cup A)((S, b) \in \text{dom}(g))\}$$

is called a **granule universe** of  $K$ . A number  $gr_K = \max\{|S| : S \in Gr_K\}$  is called a **granularity** of  $K$ .

Observe that by conditions (ii), (iii) of Definition 9,  $Gr_K = K(U)$  and it justifies its name. The condition (i) of Definition 9 says that when  $E = \emptyset$ , the  $k$ -function  $g$  is total on the set  $\{\{x\} : x \in U\} \times A$  and  $\forall x \in U \forall a \in A(g(\{x\}, a) = f(x, a))$ . We denote  $\mathcal{P}^1(U) = \{\{x\} : x \in U\}$ .

**Definition 11** Let  $I = (U, A, V_A, f)$ . Any system  $K^1 = K^1 = (\mathcal{P}^1(U), A, \emptyset, V_A, \emptyset, g) = (\mathcal{P}^1(U), A, V_A, g) = (\mathcal{P}^1(U), A, V_A, g)$  is called a **target knowledge system** based on  $I$ .

Observe that any target knowledge system has granularity one. Finally, we define the *exactness* of a knowledge system as follows.

**Definition 12** A knowledge system  $K = (K(U), A, E, V_A, V_E, g)$  is called **exact** if all its granules  $Gr_K$  form a partition of the universe  $U$ .

The system  $K_2$  from Example 4 is exact, the system  $K_3$  from Example 6 is not exact.

In our model we view data mining algorithms as certain *operators*. For example, our  $ALG_1$  is represented in the semantic model by an operator  $p_1$  acting on some subset

of a set  $\mathcal{K}$  of knowledge systems, such that  $p_1(K_0) = K_1$ .  $ALG_2$  is represented in the model by an operator  $p_2$  also acting on some (may be different) subset of the set  $\mathcal{K}$  of knowledge systems, such that  $p_2(K_0) = K_2$ . We put all the above observations into a formal notion of a semantic model.

**Definition 13** A **semantic model** is a system  $SM = (\mathcal{P}(U), \mathcal{K}, \mathcal{G})$ , where:

- $U \neq \emptyset$  is the **universe**;
- $\mathcal{K} \neq \emptyset$  is a set of knowledge systems, called also **data mining process states**;
- $\mathcal{G} \neq \emptyset$  is the set of **operators**;
- Each operator  $p \in \mathcal{G}$  is a partial function on the set of all data mining process states, i.e.  $p : \mathcal{K} \rightarrow \mathcal{K}$ .

The semantic model is always being built for a given application. The target data is represented first in a form the target information system with the universe  $U$ , as in Example 2 and then in the form of target knowledge system as in the Example 3.

The semantic model based on our Examples 2, 3, and 4 are as follows.

**Example 14**  $SM_1 = (\mathcal{P}(U), \mathcal{K}, \mathcal{G})$ , where:  $U = \{x_1, x_2, \dots, x_{14}\}$ ;  $\mathcal{K} = \{K_0, K_1, K_2\}$ ;  $\mathcal{G} = \{p_1, p_2\}$ ; Each  $p_i \in \mathcal{G}, (i = 1, 2)$  is a partial function  $p_i : \mathcal{K}_1 \rightarrow \mathcal{K}_1$ , such that  $p_1(K_0) = K_1, p_2(K_0) = K_2$ .

### Data Mining as Generalization

In order to model within our semantic model data mining as a process of generalization we first introduce the following definition of *generalization relation* based on a notion of granularity.

**Definition 15** A relation  $\preceq \subseteq \mathcal{K} \times \mathcal{K}$  is called a **generalization relation** if the following condition holds for any  $K, K' \in \mathcal{K}$ .

$$K \preceq K' \text{ if and only if } gr_K \leq gr_{K'},$$

where  $gr_K$  denotes the granularity of  $K$  (Definition 10).

Observe that for  $K_0, K_1, K_2$  from Example 14,  $gr_{K_0} = 1 \leq 5 = gr_{K_1} \leq 7 = gr_{K_2}$ , and the system  $K_2$  is the most general. But at the same time  $K_1$  is exact and  $K_2$  is not exact, so we have a trade off between exactness and generality.

As the next step we use the generalization relation to define the notion of a generalization operator as follows.

**Definition 16** An operator  $g \in \mathcal{G}$  is called a **generalization operator** if for any  $K, K' \in \mathcal{K}$  such that  $g(K) = K'$ , we have that  $K \preceq K'$ .

Observe that both operators  $p_1, p_2$  in Example 14 are generalization operators.

### Data Mining Operators

In the data mining process the preprocessing and data mining are disjoint, inclusive/exlusive categories. The pre-processing is an integral and very important stage of the data mining process and needs as careful analysis as the data mining itself. Our framework allows us distinguish two disjoint classes of operators: the preprocessing operators  $\mathcal{G}_{\text{prep}}$  and data mining proper operators  $\mathcal{G}_{\text{dm}}$  and we put

$$\mathcal{G} = \mathcal{G}_{\text{prep}} \cup \mathcal{G}_{\text{dm}}.$$

The paper [20] contains detailed formal definitions, their motivation, and discussion of these two classes. The model presented in Example 14 didn't include the preprocessing stage; it used the data mining proper operators only.

The main idea behind the concept of the operator is to capture not only the fact that data mining techniques generalize the data, but also to categorize existing methods. We want to do it in as exclusive/inclusive sense as possible. It means that we want to make sure that our categorization will be to distinguish as it should, for example clustering from classification, making at the same time all classification methods fall into one category, called the classification operator while all clustering methods would fall into the category called the clustering operator. The third category is the association analysis described in our framework by the association operator. We don't include in our analysis purely statistical methods like regression, etc ... This gives us only three data mining classes of operators to consider: classification  $\mathcal{G}_{\text{class}}$ , clustering  $\mathcal{G}_{\text{clust}}$ , and association  $\mathcal{G}_{\text{assoc}}$ . The motivation, discussion and formal definition of these classes is included in the paper [22]

**Theorem 17** *Let  $\mathcal{G}_{\text{class}}$ ,  $\mathcal{G}_{\text{clust}}$  and  $\mathcal{G}_{\text{assoc}}$  be the sets of all classification, clustering, and association operators, respectively. The following conditions hold.*

- (1)  $\mathcal{G}_{\text{class}} \neq \mathcal{G}_{\text{clust}} \neq \mathcal{G}_{\text{assoc}}$
- (2)  $\mathcal{G}_{\text{assoc}} \cap \mathcal{G}_{\text{class}} = \emptyset$ ,
- (3)  $\mathcal{G}_{\text{assoc}} \cap \mathcal{G}_{\text{clust}} = \emptyset$ .

### Data Mining Process

We model here two stages of the data mining process: pre-processing and data mining proper, as we were able to distinguish the data preprocessing and data mining proper operators as disjoint categories. We adopt the following definitions.

**Definition 18** *Any sequence  $K_1, K_2, \dots, K_n$  ( $n \geq 1$ ) of data mining states is called a **data preprocessing process**, if there is a preprocessing operator  $G \in \mathcal{G}_{\text{prep}}$ , such that  $G(K_i) = K_{i+1}$ ,  $i = 1, 2, \dots, n - 1$ .*

**Definition 19** *Any sequence  $K_1, K_2, \dots, K_n$  ( $n \geq 1$ ) of data mining states is called a **data mining proper process**, if there is a data mining proper operator  $G \in \mathcal{G}_{\text{dm}}$ , such that  $G(K_i) = K_{i+1}$ ,  $i = 1, 2, \dots, n - 1$ .*

The data mining process consists of the preprocessing process (that might be empty) and the data mining proper process. We know that the sets  $\mathcal{G}_{\text{prep}}$  and  $\mathcal{G}_{\text{dm}}$  are disjoint. This justifies the following definition.

**Definition 20** *A **data mining process** is any sequence  $K_1, K_2, \dots, K_n$  ( $n \geq 1$ ) of data mining states, such that  $K_1, \dots, K_i$  ( $0 \leq i \leq n$ ) is a preprocessing process and  $K_{i+1}, \dots, K_n$  is a data mining proper process.*

Observe that in the semantic model  $SM_1$  (Example 14) we have only two processes:  $K_0, K_1$  and  $K_0, K_2$ . Both of them are data mining proper processes.

### Descriptive Model

Given a semantic model  $SM = (\mathcal{P}(U), \mathcal{K}, \mathcal{G})$  We associate with it its descriptive counterpart defined below.

**Definition 21** *A **descriptive model** is a system  $DM = (\mathcal{L}, \mathcal{E}, DK)$  where:*

- $\mathcal{L} = (\mathcal{A}, \mathcal{E})$  is called a **descriptive language**;
- $\mathcal{A}$  is a countably infinite set called the **alphabet**;
- $\mathcal{E} \neq \emptyset$  and  $\mathcal{E} \subseteq \mathcal{A}^*$  is the set of **descriptive expressions** of  $\mathcal{L}$ ;
- $DK \neq \emptyset$  and  $DK \subseteq \mathcal{P}(\mathcal{E})$  is a set of **descriptions of knowledge states**.

As in the case of a semantic model, we build the descriptive model for a given application. We define here only a general form of the model. We assume however, that whatever the application is, the descriptions are always built in terms of attributes and values of the attributes, some logical connectives, some predicates and some extra parameters, if needed. For example, a neural network with its nodes and weights can be seen as a formal description (in an appropriate descriptive language), and the knowledge states would represent changes in parameters during the neural network training process, or a final, converged network.

The model we build here is a model for what we call a **descriptive data mining**, i. e. the data mining for which the goal of the data mining process is to produce a set

of descriptions in a language easily comprehensible to the user. For that reason we identify, for example, a decision tree constructed by the classification by the decision tree algorithm with the set of discriminant rules obtained from the tree.

In our Examples 2–14 we have used descriptions in the form  $(a = v)$  to denote that the attribute  $a$  has a value  $v$ , but one might also use, like it is often done a predicate form  $a(v)$  or  $a(x, v)$  instead.

We define the components of  $\mathcal{DM}$  in the following stages.

**Stage 1** For each  $K \in \mathcal{K}$ , we define (Sect. “[K-Descriptive Language](#)”) its own descriptive language  $\mathcal{L}_K = (\mathcal{A}_K, \mathcal{E}_K)$ .

**Stage 2** For each  $K \in \mathcal{K}$ , and descriptive expression  $F \in \mathcal{E}_K$ , we define what does it mean that  $D$  satisfied in  $K$ ; i. e. we define (Definition 25) a satisfaction relation  $\models_K$ .

**Stage 3** For each  $K \in \mathcal{K}$ , and descriptive expression  $F \in \mathcal{E}_K$ , we define what does it mean that  $D$  is true  $K$ , i. e.  $\models_K D$  (Definition 28).

**Stage 4** For each  $K \in \mathcal{K}$ , we define its set  $\mathcal{D}_K \subseteq \mathcal{P}(\mathcal{E}_K)$  of descriptions of its own knowledge (Definition 26).

**Stage 5** We use the languages  $\mathcal{L}_K$  to define the descriptive language  $\mathcal{L}$  (Definition 29).

**Stage 6** We use the descriptive expressions  $\mathcal{E}_K$  of  $\mathcal{L}_K$  to define the set  $\mathcal{E}$  of descriptive expressions of  $\mathcal{L}$  (Definition 30).

**Later Stage** We use the satisfaction relations  $\models_K$  to define the satisfaction relation  $\models$  of our granular model **GM** (Definitions 1 and 31, respectively).

### K-Descriptive Language

As Stage 1 of our construction of the descriptive model, we define for each  $K \in \mathcal{K}$ , its own description language  $\mathcal{L}_K$ . The language depends on the semantic model and the goal of the data mining process. As we have said, the descriptions produced by data mining algorithms come in different forms. We build here a model for a **descriptive data mining**, i. e. we assume that the descriptions are built from attributes and values of attributes and two logical connectives: conjunction (see descriptions associated with Examples 2–14) and implication. The implication connective is needed to model the different kind of rules that are being mined by data mining algorithms: discriminant and characteristic rules in classification analysis; association rules by association analysis; or other rules obtained by hybrid systems.

**Definition 22** For any  $K \in \mathcal{K}$ , of  $\text{SM}, K = (\mathcal{P}(U), A, E, V_A, V_E, g)$ , we define the descriptive language of

$K, \mathcal{L}_K$  as

$$\mathcal{L}_K = (\mathcal{A}_K, \mathcal{E}_K),$$

where  $\mathcal{A}_K$  is called an **alphabet**,  $\mathcal{E}_K$  the set of **descriptive expressions** such that

$$\mathcal{E}_K = \mathcal{D}_K \cup \mathcal{F}_K,$$

for  $\mathcal{D}_K$  the set of **descriptive formulae** and  $\mathcal{F}_K$  the set of **formulae** of  $\mathcal{L}_K$  (Definition 23).

The alphabet  $\mathcal{A}_K = \text{VAR}_K \cup \{\cap, \Rightarrow\} \cup \{(., )\}$ , where  $\{\cap, \Rightarrow\}$  is the set of logical connectives of  $\mathcal{L}_K$ . The set  $\text{VAR}_K$  of variables of  $K$  is also called its set of **atomic descriptions**. We put  $\text{VAR}_K = \mathcal{D}_A \cup \mathcal{D}_E$ , where  $\mathcal{D}_A = \{(a = v) : a \in A, v \in V_A\}, \mathcal{D}_E = \{(a = v) : a \in E, v \in V_E\}$ .

Atomic descriptions, i. e. the elements of  $\text{VAR}_K = \mathcal{D}_A \cup \mathcal{D}_E$  represent minimal blocks of semantical description. Elements of  $\mathcal{D}_A$  are atomic descriptions of minimal blocks built with use of the attributes of the initial target database. Elements of  $\mathcal{D}_E$  are atomic descriptions of minimal blocks built with use of knowledge attributes used (if any) during the process of data mining. We use them to define the sets of descriptions and formulae as follows.

**Definition 23** The set  $\mathcal{D}_K$  of all **descriptive formulae** of  $\mathcal{L}_K$  is the set

$$\mathcal{D}_K = \mathcal{AD}_K \cup \mathcal{ED}_K \cup \mathcal{F}_K,$$

where  $\mathcal{AD}_K, \mathcal{ED}_K, \mathcal{F}_K$  are defined as below.

The set  $\mathcal{AD}_K \subseteq \mathcal{A}_K^*$  is called the set of **data attribute descriptions** and is the smallest set such that the following conditions hold.

- (1)  $\mathcal{D}_A \subseteq \mathcal{AD}_K$  (data attribute **atomic** description);
- (2) If  $D_1, D_2 \in \mathcal{AD}_K$ , then  $D_1 \cap D_2 \in \mathcal{AD}_K$ .

The set  $\mathcal{ED}_K \subseteq \mathcal{A}_K^*$  is called the set of **knowledge attribute descriptions** and is the smallest set such that the following conditions hold.

- (1)  $\mathcal{D}_E \subseteq \mathcal{ED}_K$  (knowledge attribute **atomic** description);
- (2) If  $D_1, D_2 \in \mathcal{ED}_K$ , then  $D_1 \cap D_2 \in \mathcal{ED}_K$ .

We distinguish two categories of formulae: one  $\mathcal{F}_A$  describes a certain knowledge, the other,  $\mathcal{F}_E$ , the uncertain, or approximate, knowledge. The certain knowledge is expressed in our model in terms of attributes and values of attributes of the initial, target data only. The description

of the approximate knowledge includes the extra parameters, describing properties of granules as defined by the knowledge attributes of  $K$  (Definition 9).

**Definition 24** The set  $\mathcal{F}_K$  of **formulae** of  $\mathcal{L}_K$  is a union of two sets of formulae;  $\mathcal{F}_A, \mathcal{F}_E$ , i. e.

$$\mathcal{F}_K = \mathcal{F}_A \cup \mathcal{F}_E,$$

where  $\mathcal{F}_A = A\mathcal{D}_K \cup A\mathcal{R}_K$  and  $\mathcal{F}_E = E\mathcal{D}_K \cup E\mathcal{R}_K$  are defined below.

The set  $A\mathcal{D}_K = \{D_1 \cap D_2 : D_1 \in A\mathcal{D}_K, D_2 \in E\mathcal{D}_K\}$  is the set of **knowledge description formulae**, with knowledge attribute descriptions depicting uncertainty measures. The set  $A\mathcal{R}_K = \{(D_1 \Rightarrow D_2) : D_1, D_2 \in \mathcal{D}_A\}$  is the set of **attribute rule-formulae** depicting certain rules obtained during the data mining process. The set  $E\mathcal{R}_K = \{(D_1 \Rightarrow D_2) \cap D_3 : D_1, D_2 \in \mathcal{D}_A, D_3 \in \mathcal{D}_E\}$  is the set of **knowledge rule-formulae** depicting rules with uncertainty measures described obtained by descriptions from  $\mathcal{D}_E$ .

The formulae of the language  $\mathcal{L}_K$  are **not** yet the data mining **rules**. They only describe a syntactical form of the rules appropriate for a given application. Formulae from  $\mathcal{F}_K$  become **rules determined by  $K$**  only when they do relate semantically to  $K$ , i. e. reflect the properties of our initial target database. In this case we say that they are **true**, or **true under the measures** described by descriptions from  $\mathcal{D}_E$ . We define, in the next section the notion of truthfulness, as we always do, via notion of satisfiability.

### K- Satisfaction and K- Truth

In this section we define (Definition 25) a satisfaction relation  $\models_K \subseteq \mathcal{P}(U) \times \mathcal{E}_K$  that establishes the relationship between descriptive expressions of the  $\mathcal{L}_K$  and what they semantically represent in  $K$ . For any  $(S, F) \in \models_K$  we say that  $S$  **satisfies  $F$  in  $K$**  and write it symbolically as

$$S \models_K F.$$

We call our satisfaction relation a **k-satisfaction**, as it represents a satisfaction relation relative to the system  $K$ . As the next step we define the notion of K-truth (Definition 28) i. e. we define what does it mean that a descriptive expression  $F$  is **true in  $K$** , symbolically expressed as

$$\models_K F.$$

Let  $K = (K(U), A, E, V_A, V_E, g)$  and let  $\mathcal{L}_K = (\mathcal{A}_K, \mathcal{E}_K)$  be the description language defined by  $K$  (Definition 22), where  $\mathcal{E}_K = \mathcal{D}_K \cup \mathcal{F}_K$ .

**Definition 25** A **k-satisfaction relation**  $\models_K \subseteq \mathcal{P}(U) \times \mathcal{E}_K$  is defined by induction over the level of complexity of any descriptive expression  $F \in \mathcal{E}_K$  as follows.

1.  $F \in \mathcal{D}_K$ .
  - (i) Let  $(a = v) \in \mathcal{D}_A$  be an **atomic** attribute description. We define, for any  $S \in \mathcal{P}(U)$ , for any  $a \in A, v \in V_A$ ,  $S \models_K (a = v)$  if and only if  $(S, a) \in \text{dom}(g)$  and  $g(S, a) = v$ ;
  - (ii) Let  $(e = v) \in \mathcal{D}_E$  be an **atomic** knowledge attribute description. We define, for any  $S \in \mathcal{P}(U)$ , for any  $e \in E, v \in V_E$ ,  $S \models_K (e = v)$  if and only if  $(S, e) \in \text{dom}(g)$  and  $g(S, e) = v$ ;
2.  $F \in \mathcal{F}_K$ .
 

We extend  $\models_K$  to the set  $\mathcal{F}_A = A\mathcal{D}_K \cup A\mathcal{R}_K$

  - (iii) Let  $D \in A\mathcal{D}_K$  and  $D = D_1 \cap \dots \cap D_n$ . We define, for any  $S \in \mathcal{P}(U)$ ,  $S \models_K D$  if and only if  $\forall (1 \leq i \leq n) (S \models_K D_i)$ ;
  - (iv) Let  $(D_1 \Rightarrow D_2) \in A\mathcal{R}_K$ , i. e.  $D_1, D_2 \in A\mathcal{D}_K$ . We define, for any  $S \in \mathcal{P}(U)$ ,  $S \models_K (D_1 \Rightarrow D_2)$  if and only if  $S \models_K D_1$  and  $S \models_K D_2$ ;
  - (v) Let  $(D_1 \Rightarrow D_2) \cap D_3 \in E\mathcal{R}_K$ , i. e.  $(D_1 \Rightarrow D_2) \in A\mathcal{R}_K$  and  $D_3 \in \mathcal{D}_E$ . We define, for any  $S \in \mathcal{P}(U)$ ,  $S \models_K (D_1 \Rightarrow D_2) \cap D_3$  if and only if  $S \models_K (D_1 \Rightarrow D_2)$  and  $S \models_K D_3$ .

This ends Stage 2 of the definition of the descriptive model and we are ready for Stage 3, i. e. to define the set of all descriptions of knowledge states.

**Definition 26** Let  $K \in \mathcal{K}$ , the set  $\mathcal{D}_K \subseteq \mathcal{P}(\mathcal{E}_K)$  of **descriptions of knowledge states** of  $K$  is defined as follows

$$\mathcal{D}_K = \{F \in \mathcal{E}_K : \exists S \in \text{Gr}_K (S \models_K F)\}.$$

Now we are ready to define a notion of ***F true in  $K$*** , for any formula  $F \in \mathcal{F}_K$ , symbolically expressed by  $\models_K F$ . This notion relates the satisfaction relation  $\models_K$ , i. e. satisfiability of formulae in the system  $K$  with the initial target data as out ultimate point of reference. This connection is being established by the notion of the ***truth set*** defined below.

**Definition 27** For any attribute data description  $D \in A\mathcal{D}_K \cup E\mathcal{D}_K$  the set  $S(D) = \{x \in U : D\}$  is called a **truth set** for  $D$ .

We define  $\models_K F$  by induction over the level of complexity of any descriptive expression  $F \in \mathcal{F}_K$  as follows.

### Definition 28

1.  $F \in A\mathcal{D}_K$ .
  - (i) For any attribute data description  $D \in A\mathcal{D}_K$ , we define:  $\models_K D$  if and only if  $S(D) \models_K D$ .

- (ii) For any attribute description formula  $D \in A\mathcal{D}_K$ , we put  $\models_K D$  if and only if  $S(D) \models_K D$ .
- 2.  $F \in \mathcal{F}_A$ .
  - (iii) For an attribute rule-formula  $(D_1 \Rightarrow D_2) \in A\mathcal{R}_K$ , we define  $\models_K (D_1 \Rightarrow D_2)$  if and only if  $S(D_1) \models_K D_1, S(D_1) \models_K D_2$ , and  $S(D_1) \subseteq S(D_2)$  or  $(D_1 \cap S(D_2)) \neq \emptyset$ .
- 3.  $F \in AED_K$ .
  - (iv) For any knowledge attribute description  $D_1 \cap D_2 \in AED_K$ ,  $D_1 \in A\mathcal{D}_K$ ,  $D_2 \in E\mathcal{D}_K$  we define:  $\models_K D_1 \cap D_2$  if and only if  $S(D_1) \cap S(D_2) \models_K D_1$ .
- 4.  $F \in \mathcal{F}_E$ .
  - (v) For an attribute rule-formula  $(D_1 \Rightarrow D_2) \cap D_3 \in AER_K$ , we define  $\models_K (D_1 \Rightarrow D_2) \cap D_3$  if and only if  $S(D_1) \cap S(D_3) \models_K D_1, S(D_1) \cap S(D_3) \models_K D_2$ , and  $S(D_1) \cap S(D_3) \subseteq S(D_2) \cap S(D_3)$  or  $(D_1 \cap S(D_2) \cap S(D_3)) \neq \emptyset$ .

In the case when  $F \in \mathcal{F}_E$  and  $\models_K F$  we say that  $F$  is **true in  $K$  under the measures  $D$** , for a certain  $D \in AED_K$ .

If  $\models_K (D_1 \Rightarrow D_2)$  and the condition  $S(D_1) \subseteq S(D_2)$  holds, then we call the formula  $(D_1 \Rightarrow D_2)$  **C- true in  $K$ , or C- true under the measures  $D$** , if  $\models_K (D_1 \Rightarrow D_2) \cap D$ .

If  $\models_K (D_1 \Rightarrow D_2)$  and the condition  $(D_1) \cap S(D_2) \neq \emptyset$  holds, then we call the formula  $(D_1 \Rightarrow D_2)$  **D- true in  $K$ , or D - true under the measures  $D$** , if  $\models_K (D_1 \Rightarrow D_2) \cap D$ .

The notion of D-truth reflects the semantics needed to define the discriminant rules for the classification analysis, and C-truth is needed for the characteristic rules, hence the names. As we have said before, the knowledge attributes from the set  $E$  describe uncertainty measures for the granules  $S \in K(U)$ . The formulae that incorporate the knowledge attribute descriptions  $D \in E\mathcal{D}_K$  can be only *true* in  $K$  under some uncertainty measures (Example 5).

### Descriptive Language and Descriptive Model

We have already constructed (Sects. “[Descriptive Language and Descriptive Model](#)”) all sub-components of the Definition 21 of the descriptive model  $\mathcal{DM}$  and now we proceed to complete it as follows.

**Definition 29** We define the language of  $\mathcal{DM}$  as  $\mathcal{L} = (\mathcal{A}, \mathcal{E})$ , where  $\mathcal{A} = \bigcup \{\mathcal{A}_K : K \in \mathcal{K}\}$ , and  $\mathcal{E} = \bigcup \{\mathcal{E}_K : K \in \mathcal{K}\}$ .

The set  $\mathcal{DK}$  of all descriptions of knowledge states of the semantic model  $\mathcal{SM} = (\mathcal{P}(U), \mathcal{K}, \mathcal{G})$  is the following.

**Definition 30** Let  $\mathcal{DK}$  be the set of descriptions of knowledge states of  $K$  (Definition 26). The set  $\mathcal{DK} = \bigcup \{\mathcal{D}_K : K \in \mathcal{K}\}$  is a set of **descriptions of knowledge states of  $\mathcal{SM}$** .

This completes the definition of the descriptive model as a system  $\mathcal{DM} = (\mathcal{L}, \mathcal{E}, \mathcal{DK})$  where:  $\mathcal{L} = (\mathcal{A}, \mathcal{E})$  is a descriptive language;  $\mathcal{A}$  an alphabet and  $\mathcal{E}$  is the set of descriptive expressions (Definition 29);  $\mathcal{DK}$  is a set of descriptions of knowledge states (Definition 30).

### Granular Model Revisited: Satisfaction and Truth

The granular model is defined (Sect. “[Granular Model: Syntax and Semantics for Data Mining](#)”) as a system  $\mathbf{GM} = (\mathbf{SM}, \mathcal{DM}, \models)$  where:  $\mathbf{SM}$  is a semantic model as defined in the Definition 13;  $\mathcal{DM}$  is a descriptive model (Definition 21); and  $\models \subseteq \mathcal{P}(U) \times \mathcal{E}$  is a satisfaction relation. All the components of  $\mathbf{GM}$  except the satisfaction relation have already been defined. As the last step we define the satisfaction relation and the notion of truth in  $\mathbf{GM}$  as follows.

**Definition 31** For any  $S \in \mathcal{P}(U)$  and for any  $F \in \mathcal{E}$ ,  $S \models F$  if and only if  $\exists K \in \mathcal{K}(S \models_K F)$

**Definition 32** We say that  $F \in \mathcal{E}$  is **true in  $\mathbf{GM}$**  (symbolically  $\models F$ ) if and only if  $\exists K \in \mathcal{K}(\models_K F)$ .

### Future Directions

The models presented here set a general framework for future foundational investigations. They can be carried on three levels. One, the most general, would deal with further developments within the granular model. The second, a more specific one, would deal with applications of the methods and language developed within the granular model to specific domains. On this level one would build semantic, descriptive, and granular models for different data mining domains; i. e., build and examine specific models for classification, clustering, and for association analysis. Finally, the most specific third level would deal with building models for basic descriptive data mining algorithms: decision trees and rough sets classification, association and classification by association, and clustering. The general methodology of analysis of the data mining process developed for the granular model can hence serve as unifying language in which different data mining domains and algorithms can be studied, discussed, and compared.

### Bibliography

1. Chapman P (2000) CRISP-DM 1.0 – Step-by-step Data Mining CRISP-DM Consortium: 1.0
2. Greco S, Matarazzo B, Slowinski R, Stefanowski J (2002) Importance and interaction of conditions in decision rules. In: Proceedings of Third International Conference RSCTC'02, Malvern,

- USA. Lecture Notes in Artificial Intelligence, vol 2475. Springer, Berlin, pp 255–262
3. Hadjimichael M, Wasilewska A (1998) A hierarchical model for information generalization. In: Proceedings of the 4th Joint Conference on Information Sciences, Rough Sets, Data Mining and Granular Computing (RSMDGrC'98) North Carolina, vol II, pp 306–309
  4. Han J, Kamber M (2000) Data mining: Concepts and techniques. Morgan Kauffman
  5. Inuiguchi M, Tanino T (2003) Classification versus approximation oriented generalization of rough sets. *Bull Int Rough Set Soc* 7(1/2):19–25
  6. Komorowski J (2002) Modelling biological phenomena with rough sets. In: Proceedings of Third International Conference RSCTC'02, Malvern, USA. Lecture Notes in Artificial Intelligence, vol 2575. Springer, Berlin, p 13
  7. Lin TY (2002) Database mining on derived attributes. In: Proceedings of Third International Conference RSCTC'02, Malvern, USA. Lecture Notes in Artificial Intelligence, vol 2575. Springer, Berlin, pp 14–32
  8. Martínez FJ, Menasalvas E, Wasilewska A, Fernández C, Hadjimichael M (2002) Extension of relational management system with data mining capabilities. In: Proceedings of Third International Conference RSCTC'02, Malvern, USA. Lecture Notes in Artificial Intelligence, vol 2475. Springer, Berlin, pp 421–424
  9. Menasalvas E, Wasilewska E, Fernández C (2001) The lattice structure of the KDD process: Mathematical expression of the model and its operators. *Fundamenta Informaticae: Special Issue Int J Info Syst* 48–62
  10. Menasalvas E, Wasilewska A, Fernández C, Martínez FJ (2002) Data mining – a semantical model. In: Proceedings of 2002 World Congress on Computational Intelligence, Honolulu, pp 435–441
  11. Pawlak Z (1981) Information systems. *Theor Found Inf Syst* 6:205–218
  12. Pawlak Z (1991) Rough sets – theoretical aspects reasoning about data. Kluwer, Norwell
  13. Fayyad, Piatetsky-Shapiro, Smyth (1996) From data mining to knowledge discovery: An overview. In: Fayyad, Piatetsky-Shapiro, Smyth, Uthurusamy (eds) *Advances in Knowledge Discovery and Data Mining*. AAAI Press, Menlo Park, pp 1–34
  14. Polkowski Lech (2002) Rough sets – mathematical foundations. Physica, Heidelberg
  15. Shearer C (2000) The CRISP-DM model: The new blueprint for data mining. *J Data Warehous* 5(4):13–22
  16. Skowron A (1993) Data filtration: a rough set approach. In: Proceedings of rough sets, fuzzy sets and knowledge discovery. RSKD, pp 108–118
  17. Wasilewska A, Ruiz EM, Fernández-Baizan MC (1998) A model for RSMD implementation. 1st International Conference on Rough Sets and Current Trends in Computing (RSCTC'98), 22–26 June, Warsaw, pp 186–193
  18. Wasilewska A, Menasalvas E (2004) Data preprocessing and data mining as generalization process. In: Proceedings of ICDM'04, The Fourth IEEE International Conference on Data Mining, Brighton, 1–4 Nov, pp 133–137
  19. Wasilewska A, Menasalvas E (2004) Data mining operators. In: Proceedings of ICDM'04, The Fourth IEEE International Conference on Data Mining, Brighton, pp 209–214
  20. Wasilewska A, Menasalvas E, Schaffr C (2005) Uniform Model for Data Mining. In: Proceedings of FDM05 (Foundations of Data Mining). In: ICDM2005, Fifth IEEE International Conference on Data Mining. Austin, Texas, pp 19–27
  21. Wasilewska A, Ruiz EM (2006) Data mining as generalization: A formal model. In: Lin TY, Ohsuga S, Liu CJ, Hu X (eds) *Foundations and Novel Approaches in Data Mining. Studies in Computational Intelligence*, vol 9. Springer, Berlin, pp 99–126
  22. Wasilewska A, Ruiz EM (2008) Data preprocessing and data mining as generalization. In: Lin TY, Xie Y, Wasilewska A, Liu CJ (eds) *Data Mining: Foundations and Practice. Studies in Computational Intelligence*. Springer, pp 453–468
  23. Wasilewska A, Ruiz EM (2005) A classification model: Syntax and semantics for classification. In: Proceedings of 10th International RSFDGrC2005 Conference, Regina, Canada, August/September 2005. Lecture Notes in Artificial Intelligence, vol 2. Springer, Berlin, pp 59–68
  24. Ziarko W, Fei X (2002) VPRSM approach to WEB searching. In: Proceedings of Third International RSCTC'02 Conference, Malvern, USA, October 2002. Lecture Notes in Artificial Intelligence, vol 2475. Springer, Berlin, pp 514–522
  25. Ziarko W (1993) Variable precision rough set model. *J Comput Syst Sci* 46(1):39–59
  26. Yao JT, Yao YY (2002) Induction of classification rules by granular computing. In: Proceedings of Third International RSCTC'02 Conference, Malvern, USA, October 2002. Lecture Notes in Artificial Intelligence, vol 2475. Springer, Berlin, pp 331–338

## Granular Neural Network

YAN-QING ZHANG  
 Department of Computer Science,  
 Georgia State University, Atlanta, USA

### Article Outline

- [Glossary](#)
- [Definition of the Subject](#)
- [Introduction](#)
- [Basic Architecture](#)
- [Granular Learning Algorithms](#)
- [Applications in Bioinformatics](#)
- [Applications in Computational Web Intelligence](#)
- [Applications in Brain Informatics](#)
- [Conclusions](#)
- [Future Directions](#)
- [Bibliography](#)

### Glossary

**Granular data** Granular data include various data granules such as classes, clusters, subsets, groups, linguistic values and intervals.

**Granular neuron** An artificial neuron maps granular data inputs to granular data outputs.

**Granular link** A granular link connects two different granular neurons.

**Granular weights** A granular weight represents connection strength between two granular neurons by using a granular value that is not limited to a traditional numerical value.

**Granular neural network** An intelligent neural network consists of granular neurons and granular links that connect relevant granular neurons.

### Definition of the Subject

Granular computing is a general computation theory for effectively using granules such as classes, clusters, subsets, groups and intervals to build an efficient computational model for complex applications with huge amounts of data, information and knowledge. Relevant existing techniques are divide and conquer, cluster analysis, fuzzy sets, rough sets, Yin-Yang computing, neutrosophic computing, interval computing, quotient space theory, belief functions, machine learning, databases, and so on.

A biological neural network in the human brain consists of a huge number of biological neurons that are networked to process various data sets (numerical and non-numerical). Different complex data sets such as images, linguistic terms, numbers, patterns and sounds can be recognized by biological neural networks. In other words, biological neural networks can process granules such as data clusters, linguistic words, information classes, image groups from human sensor networks (i.e., eyes, ears, nose, etc.).

Based on granular computing and biological neural networks, a granular neural network is designed to deal with numerical-linguistic data fusion and granular knowledge discovery in numerical-linguistic databases. From a data granulation point of view, the granular neural network can process granular data in a database. From a data fusion point of view, the granular neural network makes decisions based on different kinds of granular data. From a knowledge discovery point of view, the granular neural network is able to learn internal granular relations between numerical-linguistic inputs and outputs, and predict new relations in a database. The granular neural network is also capable of greatly compressing low-level granular data to high-level granular knowledge with some compression error and a data compression rate.

### Introduction

Similar to a biological neural network in the human brain, a traditional artificial neural network consists of lots of artificial neurons that are connected by links. Each artificial neuron is simply a nonlinear mathematical mapping function, and each link has a numerical weight value showing

connection strength between two neurons. The learning algorithm uses training data sets to optimize these weights of an artificial neural network, and the trained neural network can be used in prediction and decision-making applications. Traditional artificial neural networks based on the basic principle of biological neural networks in the human brain have been widely used in a lot of applications. However, traditional artificial neural networks use numerical values like 3.14 and -23.5. Internal weights are also presented by numerical values that are not meaningful for users (i.e., the black box problem). Generally, the traditional artificial neural network has two long-term problems: (1) the black box problem: Learned knowledge in a neural network is represented by non-meaningful numerical weights (i.e., people cannot understand these numerical values because they are not linguistic values, therefore people cannot understand how the neural network makes a decision), and (2) the curse of dimensionality: When a number of inputs and outputs is increased, the number of weights is increased dramatically (even exponentially). So the traditional artificial neural network cannot be used for large-scale applications with a large number of inputs and outputs.

Granular neural networks are used to process linguistic data like fuzzy terms to do granular knowledge discovery [20]. Granular neural networks with different structures can process information granules such as fuzzy sets and rough sets by using training algorithms [10]. The granular neural network uses linguistic weights and the rules using linguistic arithmetic to speed up learning process [1]. Granular neural networks are used to real applications such as land use classification [14]. In summary, the granular neural learning algorithm can discover meaningful granular rules that can be understood by ordinary people. So the granular neural network is useful to tackle the black box problem.

A set of training data is given to a computational model, and then the problem is how the network gleans useful knowledge from these training data sets in order to use that discovered knowledge to make right decisions for real applications. The two challenging problems for both computer science and cognitive science are the black box problem (i.e., the uninterpretability of a network of numerical weights or connection strengths) and the curse of dimensionality (i.e., the intractability complex cognitive problems for linear neural networks). This proposal plans to design a new architecture of a granular neural network with a new learning algorithm, and then use data from both cognitive science and computer science to test and to improve the new learning algorithm in terms of the two challenging problems.

In recent years, hybrid neural networks such as fuzzy neural network, granular neural networks and genetic neural networks have been investigated to solve complex application problems with high dimensionality. However, a single-stage hybrid neural network suffers from curse of dimensionality. A single-stage hybrid neural network suffers from curse of dimensionality because the number of parameters of the single-stage hybrid neural network is increasing exponentially with the increasing of the number of inputs. The hierarchical network structures based on multistage fuzzy reasoning have been used to solve the dimensionality problem. The ANFIS (Adaptive-Network-Based Inference System), based on TSK fuzzy reasoning, is useful in many applications. Because it still cannot extract the commonly used fuzzy rules with both fuzzy IF part and fuzzy THEN part, it makes some difficulties in acquiring fuzzy knowledge from experts in a natural manner and learning commonly used fuzzy rules from data.

Therefore, how to design a powerful granular neural network with high learning speed, low training error and low prediction error is still a challenging problem of artificial intelligence, computational intelligence and cognitive science.

### Basic Architecture

There are different types of granular neural networks. Now two granular neural networks are introduced in the two following sections, respectively.

#### First Architecture of Granular Neural Network

A database may contain numerical values, linguistic words, images, sounds, music pieces, and texts. The lowest data granulation technique deals with raw multimedia data collected directly from a real world environment, whereas a higher data granulation technique classifies raw multimedia data into higher-level granules (i.e., classes, clusters, categories, groups, sets, etc.) to simplify data processing and data mining. For high-level data granulation, basic processing elements are these granules (not low-level raw data) [20]. In this sense, data granulation is related to data mining, data fusion, information fusion and knowledge discovery. Interestingly, the human brain has a very strong ability to process multimedia granules, compute with linguistic words, and discover useful information from multimedia data. There should be links between raw multimedia data in an outside real world and biological neural networks inside a human brain. Similarly, links should be established between multimedia granules in databases and inputs-outputs of artificial neural networks. Since an artificial neural network cannot directly process

these multimedia granules in most cases (because it is designed for directly processing numerical data), a conversion or interpretation of a granule, a link between a multimedia database and a neural network, becomes crucial to designing a granular neural network. In other words, how to convert multimedia granules into corresponding numerical features is very important for granular neural networks. For example, the linguistic data feature extraction system can convert fuzzy linguistic data such as very high, too small and almost 100 into typical numerical features [20].

Generally speaking, a granular neural network is capable of processing various granular data (granules). Granules could be a class of numbers, a cluster of images, a set of concepts, a group of objects, a category of data, etc. These granules are inputs and outputs as multimedia data are inputs and outputs of biological neural networks in the human brain. Therefore, a granular-data-based granular neural network is more useful and more effective to process multimedia granules than a conventional numerical-data-based neural network. Here, the granular neural network is made by Fuzzy Neural Networks with Knowledge Discovery [18].

Generally speaking, a granular neural network is capable of processing various granular data (granules) [20]. Granules could be a class of numbers, a cluster of images, a set of concepts, a group of objects, a category of data, etc. These granules are inputs and outputs of granular neural networks as multimedia data are inputs and outputs of biological neural networks in the human brain.

Therefore, granular-data-based granular neural networks are more useful and more effective to process multimedia granules than conventional numerical-data-based neural networks. Here, the FNN is a powerful Fuzzy Neural Network with Knowledge Discovery (FNNKD). The functions of different layers are described layer by layer as follows [20,21]:

**Layer 1: Linguistic Feature Extraction Layer** In this layer, numerical-linguistic  $X$  and  $Y$  are transformed to corresponding fuzzy feature vectors  $(a1, b1, c1, d1)$  and  $(a2, b2, c2, d2)$ , respectively.

**Layer 2: Multi-FNNKD Layer** This layer consists of 4 dedicated FNNKDs (i.e., FNNKD1, FNNKD2, FNNKD3, and FNNKD4). FNNKD1 is a  $2 \times k_1 \times 1$  fuzzy neural network which generates a crisp center  $a$  of an output fuzzy set by using  $k_1$  fuzzy rules. FNNKD2 is a  $2 \times k_2 \times 1$  fuzzy neural network which generates a crisp width  $b$  of an output fuzzy set by using  $k_2$  fuzzy rules. FNNKD3 is a  $2 \times k_3 \times 1$  fuzzy neural network which generates  $c$  of

an output fuzzy set by using  $k_3$  fuzzy rules. FNNKD4 is a  $2 \times k_4 \times 1$  fuzzy neural network which generates  $d$  of an output fuzzy set by using  $k_4$  fuzzy rules.

An  $n$ -input-1-output normal fuzzy system has  $m$  fuzzy IF-THEN rules which are described by:

IF  $x_1$  is  $A_1^k$  and ... and  $x_n$  is  $A_n^k$  THEN  $y$  is  $B^k$ , where  $x_i$  and  $y$  are input and output fuzzy linguistic variables, respectively.

### Layer 3: Output Layer

**Case 1:** A fuzzy linguistic output is  $Z$  represented by a fuzzy feature vector  $(a, b, c, d)$ .

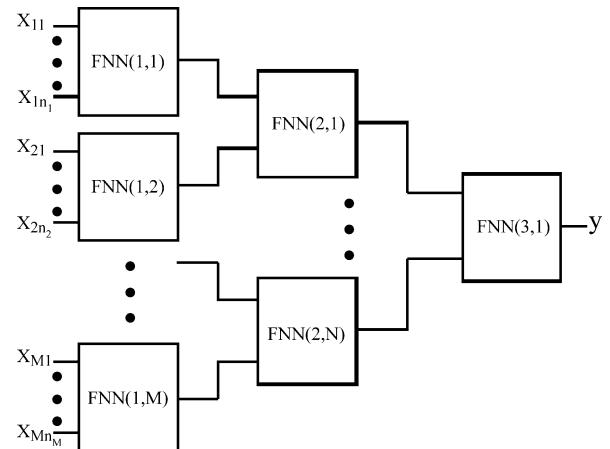
**Case 2:** A crisp numerical output is  $a$ .

For simplicity, the detailed learning algorithm is given in [19]. Once the learning procedure has been completed, all parameters for a FNNKD have been adjusted and optimized. As a result, all  $m$  fuzzy rules have been discovered from training data. Finally, the trained FNNKD can generate new values for new given input data.

### Second Architecture of Granular Neural Network

Here, we plan to use traditional artificial intelligence, new computational intelligence including fuzzy logic, neural networks, evolutionary computation, granular computing and cognitive science to investigate the new evolutionary granular cognitive neural network that can discover hidden decision-making knowledge from data (i.e., know how the human brain learns new knowledge and makes a decision). The traditional neural-learning algorithms can generate a lot of numerical weights that are not meaningful (i.e., the black box problem). Now we plan to design basic neuron granules that can contain meaningful knowledge like rules, and make new granular learning methods based on cognitive science and computational intelligence [5].

A new evolutionary granular cognitive neural network is proposed based on the normal fuzzy reasoning. The hybrid evolutionary granular cognitive neural learning algorithm using the divide-and-conquer strategy is developed to enhance learning quality in term of discovered knowledge, training error and prediction error. Simulations have shown that the evolutionary granular cognitive neural network is an effective data mining and knowledge discovery system which discovers meaningful fuzzy rules, has low training error, and generates low prediction error. There are two main problems of curse of dimensionality for a soft computing system: (1) structural dimensionality (i.e., the total number of soft rules increases exponentially with the number of input variables in single-stage reasoning process, and (2) parametric dimensionality (i.e., the total number of adjustable system pa-



Granular Neural Network, Figure 1

Architecture of a 3-stage granular neural network [17]

rameters increases exponentially with the number of input variables). For example, if a fuzzy system has  $n$  linguistic variables, each variable has  $m$  fuzzy linguistic sets, and each fuzzy linguistic set has  $k$  parameters, then the fuzzy system will have  $m^n$  fuzzy rules and  $(km)^n$  parameters. In this case, the structural complexity and the parametric complexity of the fuzzy system are  $O(m^n)$  and  $O((km)^n)$ , respectively. To solve the problems of curse of dimensionality, different hybrid soft computing architectures such as the multistage fuzzy neural network and the hierarchical system have been proposed to reduce system complexity in terms of dimensionality, time and space. In general, the key technology for solving the curse of dimensionality of a hybrid soft computing system (particularly Fuzzy Neural Network (FNN)) is the divide-and-conquer method. In other words, an original single-stage FNN with high system complexity can be divided into small multi-stage soft computing building blocks with low system complexity so as to reduce both structural dimensionality and parametric dimensionality significantly.

In general, the  $k$ th-stage FNN uses outputs generated from relevant  $k - 1$ th-stage FNNs as inputs, and then generates an output which will become an input of relevant  $k + 1$ th-stage FNNs. A final stage FNN will generate a final output. For example, Fig. 1 shows the 3-stage GCFNN. The FNN( $i, j$ ) is the  $i$ th-stage- $j$ thFNN. So FNN(1,  $j$ ) uses  $x_{1n_j}$  as inputs for  $j = 1, 2, \dots, M$ , similarly FNN(2,  $l$ ) for  $l = 1, 2, \dots, N$  uses outputs generated from relevant 1st-stage FNNs as inputs, and finally FNN(3,1) generates the final output  $y$ . Therefore, the final output  $y$  is a function of all original inputs  $x_{j1}, x_{j2}, \dots, x_{jn_j}$  for  $j = 1, 2, \dots, M$ .

Recently, hybrid neural networks such as fuzzy neural networks and granular neural networks have been inves-

tigated to solve complex application problems with high dimensionality. Such hybrid neural networks provide remarkable abilities to derive knowledge from complicated or imprecise datasets. Therefore, they have already been applied to extract patterns and detect trends that are too complex to be noticed by either human beings or other computer techniques. In this case, the trained neural networks can be thought of as an “expert” in the category of information it has been given to analyze.

In general, the key technology for solving the curse of dimensionality (including structural dimensionality and parametric dimensionality) of a hybrid soft computing system, such as a Fuzzy Neural Network (FNN), is the divide-and-conquer method. In other words, an original single-stage FNN with high system complexity can be divided into small multi-stage soft computing building blocks with low system complexity so as to reduce both structural dimensionality and parametric dimensionality significantly.

In addition, GGCNN provides parameters for deciding the number of iterations and termination criteria in order to control its learning performance. The hybrid genetic forward-wave-backward-wave learning algorithm is developed to enhance the system’s learning quality. Our simulation results generated by genetic granular cognitive fuzzy neural network are compared to human persons and are analyzed in terms of computer science and cognitive science.

### Granular Learning Algorithms

The traditional neural learning algorithms can generate a lot of numerical weights that are not meaningful (i.e., the black box problem). We plan to use traditional artificial intelligence, cognitive science and new computational intelligence including fuzzy logic, neural networks, evolutionary computation and granular computing to investigate the new evolutionary neural cognitive learning algorithms that can discover hidden decision-making knowledge from data (i.e., know how the human brain extracts rule-like relational knowledge from the inputs of experience and uses this knowledge in problem-solving and decision-making).

We plan to use traditional artificial intelligence, cognitive science and computational intelligence to investigate the new granular cognitive neural network architecture with cognitive neural granules with a small number of meaningful linguistic parameters. Each cognitive neural granule consists of a number of artificial cognitive neurons with nonlinear numerical/linguistic mapping functions. So cognitive neural granules are basic building blocks for the new granular cognitive neural network to effectively

solve the problem of curse of dimensionality. The effective topological structure of the granular cognitive neural network will be investigated.

Suppose that the GCFNN has  $K$  inputs and one output. Given input data vectors  $X^p$  (i.e.,  $X^p = (X_1^p, X_2^p, \dots, X_K^p)$ ) and an output data vector  $Y^p$  for  $p = 1, 2, \dots, L$ . The global energy function is defined by

$$GE^p = \frac{1}{2} [f(X_1^p, \dots, X_K^p) - Y^p]^2. \quad (1)$$

The basic learning algorithm for the GCFNN was described in [19]. Here, for clarity and convenience, the relevant algorithms are introduced below. The new hybrid genetic-algorithms-based learning algorithm will be proposed in Subsect. “Forward-Wave-Backward-Wave Learning”.

### Local Forward-Wave Learning

In general, suppose that a FNN has  $n$  inputs and one output. Given input data vectors  $x^p$  (i.e.,  $x^p = (x_1^p, x_2^p, \dots, x_n^p)$ ) and one-dimensional output data vector  $y^p$  for  $p = 1, 2, \dots, N$ . The energy function is defined by

$$E^p = \frac{1}{2} [f(X_1^p, \dots, X_n^p) - y^p]^2 \quad (2)$$

For simplicity, let  $E$  and  $f^p$  denote  $E^p$  and  $f(x_1^p, x_2^p, \dots, x_n^p)$ , respectively.

Based on the learning algorithm in [17,18], basic steps of the local forward-wave learning algorithm for a FNN are given below:

#### Procedure of Local-Forward-Wave-Learning

**Step 1:** Begin.

**Step 2:** Heuristic Initialization of Parameters.

**Step 3:** Gradient Descending Learning From Data.

Then we can get the following learning algorithms for  $i = 1, 2, \dots, n, k = 1, 2, \dots, m, p = 1, 2, \dots, N, t = 0, 1, 2, \dots$  and a learning rate  $\lambda > 0$ .

**Step 3.1:** Train  $b^k$

$$b^k(t+1) = b^k(t) - \lambda \frac{\partial E}{\partial b^k}|_t, \quad (3)$$

**Step 3.2:** Train  $\eta^k$

$$\eta^k(t+1) = \eta^k(t) - \lambda \frac{\partial E}{\partial \eta^k}|_t, \quad (4)$$

**Step 3.3:** Train  $a_i^k$

$$a_i^k(t+1) = a_i^k(t) - \lambda \frac{\partial E}{\partial a_i^k}|_t, \quad (5)$$

**Step 3.4:** Train  $\sigma_i^k$

$$\sigma_i^k(t+1) = \sigma_i^k(t) - \lambda \frac{\partial E}{\partial \sigma_i^k}|_t, \quad (6)$$

**Step 3.5:** Train  $wleft_i^k$  or  $wright_i^k$

$$\text{IF } x_i^k \leq a_i^k \text{ THEN } wleft_i^k(t+1) = wleft_i^k(t) - \lambda \frac{\partial E}{\partial wleft_i^k}|_t, \quad (7)$$

$$\text{ELSE } wright_i^k(t+1) = wright_i^k(t) - \lambda \frac{\partial E}{\partial wright_i^k}|_t. \quad (8)$$

**Step 4:** End.

### Global Backward-Wave Learning

Based on the traditional back-propagation learning method, the global backward-wave learning algorithm is proposed to train all local FNNs in the back-propagation manner. The following procedure is a general algorithm for updating parameters of any FNN in GCFNN.

#### Procedure of Global-Backward-Wave-Learning

**Step 1:** Begin.

**Step 2:** Calculate a back-propagation error  $\delta^l$  to the relevant  $(l-1)$ th stage FNNs based on an error  $\delta^{(l+1)}$  from the  $(l+1)$ th stage FNN.

$$\delta^l = \frac{\partial E}{\partial y^{l-1}} = \delta^{(l+1)} \frac{\partial y^l}{\partial y^{l-1}}, \quad (9)$$

where  $y^l$  and  $y^{l-1}$  are outputs of the  $l$ th stage FNN and the  $(l-1)$ th stage FNN, respectively.

**Step 3:** Update parameters using the back-propagation error  $\delta^l$ .

**Step 3.5:** Train  $wleft_i^k$  or  $wright_i^k$

**Step 4:** Discovering Fuzzy Knowledge.

**Step 5:** End.

### Forward-Wave-Backward-Wave Learning

The new hybrid learning algorithm for training is called *forward-wave-backward-wave learning algorithm*, which combines the techniques of genetic algorithms, local forward-wave learning and global backward-wave learning.

#### Procedure Genetic-Forward-Wave-Backward-Wave-Learning

**Step 1:** Begin.

**Step 2:** Partition the original training data sets input data vectors  $X^p$  and an output data vector  $Y^p$  for  $p = 1, 2, \dots, L$  into  $M$  local training data sets where  $M$  is the number of 1st-stage granular neural networks.

**Step 3:** Local Forward-Wave Learning: *Local-forward-wave-learning()*.

**Step 4:** Global Backward-Wave Learning: *Global-backward-wave-learning()*.

**Step 5:** End.

### Applications in Bioinformatics

A new granular neural network is used to form the new tertiary architecture [12]. The neurons in the new granular neural network are SVM machines that classify the input data into two classes of protein structures. The two classes are the binary classes that the SVM machines are actually trained for. The tertiary classifier's architecture is explained in subsequent sections.

The new tertiary classifier makes use of both one-versus-one as well as one-versus-rest binary classifiers. The novel architecture makes use of all the six binary classifiers in neural net architecture. The architecture is shown in the Fig. 2 [12].

There are two hidden layers; The output of the first one is the same as the output of the individual SVM.

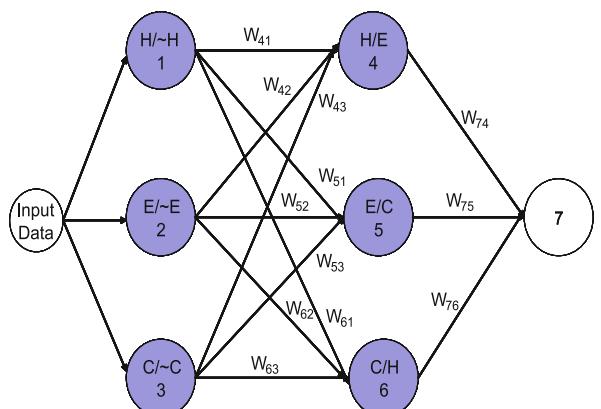
Outputs of first hidden layer

$$O_1 = \text{SVM}(H/\sim H)$$

$$O_2 = \text{SVM}(E/\sim E)$$

$$O_3 = \text{SVM}(C/\sim C).$$

The output of the second hidden layer considers the output of the first layer as well as the SVM machine stored



**Granular Neural Network, Figure 2**  
Genetic Neural Support Vector Machines

in that layer. For example the output of the neuron 4 has an SVM binary classifier that positively classifies  $H$  and negatively classifies  $E$ , the result of this SVM is combined with that of the first layer outputs. This method uses the outputs of the three one-versus-rest classifier in a single neuron.

In the formulations the output of the second hidden layer is formed by adding the output of the SVM sitting inside the neuron with the product of the weight and output of the corresponding neuron (i.e. the neuron which positively classifies the same class as the current neuron) and by subtracting the products of the other two neurons in the first layer.

The output of the second layer are calculated as

$$O_4 = \text{SVM}(H/E) + W_{41}O_1 - W_{42}O_2 - W_{43}O_3 .$$

In the above formula we add the values of the SVM that positively classify the same class ( $H$ ) and subtract those that positively classify other classes. Here  $W_{41}$  means weight between neuron 1 and neuron 4. Similarly other weight corresponds to output, input naming pattern.

Similarly outputs of other two neurons in the second hidden layer are calculated as

$$O_5 = \text{SVM}(E/C) + W_{51}O_1 - W_{52}O_2 - W_{53}O_3$$

$$O_6 = \text{SVM}(H/C) + W_{61}O_1 - W_{62}O_2 - W_{63}O_3 .$$

The final output layer does not have any SVM embedded in it. It calculates its results based on maximum of the three outputs of second hidden layer. There is only one neuron in this layer. The final output is one among the three classes ( $H, E$  or  $C$ ), which ever neuron produces the maximum output after multiplying it with appropriate weight with the second hidden layer output is considered as final output.

So the output of the third layer is as follows

If  $(\text{Max } (W_{74}O_4, W_{75}O_5, W_{76}O_6) = W_{74}O_4)$

Then

$O_7 = H$

Else If  $(\text{Max } (W_{74}O_4, W_{75}O_5, W_{76}O_6 = W_{75}O_5))$

Then

$O_7 = E$

Else

$O_7 = C$

For optimizing the weights, genetic algorithms are used. The weight range for the Genetic Granular Neural Network (GGNN) is selected to be between 0 and 1 so that the architecture performs to its full potential.

The same tertiary classifiers were demonstrated with binary classifiers using multiple windows scheme. This resulted in increase in total accuracy level, which is expected

**Granular Neural Network, Table 1**  
**Accuracy Comparisons**

Tertiary Classifier	Accuracy (%)
SVM_VOTE	62.6
SVM_REPRESNT.	64.8
GGNN	68.0

**Granular Neural Network, Table 2**  
**Meaning of data parameters**

dd/mm/yyyy	Date	Date of the day
Double	Open	Initial price
Double	Low	Lowest traded price
Double	High	Highest traded price
Double	Close	Price of the last trade
Double	Volume	Number of traded stocks

as the binary classifiers formed using multiple window scheme are better when compared to single window encoding scheme. The binary classifier used is constructed using three consecutive windows each of size 5 with gaps between the first and second window as well as between second window and third window. The results of these simulations are shown in Table 1.

## Applications in Computational Web Intelligence

### Concepts of Input Data

Database tables have been generated using stock-historical data from [www.yahoo.com](http://www.yahoo.com) site. Each line of the data set must contain 5 values date, open, high, low, close values of the stock. Each value is separated by space (see Table 2) [21].

### Overview of Implementation

A full run of the program implementation will be described, going through all the main features of the program [21]:

- Download historical data from the Internet.
- Copy the whole data into a text and run the program, which inserts the data into the database.
- A program is written through which users can buy and sell the stock shares, and the corresponding data is stored in the database.
- A program is written through which each user can see his/her own transactions.
- Algorithm, which trains the granular neural networks using the mean square error as stop criterion for learning, while never exceeding the maximum number of

cycles which can take testing data from the initial date to the user entered date, and predicts the future stock closing values.

- A program is written, which compares the predicted values with real values.

One of the most important factors here is to construct a neural network deciding on what the network will learn. A neural network must be trained on some input data. The two major problems in implementing the training are

- Defining the set of input to be used (the learning environment)
- Deciding on an algorithm.

### Performance

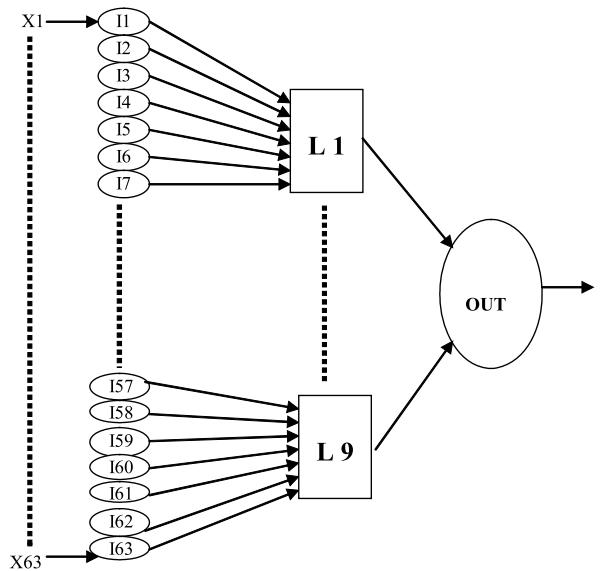
The performance of the granular neural network algorithm is compared with the performance of the BP algorithm by training the same set of data and predicting the future stock values. If the training error was set at 0.03 and the neural network was trained for dow stock data using both the algorithms. The granular neural network took 2 min 58 s to train the neural network where as BP took 2 h and 55 min. The granular neural network's average error was 1.39 where as BP gave 3.38.

The average error for granular neural network is less compared to the average error for BP algorithm. From the average error and the graph it is conclusive that, granular neural network produced closer future stock values with the real stock values compared to the BP algorithm using less training error. If the training error was set at 0.07 and the neural network was trained for cisco stock data using both the algorithms. The granular neural network took 2 min to train the neural network where as BP took 1 h and 48 min. The granular neural network's average error was 6.09 where as BP gave 7.16.

The average error for the granular neural network is less compared to the average error for BP algorithm. From the average error and the graph it is conclusive that, the granular neural network produced closer future stock values with the real stock values compared to the BP algorithm using less training error. Based on the above two simulations, the overall performance with the granular neural network technique is better than BP technique.

### Applications in Brain Informatics

Recent advances in granular computing, soft computing and cognitive science have allowed an increase understanding of normal and abnormal brain functions, especially in the research of human's pattern recognition by means of computational intelligence.



**Granular Neural Network, Figure 3**

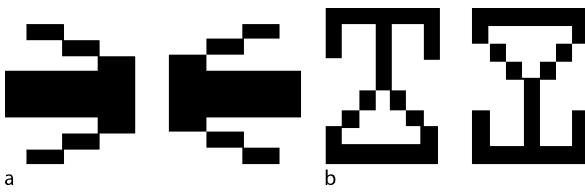
The genetic granular neural network structure for pattern recognition [5]

The simulations of the genetic granular neural network for pattern recognition have two ways: (1) testing whether a pair of noised pattern images can still be identified by the genetic granular neural network as similar as its original version after horizontal or vertical transformation, and (2) testing whether the genetic granular neural network can recognize vertically or horizontally reversed pattern pairs with noise disturbance [5].

During each test, specific pattern samples are used for training and other groups of samples are served for testing. Each pattern sample is depicted by  $7 \times 9$  pixel matrix, so totally 63 inputs are required in order to present one pattern, as shown in Fig. 3. To accommodate the input requirement, the genetic granular neural network's granular layer has 9 local genetic granular neural networks with 7 inputs, so that 63 pixel inputs are satisfied. In its second hidden layer, each the genetic granular neural network has one output; the output layer has 9 inputs and 1 output. Therefore, the total number of inputs of the system is 126 [5].

Finally, the system will make a final decision whether or not two images in each testing sample are similar or symmetric based on the maximum output of the three genetic granular neural networks. Human beings are chosen to make their subjective judgment on pattern samples, by answering what degree the two patterns are similar or symmetric to each other [5].

Group A contains 50 pairs of arrow pattern samples within the following five noise levels, which are 0–10%,



**Granular Neural Network, Figure 4**

a In Group A on the left, the second arrow pattern is transformed from the first after horizontal reversal; b In Group B on the right, the second cup-like pattern is transformed from the first after vertical reversal [5]

**Granular Neural Network, Table 3**

#### Simulation results

Group name	Different answers	Total samples	Agreement percentage
Group A	5	50	90%
Group B	22	50	56%
Total	27	100	73%

10–20%, 20–30%, 30–40%, 40–50% respectively. In each pair of arrow pattern sample, as the Fig. 4a, the second pattern is transformed from the first pattern after horizontal reversal at the first, and then added into the same level of noises.

As shown in Fig. 4b, Group B includes 50 pairs of cup-like pattern samples within the same five noise levels as Group A. In each pair of cup-like-pattern sample, the second pattern is transformed from the first pattern after vertical reversal and then added into the same level of noises.

Table 3 shows the comparisons of decisions made by human beings and the genetic granular neural network for scenario1. As can be seen, the genetic granular neural network and the human brain make 73% identical choices with total 100 testing data. In Group A, Only 10% of 50 testing samples are different between the genetic granular neural network and human; while, in Group B, almost half of the testing samples, which are more than four times higher than Group A, introduce contradictive decisions. Thus, human and computer system in Group B make much more different answers than in Group A. It may indicate that the horizontal reversal causes more confusions than vertical reversal for human or the genetic granular neural network [5].

According to Table 4 and Table 5, we discovered an interesting phenomenon: As noise level increases, the disagreement on pattern's similarity between human and GGCFNN tends to decrease. More specifically, as shown in Table 4, disagreement percentage is exactly 0% in (40%, 50%) while it increases to 16.67% in (0%, 30%). In Table 5,

**Granular Neural Network, Table 4**

Different answers to vertical reversal within 5 noise levels

Noise level	Different answers	Total samples	Disagreement percentage
40–50% noise	0	10	0%
30–40% noise	0	10	10%
20–30% noise	1	10	10%
10–20% noise	1	10	10%
0–10% noise	3	10	30%

**Granular Neural Network, Table 5**

Different answers to horizontal reversal within 5 noise levels

Noise level	Different answers	Total samples	Disagreement percentage
40–50% noise	3	10	30%
30–40% noise	4	10	40%
20–30% noise	4	10	40%
10–20% noise	4	10	40%
0–10% noise	7	10	70%

disagreement percentage is increased from 36% within the noise level (20%, 50%) to 55% in (0%, 20%). Additionally, the disagreement percentages experience prominent increases or decreases within low noise levels.

The genetic granular neural network has exhibited good learning performance when stimulating human's pattern recognition in term of symmetry and similarity. It achieved higher agreement ratios when recognizing symmetrical patterns than similar patterns, and vertical transformation than horizontal transformation, even within high noise levels. The preliminary experiments indicated that the genetic granular neural network is able to simulate the human's recognition ability and adapt its learning experience to new patterns and then achieve quite similar results as human beings.

## Conclusions

The human brain has high intelligence to recognize different geometrical patterns in terms of the similarity and symmetry, but a systematic framework of biological neural network has not been established. The granular neural network aims at simulating functionality of the human brain in terms of processing granular data such as linguistic terms, clusters, sets and classes.

Since the human brain consists of biological neural networks that are the major components performing intelligent tasks, it is very important to develop advanced algorithms and mathematical models for the granular neural networks.

## Future Directions

In the future, the challenging problem is how to design more powerful granular neural networks for more complex applications effectively and efficiently. A lot of advanced intelligent techniques such as type-1/type-2 fuzzy sets [4,15], rough sets [6,7,8], granular computing [8,9], granular systems [16], granular support vector machines [13], granular kernel machines [3] and neuro-fuzzy systems [2,17,19] can be merged to design a hybrid granular neural networks with various functions and outstanding performance.

## Bibliography

1. Dick S, Kandel A (2001) Granular weights in a neural network. In: IFSA World Congress and 20th NAFIPS International Conference, vol 3, Vancouver, July 2001, pp 1708–1713
2. Jang J-SR (1993) ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23(3):665–685
3. Jin B, Zhang Y-Q, Wang BH (2007) Granular kernel trees with parallel genetic algorithms for drug activity comparisons. *Int J Data Mining Bioinform* 1(3):270–285
4. Karnik NN, Mendel JM, Liang Q (1999) Type-2 fuzzy logic systems. *IEEE Trans Fuzzy Syst* 7:643–658
5. Lin C, Li J, Barrett N, Zhang Y-Q, Washburn DA (2006) Genetic granular cognitive fuzzy neural networks and human brains for pattern recognition. In: The WICI International Workshop on Web Intelligence (WI) meets Brain Informatics (BI) (WlmbI 2006), Beijing, 15–16 December 2006
6. Lin TY (1997) Granular computing: From rough sets and neighborhood systems to information granulation and computing in words. In: European Congress on Intelligent Techniques and Soft Computing, 8–12 September 1997, pp 1602–1606
7. Lin TY (1999) Granular computing: Fuzzy logic and rough sets. In: Zadeh L, Kacprzyk J (eds) Computing with Words in Information/Intelligent Systems. Physica, Heidelberg, pp 184–200
8. Lin TY, Yao YY, Zadeh L (2002) Data mining, rough sets and granular computing. Physica, Heidelberg
9. Pedrycz W (2001) Granular computing: An emerging paradigm. Physica, Heidelberg
10. Pedrycz W, Vukovich G (2001) Granular neural networks. *Neurocomputing* 36:205–224
11. Pedrycz W, Kandel A, Zhang Y-Q (1997) Neurofuzzy Systems. In: Dubois D, Prade H (eds) Fuzzy systems: Modeling and control. Kluwer, Norwell, pp 311–380
12. Reyaz-Ahmed A, Zhang Y-Q (2007) Protein secondary structure prediction using genetic neural support vector machines. In: Proc of IEEE 7th International Conference on Bioinformatics and BioEngineering, Boston, 14–17 October 2007, pp 1355–1359
13. Tang YC, Jin B, Zhang Y-Q (2005) Granular support vector machines with association rules mining for protein homology prediction. *Artif Intell Med* (Special Issue on Computational Intelligence Techniques in Bioinformatics) 35(1–2):121–134
14. Vasilakos A, Stathakis D (2005) Granular neural networks for land use classification. *Soft Comput* 9:332–340
15. Zadeh LA (1979) Fuzzy sets and information granulation. In: Gupta N, Ragade R, Yager R (eds) Advances in fuzzy set theory and applications. Elsevier, St. Louis, pp 3–18
16. Zhang Y-Q (2005) Constructive granular systems with universal approximation and fast knowledge discovery. *IEEE Trans Fuzzy Syst* 13(1):48–57
17. Zhang Y-Q, Chung F (2002) Fuzzy neural network tree with heuristic back-propagation learning. In: Proc of IJCNN of World Congress on Computational Intelligence 2002, Honolulu, May 2002, pp 553–558
18. Zhang Y-Q, Kandel A (1998) Compensatory genetic fuzzy neural networks and their applications. In: Series in machine perception artificial intelligence, vol 30. World Scientific, Singapore
19. Zhang Y-Q, Kandel A (1998) Compensatory neurofuzzy systems with fast learning algorithms. *IEEE Trans Neural Netw* 9(1):83–105
20. Zhang Y-Q, Fraser MD, Gagliano RA, Kandel A (2000) Granular neural networks for numerical-linguistic data fusion and knowledge discovery. *IEEE Trans Neural Netw* 11(3):658–667
21. Zhang Y-Q, Akkaladevi S, Vachtsevanos G, Lin TY (2001) Fuzzy neural web agents for stock prediction. In: Proc of FLINT2001, UC Berkeley, 14–18 August 2001, pp 101–105

## Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems

LECH POLKOWSKI

Polish-Japanese Institute of Information Technology,  
Warsaw, Poland

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Similarity](#)

[Information and Decision Systems](#)

[Granulation of Knowledge: General Ideas  
and Technique of Rough Inclusions](#)

[Calculus on Granules](#)

[Some Principal Fields of Applications](#)

[Complexity Issues](#)

[Future Directions](#)

[Bibliography](#)

### Glossary

**Knowledge** The notion of knowledge can be described by positing the structure of the perceivable world as a system of states of things. These states are related among themselves by relations which in turn form a network of interdependent combinations. States and relations,

as well as dependencies among them, are reflected in knowledge: things are reflected in objects or notions, relations in notions or concepts, states of things in sentences. Sentences form knowledge [6]; for instance, in Artificial Intelligence, a judiciously chosen set of sentences forms a knowledge base for a logical agent (see [52]).

**Knowledge representation** A chosen symbolic system (language) by means of which notions are encoded and reasoning is formalized.

**Description logic** A language of concepts commonly used for knowledge representation in Artificial Intelligence logical systems (see [2]). The variant which is interesting to us is constructed in the attribute-value context in which objects are described in terms of chosen attributes (features) and their values. Primitive formulas of *description logic* are *descriptors* of the form  $(a = v)$  where  $a$  is an attribute and  $v$  is one of its values. Descriptors are interpreted in the universe  $U$  of objects: the meaning  $[(a = v)]$  of the descriptor  $(a = v)$  is defined as the set  $\{u \in U : a(u) = v\}$ . Descriptors are extended to formulas of description logic with the help of connectives of sentential calculus: when  $\alpha, \beta$  are formulas then  $\alpha \vee \beta$ ,  $\alpha \wedge \beta$ ,  $\neg\alpha$ ,  $\alpha \Rightarrow \beta$  are formulas as well. The meaning of formulas is defined by recursion:  $[\alpha \vee \beta] = [\alpha] \cup [\beta]$ ,  $[\alpha \wedge \beta] = [\alpha] \cap [\beta]$ ,  $[\neg\alpha] = U \setminus [\alpha]$ . The meaning of implication  $\alpha \Rightarrow \beta$  depends on the chosen interpretation of implication; when it is interpreted as in sentential calculus, that is, as  $\neg\alpha \vee \beta$ , the meaning is already defined by the given formulas for  $\vee, \neg$ .

**Rough sets: knowledge as classification** Rough set theory proposed by Pawlak [30,32], understands knowledge as classification: a set of equivalence relations, each of which is understood as a certain classification of objects into its equivalence classes. Together, these relations constitute the knowledge base  $\mathcal{R}$ . Objects that belong in one class  $[u]_R$  of a relation  $R$  in the knowledge base are *R-indiscernible*. Objects that belong in the intersection  $\bigcap_{R \in \mathcal{R}} [u]_R$  are *R-indiscernible*: they cannot be discerned by means of the knowledge available in  $\mathcal{R}$ .

Indiscernibility relations are formally defined for sets of relations: given  $\mathcal{T} \subseteq \mathcal{R}$ , the  $\mathcal{T}$ -indiscernibility relation  $\text{ind}(\mathcal{T})$  is defined as the intersection  $\bigcap_{R \in \mathcal{T}} R$ . Concepts are divided into two classes:  $\mathcal{T}$ -exact concepts are sets of objects which can be expressed as unions of  $\mathcal{T}$ -indiscernibility classes; other concepts are  $\mathcal{T}$ -rough. For each concept  $X$ , there exist the greatest exact concept  $\underline{\mathcal{T}}$  contained in  $X$  (the  $\mathcal{T}$ -lower approximation to  $X$ ) and the smallest exact concept  $\overline{\mathcal{T}}$

which contains  $X$  (the  $\mathcal{T}$ -upper approximation to  $X$ ). Rough concepts are perceived as uncertain sets sandwiched between a lower approximation and an upper approximation.

**Information system** A system commonly used for representing information about a given world fragment. An information system can be represented as a collection of pairs of the form  $(U, A)$  where  $U$  is a set of objects – representing things – and  $A$  is a set of attributes; each attribute  $a$  is modeled as a mapping  $a : U \rightarrow V_a$  from the set of objects into the value set  $V_a$ . For an attribute  $a$ , the equivalence relation  $R_a$  is the set  $\{(u, v) \in U \times U : a(u) = a(v)\}$  and the collection  $\mathcal{R} = \{R_a : a \in A\}$  is the knowledge base. For each set  $B \subseteq A$  of attributes, the  $B$ -indiscernibility relation  $\text{ind}(B)$  is the set  $\{(u, v) \in U \times U : a(u) = a(v) \text{ for each } a \in B\}$ . Each object  $u \in U$  can be described in two ways: first, by means of its indiscernibility class  $[u]_A = \{v \in U : (u, v) \in \text{ind}(A)\}$ ; next, by means of its descriptor logic formula  $\phi^A(u) : \bigwedge_{a \in A} (a = a(u))$ ; as  $[\phi^A(u)] = [u]_A$ , both descriptions are equivalent and  $u$  is described up to its equivalence class.

**Decision system** A particular form of an information system, a triple  $(U, A, d)$  in which  $d$  is the *decision*, the attribute not in  $A$ , that expresses the evaluation of objects by an external oracle, an expert. Attributes in  $A$  are called *conditional* which expresses the fact that they are set by us as conditions describing objects in our language. The principal problem related to a decision system is to find a description of dependence between conditional knowledge given by  $(U, A)$  and the expert knowledge  $(U, \{d\})$ .

**Decision rule** A *decision rule* is a formula in descriptor language that expresses a particular relation among conditional attributes in the attribute set  $A$  and the decision  $d$ , of the form:  $\bigwedge_{a \in A} (a = v_a) \Rightarrow (d = v)$  with the semantics defined in (Glossary: Description logic). The formula is *true*, or *certain* in case  $[\bigwedge_{a \in A} (a = v_a)] = \bigcap_{a \in A} [(a = v_a)] \subseteq [(d = v)]$ . Otherwise, the formula is *partially true*, or *possible*. An object  $o$  which satisfies the rule  $a(o) = v_a$  for  $a \in A$  can be assigned to the class  $[(d = v)]$ ; often a partial match based on a chosen similarity measure must be performed.

**Similarity** Similarity relations are weaker forms of indiscernibility: indiscernible objects are similar but not necessarily vice versa an example attributable to Henri Poincaré [35] explains this best: assume that points  $x, y$  in the real line are similar whenever their distance is less than a fixed threshold  $\delta$ :  $|x - y| < \delta$ . Then certainly each  $x$  is similar to  $x$ , and if  $x$  is similar to  $y$  then

$y$  is similar to  $x$ , but from  $x$  similar to  $y$  and  $y$  similar to  $z$  it need not follow that  $x$  is similar to  $z$ :  $|x - z|$  can be close to  $2\delta$ . Relations of this kind are reflexive and symmetric: they were called *tolerance relations* in Zeeman [67]. Tolerance relations in the framework of rough set theory information systems were considered in Nieminen [26] and tolerance reducts and related notions in information systems were discussed in Polkowski, Skowron and Zytkow [50]. Some authors relax tolerance relations to similarity relations that need not be symmetric: for instance, rough inclusions need not be symmetric [37].

In many applications, metrics are taken as similarity measures: a metric on a set  $U$  is a function  $d: U \times U \rightarrow R^+$  into non-negative reals such that 1.  $d(x, y) = 0$  if and only if  $x = y$ . 2.  $d(x, y) = d(y, x)$ . 3.  $d(x, y) \leq d(x, z) + d(z, y)$  for each  $x, y, z$  in  $X$  (the *triangle inequality*); similarity induced by a metric  $d$  is usually of the form  $d(x, y) < \delta$  for a certain  $\delta > 0$ , Poincaré-style.

**Rough inclusion** A ternary relation  $\mu$  on a set  $U \times U \times [0, 1]$  which satisfies conditions: 1.  $\mu(x, x, 1)$ . 2.  $\mu(x, y, 1)$  is a binary partial order relation on the set  $X$ . 3.  $\mu(x, y, 1)$  implies that for each object  $z$  in  $U$ : if  $\mu(z, x, r)$  then  $\mu(z, y, r)$ . 4.  $\mu(x, y, r)$  and  $s < r$  imply that  $\mu(x, y, s)$ . The formula  $\mu(x, y, r)$  is read as “the object  $x$  is a part in object  $y$  to a degree at least  $r$ ”. The partial containment idea encompasses the idea of an exact part, that is, mereological theory of concepts Leśniewski [15]. Similarity induced by a rough inclusion can be introduced as an asymmetric relation:  $\mu(x, y, r)$ , or a symmetric one:  $\mu(x, y, r) \wedge \mu(y, x, r)$ .

**Granulation of knowledge** The concept of granulation was proposed by Zadeh [67]. In fuzzy calculus, granules are naturally induced as the inverse images of fuzzy membership functions and fuzzy computing is a fortiori computing with granules. A similar case happens with rough sets, as primitive objects in this paradigm are indiscernibility classes that are elementary granules, whereas their unions are granules of knowledge; reasoning in rough sets means, by definition, reasoning with granules. Each similarity relation induces granules understood as its classes; in comparison to indiscernibility relations, granules induced by similarity relations are more intricate due to lack of transitivity: relations among granules are more difficult to ascertain. In defining granules, attention is focused on the property that singles out objects belonging in the granule; usually, this property is a relation to an object chosen as the granule center: this means that distinct objects in the granule need not

be in the relation. This feature distinguishes granules from *clusters*: aggregates in which any two objects have the defining property, for example, distance less than the chosen threshold. Also, relations between pairs of distinct granules are difficult to ascertain, contrary to the case of (for example) clusters, in which case the distance between distinct clusters is kept above a fixed threshold.

**Classification of objects** Classification entails assigning to each element in a set of objects (test sample) a class (a decision) to which the given element should belong; it is accomplished on the basis of knowledge induced from the given collection of examples (the training sample). To perform this task, objects are usually mapped onto vectors in a multi-dimensional real vector space (feature space). Paradigms invented to perform classification are many, from heuristics based on cognitive networks such as neural networks, through probabilistic tools such as Bayesian networks to tools based on distance such as k-nn methods and prototype methods (see [7,12]).

**Fusion of knowledge** A combination of knowledge coming from few distinct sources. For example, in a mobile robot, fusion of knowledge by means of a Kalman filter (see [52]), consists in producing an output – a robot’s predicted position at the current moment – on the basis of inputs: a robot’s position estimated as the last moment preceding the current moment, the current controls, and sensor readings in the current position.

**Reasoning** Processes of reasoning include an effort by means of which sentences are created; various forms of reasoning depend on the chosen system of notions, symbolic representation of notions, forms of manipulating symbols (see [6]).

**Mereology** Mereology, see Leśniewski [15], is a theory of concepts based on the notion of a part instead – as with naive set theory – on the notion of an element. The relation of being a part is irreflexive and transitive. A secondary notion of an ingredient means either being a part or being the whole object and it sets a partial order on objects. By means of the notion of an ingredient the notion of a class operator is introduced in Leśniewski [15] which converts ontological notions, that is, collections of objects into an individual object; it is used by us in granule formation.

## Definition of the Subject

The creator of Fuzzy Set Theory, Lotfi A. Zadeh, proposed computing with granules in Zadeh [66]. The idea was natural, as fuzzy reasonings are carried out in terms of fuzzy

membership functions. A fuzzy membership function  $\mu_X$  maps a universe  $U$  of objects into the interval  $[0, 1]$  and it represents the membership in a set  $X$  as a membership to a degree. The value  $\mu_X(x) = r$  is interpreted as the statement that the object  $x$  is an element of the set  $X$  to the degree of  $r$ . The mapping  $U \rightarrow U/\mu_X$  which sends each object  $x$  to its fibre  $\mu_X^{-1}(\mu_X(x))$  includes into the granule  $g_\mu(r)$  all objects that belong in  $X$  to the degree  $r$ . All fuzzy constructs are then expressed in terms of those granules. In this sense, fuzzy reasoning is, in a natural way, reasoning with granules. Granules in this reasoning are constructed in a uniform way, that is, all objects in a granule share the same property of external character: they belong to an “oracle”  $X$  to the same degree; changing  $X$  produces a variety of granules to reason with. The relation forming any granule is an equivalence  $R_X$ : the universe is decomposed into fibres of a fuzzy membership function  $\mu_X$ , and a knowledge base is composed of all relations  $R_X$  for subsets  $X \subseteq U$  of the universe of objects.

The same conclusion concerns rough set reasoning: elementary objects in reasoning are indiscernibility classes – elementary granules which are elements of a partition of the universe of objects by an indiscernibility relation  $\text{ind}(B)$  (see Glossary: Information systems). These granules are the smallest objects which can be described in terms of attributes and their values, that is, by description logic, and they are used in forming description of objects, in building decision rules and then classifiers as well as control algorithms. Lin [16,17] was the first to recognize the topological character of granules and to form the basic notion of a neighborhood system as a collection of granules on the universe of objects (see also [18,19,20,21,22]). Lin [17] recognized the import of tolerance relations by discussing tolerance induced neighborhoods.

In all hybrid approaches involving fuzzy or rough sets along with neural networks, genetic algorithms etc., one is therefore bound to compute with granules, testimony to the importance of granular structures.

In search of adequate similarity relations, various forms of granules have been proposed and considered as well as experimentally verified as to their effectiveness. In information systems, *templates* were proposed as granules, that is, generalized descriptors of the form  $(a \in W_a)$  where  $W_a \subseteq V_a$  with the meaning  $[(a \in W_a)] = \{u \in U: a(u) \in W_a\}$  [27]. Clearly, templates are aggregates in the ontological sense of descriptors, that is, they form “big” granules. Their usage is motivated by their greater descriptive force *vis-a-vis* descriptors; a judicious choice of sets  $W_a$  should allow construction of a similarity relation that accurately reflects the decision’s dependence on conditional attributes.

Rough inclusions have been considered as a means for granule construction; a rough inclusion is (see Glossary: Rough inclusions) a generic term for a ternary predicate that formalizes the phrase: “the object  $x$  is a part of the object  $y$  to a degree  $r$ ”, introduced in [47,48]. Granulation by means of rough inclusions has been studied, for example, in [37,38,39,40,41,42,43,44]. The idea of granule formation comes from mereology and it rests on using the class operator (see Glossary: Mereology).

Granules formed by rough inclusions are used in fusion of knowledge models, in rough-neural computing and in building many-valued logics reflecting the rough set ideology in reasoning.

## Introduction

Granulation of knowledge can be considered from a few angles:

1. General purpose of granulation.
2. Granules from binary relations.
3. Granules in information systems from indiscernibility.
4. Granules from generalized descriptors.
5. Granules from rough inclusions – mereological approach.

## General Purpose of Granulation

Granulation of knowledge comes into existence for a few reasons: the principal one is founded on the underlying assumption of basically all paradigms, *viz.*, that reality exhibits a fundamental continuity, that is, objects with identical descriptions in a given paradigm should exhibit the same properties with respect to classification or decision making. For instance, fuzzy set theory assumes that objects with identical membership descriptions should behave identically and rough set theory assumes that objects indiscernible with respect to a group of attributes should behave identically; in particular, they should fall into the same decision class.

Thus, granulation is implied by assumptions of the paradigm and is unavoidable once the paradigm is accepted and applied. Granules induced in the given paradigm by necessity form the first level of granulation.

In the search for the most appropriate similarities for use in applications such as classification or decision making, more complex granules are constructed, for example, as unions of granules of the first level, or more complex functions of them, resulting, for example, in fusion of various granules from distinct sources.

Among granules of the first two levels, some kinds can be exhibited by means of various operators, for example, the class operator associated with a rough inclusion.

## Granules from Binary Relations

Granulation on the basis of binary general relations as well as, for example, tolerance relations, has been studied by T. Y. Lin [16,17,18,19,20,21,22,23] in particular as an important notion of neighborhood systems; see also Yi Yu Yao [64,65]. This extends the approach based on indiscernibility as a special case. A general form of this approach, according to Yao, exploits the classical notion of Galois connection: two mappings  $f: X \rightarrow Y, g: Y \rightarrow X$  form a Galois connection between ordered sets  $(X, \prec)$  and  $(Y, \prec)$  if and only if the equivalence  $x \prec g(y) \Leftrightarrow f(x) \prec y$  holds. In an information system  $(U, A)$ , for a binary relation  $R$  on  $U$ , one considers the sets  $xR = \{y \in U: xRy\}$  and  $Rx = \{y \in U: yRx\}$ , called, respectively, the *successor neighborhood* and the *predecessor neighborhood* of  $x$ . These sets are considered as granules formed by a specific relation of being affine to  $x$  in the sense of the relation  $R$ . Other forms of granulation can be obtained by comparing objects with identical neighborhoods, for example,  $x \equiv y \Leftrightarrow xR = yR$  etc. Saturation of sets of objects  $X, Y$  with respect to the relation  $R$  leads to sets  $X^*, Y^*$  such that  $X * R = Y^* \Leftrightarrow RY^* = X^*$ , forming a Galois connection. This approach is closely related to the Formal Concept Analysis of Wille [61].

## Granules in Information Systems from Indiscernibility

Granules based on indiscernibility in information/decision systems are constructed as indiscernibility classes: given an information system  $(U, A)$  (see Glossary: Information systems; Decision systems), and the collection  $\text{IND} = \{\text{ind}(B): B \subseteq A\}$  of indiscernibility relations, each elementary granule is of the form  $[u]_B = \{v \in U: (u, v) \in \text{ind}(B)\}$  for some  $B$ . Among those granules, there are minimal ones: granules of the form  $[u]_A$  induced from the set  $A$  of all attributes. Granules  $[u]_A$  form the finest partition of the universe  $U$ ; given the class  $[u]_B$  and the class  $[u]_{A \setminus B}$ , we have  $[u]_A = [u]_B \cap [u]_{A \setminus B}$  hence  $[u]_B = \bigcup_{v \in [u]_B \cap \text{DIS}_{A \setminus B}(u)} [v]_A$ , where  $v \in \text{DIS}_{A \setminus B}(u)$  if and only if there exists an attribute  $a \in A \setminus B$  such that  $a(u) \neq a(v)$ . It is manifest that granules  $[u]_B$  can be arranged into a tree with the root  $[u]_\emptyset = U$  and leaves of the form  $[u]_A$ .

Granules based on indiscernibility form a complete Boolean algebra generated by atoms of the form  $[u]_A$ : unions of these atomic granules are closed on intersections and complements and these operations induce into granules the structure of a field of sets.

Atomic granules are aggregated into some important unions by approximation operators: given a concept

$X \subseteq U$ , the lower approximation  $\underline{BX}$  to  $X$  over the set  $B$  of attributes is defined as the union  $\bigcup\{[u]_B: [u]_B \subseteq X\}$ ; the operator  $L_B: \text{Concepts} \rightarrow \text{Granules}$  sending  $X$  to  $\underline{BX}$  is monotone increasing and idempotent:  $X \subseteq Y$  implies  $\underline{X} \subseteq \underline{Y}$  and  $L \circ L = L$ . Similarly, the upper approximation  $\overline{BX} = \bigcup\{[u]_B: [u]_B \cap X \neq \emptyset\}$  to  $X$  over  $B$ , brings some elementary granules into the union; the operator  $U^B$  sending concepts into upper approximations is also monotone increasing and idempotent.

## Granules from Generalized Descriptors

Some authors have made use of generalized descriptors called *templates* (see Nguyen S H [27]). A template is a formula  $T: (a \in W_a)$ , where  $W_a \subseteq V_a$  (see Glossary: Information systems) with the meaning  $g(T): \{u \in U: a(u) \in W_a\}$ . The granule  $g(T)$  can be represented as the union  $\bigcup_{u \in W_a} [u]_a$ ; granules of the form  $[u]_a$  are also called *blocks*, see Grzymala-Busse [10,11].

## Granules from Rough Inclusions – Mereological Approach

Rough inclusions (see Glossary: Rough inclusions) are ternary predicates of the form  $\mu(x, y, r)$  which read: “ $x$  is a part of  $y$  to degree at least  $r$ ”; intuitively, one can expect of  $\mu$  the fulfillment of conditions:

- a.  $\mu(x, x, 1)$ ;
- b.  $\mu(x, y, 1)$  if and only if  $x \text{ ing}_\pi y$ , where  $\pi$  is a part relation (see Glossary: Mereology) and  $\text{ing}_\pi$  is the associated ingredient relation:  $x \text{ ing}_\pi y$  if and only if  $x \pi y$  or  $x = y$ ;
- c. If  $\mu(x, y, 1)$  then for each  $z$ : if  $\mu(z, x, r)$  then  $\mu(z, y, r)$  which is a condition of monotonicity of  $\mu$ ;
- d. If  $r > s$  and  $\mu(x, y, r)$  then  $\mu(x, y, s)$ .

Granule  $g_r(u)$  of the radius  $r$  about the center  $u$  is defined as the class of property  $\Phi(v): \mu(v, u, r)$ . The class operator

Cl<sub>s</sub>: Ontological entities (collections of objects)  
 → Mereological entities (individual objects)

acts on collections, or properties of objects and yields individual objects; an example of such a class operator in set theory with the relation  $\subset$  of part and the associated relation  $\subseteq$  of ingredient is the union of a family of sets operator  $\bigcup$ : it converts a family (a property) of sets  $\mathcal{F}$  into its union  $\bigcup \mathcal{F}$  which is a single set.

The class operator Cl<sub>s</sub> applied to a non-vacuous property  $\Phi$  of objects, yields the object Cl<sub>s</sub>  $\Phi$  which is defined

as the unique object with the properties:

1. If  $\Phi(u)$  then  $\text{ing}_\pi \text{Cls } \Phi$ ;
2. If  $u \text{ ing}_\pi \text{Cls } \Phi$  then there exist objects  $v, w$  such that  $v \text{ ing}_\pi u, v \text{ ing}_\pi w, \Phi(w)$ .

In other words,  $\text{Cls } \Phi$  absorbs all objects whose each part has a part in common with an object which satisfies  $\Phi$ .

Hence:  $g_r(u)$  is  $\text{Cls}\{v: \mu(v, u, r)\}$ . Properties of granules defined in this way depend on properties of rough inclusion  $\mu$ .

There are many rough inclusions to select from, and most are defined from one of three sources (see Polkowski [37,38,39,40,41,42,43,44])

**A** From archimedean t-norms: an archimedean t-norm  $t$  see Polkowski [36], admits a representation  $t(x, y) = g_t(f_t(x) + f_t(y))$  with a continuous decreasing  $f_t: [0, 1] \rightarrow [0, 1]$  and its pseudo-inverse  $g_t$ , see Ling [24] or Polkowski [36].

We define the rough inclusion  $\mu_t$  as follows:

$$\mu_t(u, v, r) \quad \text{if and only if} \quad g_t\left(\frac{|\text{DIS}_A(u, v)|}{|A|}\right) \geq r,$$

where  $\text{DIS}_A(u, v) = \{a \in A: a(u) \neq a(v)\}$ .

The most important example of rough inclusions obtained in this way is the rough inclusion  $\mu_L$  induced from the Łukasiewicz t-norm  $t_L(x, y) = \max\{0, x + y - 1\}$  for which  $f(x) = 1 - x$  and  $g(y) = 1 - y$ , see Ling [24] or Polkowski [36]. Introducing the complement  $\text{IND}_A(u, v) = U \times U \setminus \text{DIS}_A(u, v)$ , one can write the rough inclusion  $\mu_L$  as,

$$\mu_L(u, v, r) \quad \text{if and only if} \quad \frac{|\text{IND}_A(u, v)|}{|A|} \geq r$$

Thus, objects  $u, v$  are similar at least to degree  $r$  if and only if the probability of randomly choosing an attribute which does not discern between  $u$  and  $v$  is at least  $r$ . This rough inclusion is also induced by the relative Hamming distance on information sets of objects in the set  $U$  of the information system (see [43]).

**B** From continuous t-norms (see [43]): a continuous t-norm  $t$  induces the *residual implication*  $\Rightarrow_t$  by the formula,

$$x \Rightarrow_t y \leq z \quad \text{if and only if} \quad t(x, z) \leq y,$$

and it is well-known that  $x \Rightarrow_t y = 1$  if and only if  $x \leq y$ . Thus, residual implications are proper candidates for rough inclusions, once proper set functions

$\Phi(u, v), \Psi(u, v)$  are found such that  $\mu_t(u, v, r)$  holds if and only if  $\Phi(u, v) \Rightarrow_t \Psi(u, v) \geq r$  holds. Examples are given in Polkowski [43].

**C** It is evident that for a metric  $d$  on the universe  $U$ , the predicate  $\mu_d(u, v, r)$  is satisfied if and only if  $d(u, v) \leq 1 - r$  is a rough inclusion. In particular, for the Hamming metric relative to the set  $A$  of attributes,  $h(u, v) = \frac{|\{a \in A: a(u) \neq a(v)\}|}{|A|}$ , the rough inclusion  $\mu_h$  coincides with  $\mu_L$ .

A rough inclusion  $\mu_t$  is *transitive* in cases where following the property holds:

$$\text{If } \mu_t(u, v, r) \text{ and } \mu(v, w, s) \text{ then } \mu(u, w, t(r, s)).$$

It has been proved that rough inclusions obtained by methods 1 and 2 are transitive [43]. One can also verify that rough inclusions obtained from metrics by method 3 are transitive with the Łukasiewicz rough inclusion as the transitivity measure: for a metric  $d$  and  $\mu_d(u, v, r), \mu_d(v, w, s)$ , it follows that  $d(u, v) \leq 1 - r$  and  $d(v, w, s) \leq 1 - s$  and the triangle inequality implies that  $d(u, w) \leq (1 - r) + (1 - s)$  hence  $\mu_d(u, w, r + s - 1)$ , that is,  $\mu_d(u, w, t_L(r, s))$ .

Reasoning about granules defined from rough inclusions is carried out on lines of the mereological deduction rule (Leśniewski [15]):

(D(eduction) R(ule) in M(ereology)) For objects  $x, y$ : if for each object  $z$ , from  $z \text{ ing}_\pi x$  it follows that there exists an object  $w$  such that  $w \text{ ing}_\pi z$  and  $z \text{ ing}_\pi y$  then  $x \text{ ing}_\pi y$ .

## Similarity

Similarity relations came to the attention of eminent theorists in the beginning of the twentieth century due to their interest in problems regarding mechanisms of thinking and perception. Henri Poincaré [35] considered relations of discernibility, for example, two points in space are *indiscernible* if and only if their distance is less than a fixed threshold value  $\delta$ . This relation is certainly reflexive (a point cannot be discerned from itself) and symmetric, but may lack the transitivity property; again, in connection with perception mechanisms, these relations were called *tolerance relations* in [67].

Formally, a *tolerance relation*  $\tau$  is a relation which is reflexive:  $\tau(x, x)$  for each  $x$ , and symmetric: if  $\tau(x, y)$  then  $\tau(y, x)$  for each pair  $x, y$ .

Tolerance relations are weaker than equivalence relations of, for example, indiscernibility (see Glossary: Information systems); lack of transitivity makes them more difficult to analyze. In analogy to equivalence relations, one

introduces *tolerance classes*: a tolerance class  $c_\tau$  is defined as a maximal set with the property: for each pair  $x, y \in c_\tau$  the relation  $\tau(x, y)$  holds.

Contrary to equivalence classes, distinct tolerance classes can intersect: the collection of all pairwise distinct tolerance classes is a *covering* of the universe  $U$  of objects.

From the maximality principle in set theory (Teichmuller's, Vaught's or Zorn's, see [14]) it follows that for each pair  $x, y$  with  $\tau(x, y)$  there exists a tolerance class  $c_{\tau(x,y)}$  which contains both  $x, y$ .

We denote by the symbol  $C_\tau(x)$  the collection of all tolerance classes which contain  $x$ . Then the principal property of tolerance relations holds: For each pair  $x, y$  of objects in the universe  $U$  and a tolerance relation  $\tau$  on this universe:  $\tau(x, y)$  if and only if  $C_\tau(x) \cap C_\tau(y) \neq \emptyset$ .

This fact shows that a canonical tolerance relation is defined as the non-empty intersection of sets, and each tolerance relation can be reduced to it.

Tolerance relations in information systems should factor through indiscernibility: if  $\tau(u, v)$  and  $u \text{ ind}(A) u'$ ,  $v \text{ ind}(A) v'$  then  $\tau(u', v')$ . Thus, tolerance relations in information systems should be defined in terms of attributes and their values.

Examples are: metrics-induced tolerance relations, for example,  $\tau(u, v)$  if and only if  $\max_a |a(u) - a(v)| < \delta$  for a fixed value of  $\delta$  (tolerance based on the Manhattan metric) or  $\tau(u, v)$  if and only if  $\sum_a (|a(u) - a(v)|^2)^{\frac{1}{2}} < \delta$  (tolerance based on the Euclidean metric).

Rough inclusions also supply examples of tolerance relations: given a rough inclusion  $\mu$ , and a real number  $r \in [0, 1]$ , the tolerance relation  $\tau_{\mu,r}$  can be defined as follows:  $\tau_{\mu,r}(u, v)$  if and only if  $\mu(u, v, r)$  and  $\mu(v, u, r)$ .

Weaker yet is the class of similarity relations which are merely reflexive: this kind of similarity relations encompasses relations of rough inclusions induced by asymmetric part relations as well as relations discussed in Słowiński et al. [8].

A reflexive similarity relation  $\tau$  can be described in terms of classes  $C_\tau(x) = \{y \in U: \tau(x, y)\}$  for  $x \in U$ . Then:  $\tau(x, y)$  if and only if  $y \in C_\tau(x)$ . Thus, reflexive similarity relations can be canonically described as relations of membership between objects and their collections.

## Information and Decision Systems

The scope of our discussion is limited to knowledge representation in the form of information systems (see Glossary: Information systems, Decision systems, Description logic, Decision rules).

Information systems are means for formal rendering of data usually given in the form of data tables: rows in

the table correspond to objects being described, columns correspond to features used in object descriptions.

Each data table involves the set  $U$  of objects, called often the *universe of objects* and the set  $A$  of *attributes* formally representing features used in object descriptions. The pair  $(U, A)$  is called the *information system*.

For an object  $u \in U$  and a set of attributes  $B$ , the *information set* of  $u$  is the set  $\inf_B(u) = \{(a = a(u)): a \in B\}$ .

Thus, the universe  $U$  can be represented as the set  $\text{Inf}_B = \{\inf_B(u): u \in U\}$ .

A principal assumption of the information system  $(U, A)$  is that all information about objects is supplied, and constructs derived from it should extend to the world of possible objects.

Therefore: objects  $u, v$  which satisfy the condition  $\inf_B(u) = \inf_B(v)$  are regarded as *indiscernible* over  $B$  and perceived through the filter of attributes in  $B$  should be treated as identical with regard to description over  $B$ .

Formal rendering of this postulate is effected by means of *B-indiscernibility relation*  $\text{ind}(B)$  (see Glossary: Rough sets: knowledge as classification):  $\text{ind}(B) = \{(u, v) \in U \times U: \inf_B(u) = \inf_B(v)\}$ . This is an equivalence relation.

Attribute sets produce a variety of indiscernibility relations with obvious properties such as:  $\text{ind}(C) \subseteq \text{ind}(D)$  if and only if  $D \subseteq C$ ;  $\text{ind}(C \cap D) = \text{ind}(C) \cap \text{ind}(D)$ ;  $\text{ind}(A) \subseteq \text{ind}(C) \subseteq \text{ind}(\emptyset) = U \times U$  for each  $C \subseteq A$ .

Classes  $[u]_B = \{v \in U: (u, v) \in \text{ind}(B)\}$  are minimal sets which can be described exactly over the set  $B$ : the class  $[u]_B$  is the meaning of the descriptor formula  $\bigwedge_{a \in B} (a = a(u))$ .

Indiscernibility relations form the knowledge base from which inferences are derived about relations among objects, attributes and attribute sets.

An important example of a property of attribute sets is the reduct property [31,54]: a set  $B$  of attributes is a *reduct* if and only if  $B$  is a minimal (with respect to set inclusion) set of attributes such that  $\text{ind}(B) = \text{ind}(A)$ . Thus, each reduct fully preserves both classification and knowledge.

The meaning of a reduct is also shown by means of dependencies among sets of attributes. The *functional dependency* between sets  $C$  and  $D$  of attributes, see [28,29,31], means that there exists a mapping  $f_{C,D}: \text{Inf}_C \rightarrow \text{Inf}_D$  with the property that  $f(\inf_C(u)) = \inf_D(u)$  for each  $u \in U$ . In this case,  $D$  depends functionally on  $C$ , in symbols,  $C \mapsto D$ .

Functional dependence  $C \mapsto D$  means equivalently that  $\text{ind}_C \subseteq \text{ind}_D$ ; in particular, as  $\text{ind}_B \subseteq \text{ind}_C$  for each reduct  $B$ , the functional dependence  $B \mapsto C$  takes place for each set  $C$  of attributes: each reduct functionally determines all values of attributes.

Finding reducts is computationally difficult, see [54]. In theory, reducts can be found as prime implicants of the Boolean discernibility function. Decision systems are (see Glossary: Decision systems) information systems of the form  $(U, A \cup \{d\})$  where  $d \notin A$  is a *decision*.

The most important problem concerning decision systems consists in finding a set of *decision rules* (see Glossary: Decision rules) which adequately describes the decision attribute in terms of attributes in  $A$  (*conditional attributes*).

Classification methods can be divided, according to the adopted methodology, into classifiers based on reducts and decision rules, classifiers based on templates and similarity, classifiers based on descriptor search, classifiers based on granular descriptors, and hybrid classifiers.

For a decision system  $(U, A, d)$ , classifiers are sets of decision rules. Induction of rules has been a subject of research in rough set theory since its beginning. In most general terms, building a classifier consists of searching the pool of descriptors for conjuncts that describe decision classes sufficiently well. As distinguished in Stefanowski [58], there are three main kinds of classifiers to search for: *minimal*, that is, consisting of minimum possible number of rules describing decision classes in the universe, *exhaustive*, that is, consisting of all possible rules, and *satisfactory*, that is, containing rules tailored to a specific use. Classifiers are evaluated globally with respect to their ability to properly classify objects, usually by *error* which is the ratio of the number of correctly classified objects to the number of test objects, *total accuracy* being the ratio of the number of correctly classified cases to the number of recognized cases, and *total coverage*, the ratio of the number of recognized test cases to the number of test cases.

Minimum size algorithms include the LEM2 algorithm of Grzymala-Busse [10,11] and the covering algorithm in the RSES package [51]; exhaustive algorithms include, for example, the LERS system of Grzymala-Busse [9], systems based on discernibility matrices and Boolean reasoning by Skowron [53], Bazan [3], Bazan et al. [4], implemented in the RSES package [51].

Minimal consistent sets of rules were introduced in Skowron and Rauszer [54]; they were shown to coincide with rules induced on the basis of local reducts in Wróblewski [63]. Further developments include dynamic rules, approximate rules, and relevant rules as described in [3,4] as well as local rules [3,4], effective in implementations of algorithms based on minimal consistent sets of rules. Rough set based classification algorithms, especially those implemented in the RSES system [51], were discussed extensively in [3]; Skowron and Stepaniuk [55] and Skowron and Swiniarski [56] discussed rough set classifi-

fiers along with some attempt at analysis of granulation in the process of knowledge discovery.

In [3], a number of techniques were verified in experiments with real data, based on various strategies:

discretization of attributes (codes: N-no discretization, S-standard discretization, D-cut selection by dynamic reducts, G-cut selection by generalized dynamic reducts);

dynamic selection of attributes (codes: N-no selection, D-selection by dynamic reducts, G-selection based on generalized dynamic reducts);

decision rule choice (codes: A-optimal decision rules, G-decision rules on basis of approximate reducts computed by Johnson's algorithm, simulated annealing and Boltzmann machines etc., N-without computing of decision rules);

approximation of decision rules (codes: N-consistent decision rules, P-approximate rules obtained by descriptor dropping);

negotiations among rules (codes: S-based on strength, M-based on maximal strength, R-based on global strength, D-based on stability).

Any choice of a strategy in particular areas yields a compound strategy denoted with the alias being concatenation of symbols of strategies chosen in consecutive areas, for example, NNAND etc.

We record here in Table 1 an excerpt from the comparison (Table 8, 9, 10 in [3]) of the best of these strategies with results based on other paradigms in classification for two sets of data: Diabetes and Australian credit from the UCI Repository [55].

An adaptive method of classifier construction was proposed in Wróblewski [63]; reducts are determined by means of a genetic algorithm and in turn reducts induce sub-tables of data regarded as classifying agents; choice of optimal ensembles of agents is done by a genetic algorithm.

Classifiers constructed by means of similarity relations are based on templates matching a given object, or closest to it, with respect to a certain distance function, or on coverings of the universe of objects by tolerance classes and assigning the decision value on basis of some of them [27]. We include in Table 2 excerpts from classification results in [27].

A combination of rough set methods with the k-nearest neighbor idea is a further refinement of classification based on similarity or analogy in Wojna [62]. In this approach, training set objects are endowed with a metric, and the test objects are classified by a vote of k nearest training objects for some k that is subject to optimization.

**Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Table 1**

A comparison of errors in classification by rough set and other paradigms

Paradigm	System/method	Diabetes	Austr. credit
Stat. Methods	Logdisc	0.223	0.141
Stat. Methods	SMART	0.232	0.158
Neural Nets	Backpropagation2	0.248	0.154
Neural Networks	RBF	0.243	0.145
Decision Trees	CART	0.255	0.145
Decision Trees	C4.5	0.270	0.155
Decision Trees	ITrule	0.245	0.137
Decision Rules	CN2	0.289	0.204
Rough Sets	NNANR	0.335	0.140
Rough Sets	DNANR	0.280	0.165
Rough Sets	best result	0.255(DNAPM)	0.130 (SNAPM)

**Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Table 2**

Accuracy of classification by template and similarity methods

Paradigm	System/method	Diabetes	Austr. credit
Rough Sets	Simple.templ./Hamming	0.6156	0.8217
Rough Sets	Gen.templ./Hamming	0.742	0.855
Rough Sets	Simple.templ./Euclidean	0.6312	0.8753
Rough Sets	Gen.templ./Euclidean	0.7006	0.8753
Rough Sets	Match.tolerance	0.757	0.8747
Rough Sets	Clos.tolerance	0.743	0.8246

**Granulation of Knowledge: General Ideas and Technique of Rough Inclusions**

Granulation of knowledge means that, given a representation of knowledge, for example, an information system, a fortiori, the set of information vectors of objects in its universe, these information vectors will undergo the process of aggregation, that is, of granule formation (granules being collections of information vectors). The process of granulation should be driven by a similarity relation on information vectors, and as the result of this process, similar to a satisfactory degree, information vectors/objects will find themselves in one single granule. Parameters of this process are an object  $u$  around which the granule is formed (the granule center) and the degree  $r \in [0, 1]$  to which objects in the granule are similar to  $u$ .

**A General Form of Similarity**

For our purposes of granulation of knowledge, we introduce a general form of graded similarity in the form a relation  $\tau \subseteq U \times U \times [0, 1]$  on triples of the form  $(u, v, r)$  with  $u, v \in U$ , the universe of an information system

$(U, A)$ , and  $r \in [0, 1]$ . We require of the similarity  $\tau$  the following:

1.  $\tau(u, u, 1)$ ;
2.  $r < s$  implies if  $\tau(u, v, s)$  then  $\tau(u, v, r)$ ;
3.  $\tau(u, v, 1)$  is a partial ordering of the universe  $U$ .

**Granulation: Tools**

Given a similarity  $\tau$ , we form a basic granule  $g_\tau(u, r)$  of the radius  $r$  about an object  $u$ , by collecting together all objects  $v$  with the property that  $\tau(v, u, r)$ . Then, by definition of  $\tau$ , one has the following properties of granules:

- a.  $u \in g_\tau(u, r)$  for each  $r \in [0, 1]$ ;
- b.  $v \in g_\tau(u, r)$  implies  $v \in g_\tau(u, s)$  for each  $s < r$ ;

Granulation thus requires:

- a similarity measure,
- an aggregation mechanism.

**Granulation by Means of Rough Inclusions**

In this section, we apply the general ideas derived in Sect. “Similarity” and Sect. “Information and Decision Systems” in a specific context of rough inclusions (see Glossary: Rough Inclusions).

**Mereological Theory of Concepts** As a first step to granulation by means of rough inclusions, we introduce the reader to the mereological theory of concepts which is the basis of our approach to granulation.

Mereology, as conceived by Stanisław Leśniewski [15], is based on the primitive binary relation of *part*,  $\pi$ , on the universe  $U$  of objects, which is subject to the following requirements:

- (P1)  $u \pi v$  implies  $u \neq v$ ;  
 (P2)  $u \pi v$  and  $v \pi w$  imply  $u \pi w$ .

Thus, the part relation  $\pi$  is non-reflexive and transitive.

Given a part relation, which by (P1) means a proper part relation, one forms an improper part counterpart, called the ingredient, *ing*, defined from the part relation as follows:

- (I)  $v \text{ ing } u$  if and only if  $v \pi u$  or  $v = u$ .

Thus, the ingredient relation is a partial ordering on the universe  $U$ :

- (U1)  $u \text{ ing } u$  for each  $u$ ;  
 (U2)  $u \text{ ing } v$  and  $v \text{ ing } u$  imply  $v = u$ ;  
 (U3)  $u \text{ ing } v$  and  $v \text{ ing } w$  imply  $u \text{ ing } w$ .

Mereology provides an operator, the *class operator*, whose purpose is to represent collections of objects as an individual object; clearly, this will be also our principal granulation tool when formal properties of granules are involved. The class operator  $\text{Cls}$  is applied to any non-empty collection (a property)  $\Phi$  of individual objects in  $U$ , making this collection into an individual  $\text{Cls } \Phi$ ; its definition is as follows Leśniewski [15]:

- (Cl1) If  $v \in \Phi$  then  $v \text{ ing } \text{Cls } \Phi$ ;  
 (Cl2) If  $v \text{ ing } \text{Cls } \Phi$  then for some  $w, z \in U$  one has that  $w \text{ ing } v, w \text{ ing } z, z \in \Phi$ .

The object  $\text{Cls } \Phi$  is unique for a given  $\Phi$ .

As an example, the reader will verify that the relation  $\subset$  of the proper set/concept containment is a part relation, the relation  $\subseteq$  of the improper containment is the adjoint relation of an ingredient, and given a non-empty family of sets  $\Phi$ , the union  $\bigcup \Phi$  is the class  $\text{Cls } \Phi$  with respect to the ingredient relation  $\subseteq$ . Thus, the mereological context generalizes the standard set-theoretical context adopted prevalently by computer science.

**Rough Inclusions** In the process of development of rough set theory, it has turned out that indiscernibility should be relaxed to similarity: in Polkowski, Skowron and Zytkow [50], attention was focused on tolerance relations, that is, relations which are reflexive and symmetric but need not be transitive. An example of such a relation was given in Poincaré [35]: given a metric  $\rho$  and a fixed small positive  $\delta$ , one declares points  $x, y$  in the relation  $\text{sim}(\delta)$  if and only if  $\rho(x, y) < \delta$ . The relation  $\text{sim}(\delta)$  is a tolerance relation but it is equivalence for non-archimedean  $\rho$ 's only, that is, when  $\rho(x, y) \leq \max\{\rho(x, z), \rho(z, y)\}$ .

We continue this example by introducing a graded version of  $\text{sim}(\delta)$ , viz., for a real number  $r \in [0, 1]$ , we define the relation  $\text{sim}(\delta, r)$  by letting,

$$\text{sim}(\delta, r)(x, y) \text{ iff } \rho(x, y) \leq 1 - r. \quad (1)$$

The collection  $\text{sim}(\delta, r)$  of relations have the following properties evident by the properties of the metric  $\rho$ :

- (SIM1)  $\text{sim}(\delta, 1)(x, y)$  iff  $x = y$ ;  
 (SIM2)  $\text{sim}(\delta, 1)(x, y)$  and  $\text{sim}(\delta, r)(z, x)$  imply  $\text{sim}(\delta, r)(z, y)$ ;  
 (SIM3)  $\text{sim}(\delta, r)(x, y)$  and  $s < r$  imply  $\text{sim}(\delta, s)(x, y)$ .

Properties (SIM1)–(SIM3) induced by the metric  $\rho$  refer to the ingredient relation = whose corresponding relation of part is empty; a generalization can thus be obtained by replacing the identity with an ingredient relation *ing* in a mereological universe  $(U, \pi)$ .

Consequently, a relation  $\mu(u, v, r)$  is defined that satisfies the following conditions:

- (RM1)  $\mu(u, v, 1)$  iff  $u \text{ ing } v$ ;  
 (RM2)  $\mu(u, v, 1)$  and  $\mu(w, u, r)$  imply  $\mu(w, v, r)$ ;  
 (RM3)  $\mu(u, v, r)$  and  $s < r$  imply  $\mu(u, v, s)$ .

Any relation  $\mu$  which satisfies the conditions (RM1)–(RM3) is called a *rough inclusion*. This relation is a similarity relation which is not necessarily symmetric, but is reflexive. It is read as “the relation of a part to a degree”.

**Rough Inclusions: Case of Information Systems** The problem of methods by which rough inclusions could be introduced in information/decision systems has been studied in Polkowski [37,38,39,40,41,42,43,44]. Here we recapitulate the results and add new ones. We recall that an *information system* is a method of representing knowledge about a certain phenomenon in the form of a table of data; formally, it is a pair  $(U, A)$  where  $U$  is a set of *objects* and  $A$  is a set of *conditional attributes*; any object  $u \in U$  is described by means of its *information set*  $\text{Inf}(u) = \{(a, a(u)): a \in A\}$ . The indiscernibility relation  $\text{Ind}$ , definable sets and non-definable sets are defined from  $\text{Inf}$  as indicated in (Glossary: Information Systems).

**Rough Inclusions from Metrics** As observed in Sect. “Introduction”, any metric  $\rho$  defines a rough inclusion  $\mu_\rho$  by means of the equivalence  $\mu_\rho(u, v, r) \Leftrightarrow \rho(u, v) \leq 1 - r$ .

A very important example of a rough inclusion obtained on these lines is the rough inclusion  $\mu_h$  with  $h(u, v)$  being the reduced Hamming distance on information vectors of  $u$  and  $v$ , that is,  $h(u, v) = \frac{|\{a \in A: (a, a(u)) \neq (a, a(v))\}|}{|A|}$ ,  $|X|$  denoting the cardinality of the set  $X$ .

Thus,  $\mu_h(u, v, r)$  if and only if  $h(u, v) \leq 1 - r$ ; introducing sets  $\text{DIS}(u, v) = \{a \in A: (a, a(u)) \neq (a, a(v))\}$  and  $\text{IND}(u, v) = A \setminus \text{DIS}(u, v) = \{a \in A: a(u) = a(v)\}$ , one can write down the formula for  $\mu_h$  either as,

$$\mu_h(u, v, r) \Leftrightarrow \frac{|\text{DIS}(u, v)|}{|A|} \leq 1 - r, \quad (2)$$

or,

$$\mu_h(u, v, r) \Leftrightarrow \frac{|\text{IND}(u, v)|}{|A|} \geq r. \quad (3)$$

The formula (3) witnesses that the rough inclusion  $\mu_h$  is an extension of the indiscernibility relation  $\text{Ind}$  to a graded indiscernibility.

In a similar manner one should be able to compute rough inclusions induced by other metrics standardly used on information sets such as Euclidean, Manhattan etc.

Rough inclusions induced by metrics possess an important property of *functional transitivity* expressed in general form by the rule,

$$\frac{\mu_\rho(u, v, r), \mu_\rho(v, w, s)}{\mu_\rho(u, w, L(r, s))}, \quad (4)$$

where  $L(r, s) = \max\{0, r + s - 1\}$  is the Łukasiewicz t-norm, see Polkowski [66]. We offer a short proof of this fact: assuming that  $\mu_\rho(u, v, r), \mu_\rho(v, w, s)$  which means in terms of the metric  $\rho$  that  $\rho(u, v) \leq 1 - r, \rho(v, w) \leq 1 - s$ ; by the triangle inequality,  $\rho(u, w) \leq (1 - r) + (1 - s)$ , that is,  $\mu_\rho(u, w, r + s - 1)$ .

**Rough Inclusions from Functors of Many-Valued Logics** A functor (t-norm)  $t: [0, 1] \times [0, 1] \rightarrow [0, 1]$  is *non-archimedean* only when the equality  $t(x, x) = x$  holds for  $x = 0, 1$ ; it is known (see e.g. [36]) that the only such t-norms are the Łukasiewicz  $L$  and the product t-norm  $P(x, y) = x \cdot y$ .

Each of these t-norms admits a functional representation:  $t(x, y) = g(f(x) + f(y))$  (see, for example, Polkowski [36]).

One defines a rough inclusion  $\mu_t$  by letting (see Polkowski [35,37,38,39,40,41,42,43]),

$$\mu_t(u, v, r) \Leftrightarrow g\left(\frac{|\text{DIS}(u, v)|}{|A|}\right) \geq r. \quad (5)$$

In particular, in case of the t-norm  $L$ , one has  $g(x) = 1 - x$  (see, for example, [36]), and thus the rough inclusion  $\mu_L$  is expressed by means of the formula (3).

Another systematic method for defining rough inclusions is by means of residual implications of continuous t-norms (see Polkowski [38,39,40,41,42,43,44]).

For a continuous t-norm  $t$ , the *residual implication*  $x \Rightarrow_t y$  is a mapping from the square  $[0, 1]^2$  into  $[0, 1]$  defined as follows (see, for example, Polkowski [36]),

$$x \Rightarrow_t y \geq z \text{ iff } t(x, z) \leq y; \quad (6)$$

thus,  $x \Rightarrow_t y = \max\{z: t(x, z) \leq y\}$ .

**Proposition 1** *The residual implication  $x \Rightarrow_t y$  does induce a rough inclusion  $\mu_t^{\Rightarrow}$  by means of the formula:  $\mu_t^{\Rightarrow}(x, y, r)$  if and only if  $x \Rightarrow_t y \geq r$  for every continuous t-norm  $t$ .*

We include a short argument for the sake of completeness; clearly,  $\mu_t^{\Rightarrow}(x, x, 1)$  holds as  $x \Rightarrow_t y \geq 1$  is equivalent to  $x \leq y$ . Assuming  $\mu_t^{\Rightarrow}(x, y, 1)$ , that is,  $x \leq y$ , and  $\mu_t^{\Rightarrow}(z, x, r)$ , that is,  $z \Rightarrow x \geq r$  hence  $t(z, r) \leq x$  we have  $t(z, r) \leq y$ , that is,  $z \Rightarrow y \geq r$  so finally  $\mu_t^{\Rightarrow}(x, y, r)$ . Clearly, by definition, from  $\mu_t^{\Rightarrow}(x, y, r)$  and  $s < r$  it does follow that  $\mu_t^{\Rightarrow}(x, y, s)$ .

We recall here the basic cases of rough inclusions obtained from the most frequently applied t-norms. In all cases,  $\mu_t^{\Rightarrow}(x, y, 1)$  if  $x \leq y$ , so the associated part relation is  $\leq$  and the underlying part relation is  $<$ . For  $r < 1$ , that is,  $x > y$ , one has

Case 1  $t = L$ ; in this case  $x \Rightarrow_L y = \min\{1, 1 - x + y\}$ , hence  $\mu_L^{\Rightarrow}(x, y, r)$  if and only if  $1 - x + y \geq r$ .

Case 2  $t = P$  where  $P(x, y) = x \cdot y$ ; in this case,  $x \Rightarrow_P y = \frac{y}{x}$  when  $x \neq 0$  and 1 when  $x = 0$  hence  $\mu_P^{\Rightarrow}(x, y, r)$  if and only if  $y \geq x \cdot r$ .

Case 3  $t = \min(x, y)$ ; in this case  $x \Rightarrow_{\min} y$  is  $y$  hence  $\mu(x, y, r)$  if and only if  $y \geq r$ .

**Transitivity of Rough Inclusions** It has been proved [37,38,39,40,41,42,43,44] that all rough inclusions induced from either non-archimedean or continuous t-norms  $t$  in the manner as above are transitive in the sense of the formula:

$$\frac{\mu(u, v, r), \mu(v, w, s)}{\mu(u, w, t(r, s))}.$$

**Applications to Granulation of Knowledge** Formal theory of rough inclusions allows for a formal mechanism of granulation of knowledge; we assume an information system  $(U, A)$  is given. Granulation of knowledge, proposed as a paradigm by L.A. Zadeh [66], means grouping objects into collections called granules, objects within a granule being similar with respect to a chosen measure; granular computing means computing with granules in place of objects.

The mechanism of granule formation based on rough inclusions has been presented by the author in a few works (see [37,38,39,40,41,42,43,44]) and we recall it here. We assume an information system  $(U, A)$  is given. The basic tool in establishing properties of granules is the class operator of mereology, along with the Leśniewski deduction rule (DRM) of Sect. “Introduction”.

Given a rough inclusion  $\mu$  on the universe  $U$ , for each object  $u$  and each  $r \in [0, 1]$ , the granule  $g_\mu(u, r)$  of the radius  $r$  about  $u$  relative to  $\mu$  is defined as the class of the property  $\Phi(u, r, \mu) = \{v: \mu(v, u, r)\}$ :

$$g_\mu(u, r) \text{ is } \text{Cls } \Phi(u, r, \mu). \quad (7)$$

In case of t-norm-induced rough inclusions, by their transitivity, the following important property holds (see Polkowski [43]):

$$v \text{ ing } g_\mu(u, r) \text{ iff } \mu(v, u, r), \quad (8)$$

That is, the granule  $g_\mu(u, r)$  is  $\{v: \Phi(u, r, \mu)(v)\}$  with no synergy effect.

**Rough Inclusions on Granules** because mereology operates (due to the class operator) only on the level of individuals, one can extend rough inclusions from objects to granules. The formula for extending a rough inclusion  $\mu$  to a rough inclusion  $\bar{\mu}$  on granules is a modification of the mereological deduction rule of Sect. “Introduction”: (DRMG):  $\bar{\mu}(g, h, r)$  if and only if for each object  $z$  with  $z$  ing  $g$  there exists an object  $w$  with  $w$  ing  $h$  such that  $\mu(z, w, r)$ .

The fact that  $\bar{\mu}$  is a rough inclusion can be established by means of the rule (DRMG).

**Modifications and Variants of Rough Inclusions** In applications to be presented in this special section, some modified rough inclusions or weaker similarity measures will be instrumental. We include a discussion of them here.

**Modification by Means of Metrics on Attribute Values** For the rough inclusion  $\mu_L$ , the formula  $\mu_L(v, u, r)$  means that  $\frac{|\text{IND}(v, u)|}{|A|} \geq r$ , that is, at least  $r \cdot 100$  percent of attributes agree on  $u$  and  $v$ ; an extension of this rough inclusion depends on a chosen metric  $\rho$  bounded by 1 in the attribute value space  $V$  (we assume a simple case that  $\rho$  works for each attribute).

Then, given an  $\varepsilon \in [0, 1]$ , we let  $\mu^\varepsilon(v, u, r)$  if and only if  $|\{a \in A: \rho(a(v), a(u)) < \varepsilon\}| \geq r \cdot |A|$ ; it is manifest that  $\mu^\varepsilon$  is a rough inclusion if  $\rho$  is a non-archimedean met-

ric, that is,  $\rho(u, w) \leq \max\{\rho(u, v), \rho(v, w)\}$ ; otherwise the monotonicity condition rm2 of Sect. “Calculus on Granules” need not be satisfied and this takes place with most popular metrics (Euclidean, Manhattan etc.).

In this case, a remedy is to define a rough inclusion  $\mu^*$  as follows:  $\mu^*(v, u, r)$  if and only if there exists an  $\varepsilon$  such that  $\mu^\varepsilon(v, u, r)$ . It is easy, then, to verify that  $\mu^*$  is a rough inclusion.

**Modification by transfer** Assume that  $\Rightarrow_t$  is chosen, and for an information system  $(U, A)$ , with an ingredient relation  $\text{ing}$  on  $U$ , a mapping  $\phi: U \rightarrow [0, 1]$  is given such that  $\phi(u) \leq \phi(v)$  if and only if  $u$  ing  $v$ . Then, the relation,

$$\mu_\phi(v, u, r) \text{ iff } \phi(u) \Rightarrow_t \phi(v) \geq r, \quad (9)$$

is a rough inclusion on  $U$ . We include a short proof of this fact: that  $\mu_\phi(u, v, 1)$  is equivalent to  $\phi(u) \leq \phi(v)$  hence to  $u$  ing  $v$  was observed in Subsect. “Granules from rough inclusions – mereological approach” b. Assuming that  $\mu_\phi(u, v, 1)$  and  $\mu_\phi(w, u, r)$ , the proof that (RM2) (see Sect. “Granulation of Knowledge: General Ideas and Technique of Rough Inclusions”) holds, that is,  $\mu_\phi(w, v, r)$  proceeds like the proof of (RM2) in Proposition 1 of Sect. “Granulation of Knowledge: General Ideas and Technique of Rough Inclusions”. Finally, the property (RM3) is evident.

As  $\phi$  depends on one argument, let us select an object  $s \in U$ , and consider:

1.  $\phi_1 = \text{dis}(u) = \frac{|\{a \in A: a(s) \neq a(u)\}|}{|A|}$ ;
2.  $\phi_2 = \text{ind}(u) = \frac{|\{a \in A: a(s) = a(u)\}|}{|A|}$ ;
3.  $\phi_3 = \text{dis}_\varepsilon(u) = \frac{|\{a \in A: \rho(a(s), a(u)) \geq \varepsilon\}|}{|A|}$ ;
4.  $\phi_4 = \text{ind}_\varepsilon(u) = \frac{|\{a \in A: \rho(a(s), a(u)) \leq \varepsilon\}|}{|A|}$ , where  $\rho$  is a chosen metric on the set of attribute values  $V$ , and  $\varepsilon$  is a chosen threshold in  $[0, 1]$ .

**Proposition 2** *In all cases  $i = 1, 2, 3, 4$ , the relation  $\mu_{\phi_i}(v, u, r)$  defined with  $\phi_i$  as  $\phi$  is a rough inclusion.*

The reference object  $s$  can be chosen as an “ideal object,” for example, whose conditional class is contained in its decision class, etc. The formula (9) can be generalized by considering a set  $S = \{s_1, \dots, s_k\}$  of reference objects.

Comparison of objects  $u, v$  on the lines of this section, do not necessarily lead to rough inclusions due to a possible violation of property (RM2); yet, such variants are of importance as they allow for a direct comparison among objects, rules and granules. We introduce a generalization of rough inclusions:

For given objects  $u, v$ , and  $\varepsilon \in [0, 1]$ , we introduce the factors:

$$\text{dis}_\varepsilon(u, v) = \frac{|\{a \in A: \rho(a(u), a(v)) \geq \varepsilon\}|}{|A|},$$

and

$$\text{ind}_\varepsilon(u, v) = \frac{|\{a \in A: \rho(a(u), a(v)) < \varepsilon\}|}{|A|},$$

where  $\rho$  is a metric on attribute value sets.

Then, we modify the formula (9) to the form,

$$\nu(u, v, r) \text{ iff } \text{dis}_\varepsilon(u, v) \rightarrow_t \text{ind}_\varepsilon(u, v) \geq r. \quad (10)$$

Clearly,  $\nu$  has properties: 1.  $\nu(u, u, 1)$ ; 2.  $\nu(u, v, r)$  and  $s < r$  imply  $\nu(u, v, s)$  but monotonicity property (RM2) need not hold. Rough inclusions and their weak variants will be essentially exploited in data mining tasks presented in the sequel.

### Calculus on Granules

We recall that we define a granule  $g_\mu(u, r)$  about  $u \in U$  of the radius  $r$ , relative to the rough inclusion  $\mu$ , as follows:

$$g_\mu(u, r) \text{ is } \text{Cls}_\mu \Pi^\mu(u, r), \quad (11)$$

where  $\Pi^\mu(u, r)$  is the property defined by means of,

$$\Pi^\mu(u, r)(v) \text{ iff } \mu(v, u, r). \quad (12)$$

General properties of granules are collected below,

- 1. if  $y \text{ ing } x$  then  $y \text{ ing } g_{r x}$ ;
- 2. if  $y \text{ ing } g_{r x}$  and  $z \text{ ing } y$  then  $z \text{ ing } g_{r x}$ ;
- 3. if  $\mu(y, x, r)$  then  $y \text{ ing } g_{r x}$ ;
- 4. if  $s < r$  then  $g_{r x} \text{ ing } g_{s x}$ ,

which follow from properties (RM1)–(RM3) of rough inclusions (see sect. V) and the fact that  $\text{ing}$  is a partial order, and, in particular, is transitive.

More generally,

$$\text{if } y \text{ ing } g_{r x} \text{ then } g_s y \text{ ing } g_{T(r,s)x}. \quad (14)$$

The last statement follows directly from the transitivity property of rough inclusions and from the deduction rule (DRM) of Sect. “Introduction”.

It is natural to regard a granule system  $\{g_r^{\mu_t}(x): x \in U; r \in (0, 1)\}$  as a neighborhood system for a topology on  $U$  that may be called the *granular topology*; we refer here to the idea of a neighborhood system introduced by Lin [16,17].

In order to make this idea explicit, we define classes of the form  $N^T(x, r) = \text{Cls}(\psi_{r,x}^{\mu_T})$ , where

$$\psi_{r,x}^{\mu_T}(y) \Leftrightarrow \exists s > r. \mu_T(y, x, s). \quad (15)$$

We declare the system  $\{N^T(x, r): x \in U; r \in (0, 1)\}$  to be a neighborhood basis for a topology  $\theta_\mu$ . This is justified by the following

**Theorem 1** *Here are properties of the system  $\{N^T(x, r): x \in U; r \in (0, 1)\}$ :*

1.  $y \text{ ingr } N^t(x, r) \Rightarrow \exists \delta > 0. N^t(y, \delta) \text{ ingr } N^t(x, r);$
2.  $s > r \Rightarrow N^t(x, s) \text{ ingr } N^t(x, r);$
3.  $z \text{ ingr } N^t(x, r) \wedge z \text{ ingr } N^t(y, s) \Rightarrow \exists \delta > 0. N^t(z, \delta) \text{ ingr } N^t(x, r) \wedge N^t(z, \delta) \text{ ingr } N^t(y, s).$

An argument for (16) is as follows:

For Property 1.  $y \text{ ingr } N^t(x, r)$  implies that there exists an  $s > r$  such that  $\mu_t(y, x, s)$ . Let  $\delta < 1$  be such that  $t(u, s) > r$  whenever  $u > \delta$ ;  $\delta$  exists by continuity of  $t$  and the identity  $t(1, s) = s$ . Thus, if  $z \text{ ingr } N^t(y, \delta)$ , then  $\mu_t(z, y, \eta)$  with  $\eta > \delta$  and  $\mu_t(z, x, t(\eta, s))$  hence  $z \text{ ingr } N^t(x, r)$ .

Property 2. follows by (RM3) and Property 3 is a corollary to properties 1 and 2. This concludes the argument for (16).

Granule systems as defined above form a basis for applications where approximate reasoning is a crucial ingredient.

We begin with a basic application in which approximate reasoning itself is codified as a many-world (intensional) logic where granules serve as possible worlds.

### Some Principal Fields of Applications

Applications of the proposed mechanism of granulation have been indicated in many areas of reasoning under uncertainty, in particular in:

- Reasoning about uncertainty in information/decision systems by means of rough mereological granular logics;
- Reasoning in distributed systems (fusion of knowledge);
- Reasoning in cognitive schemes (rough neural hybrid systems);
- Granular classifiers.

We devote some space to these applications.

## Reasoning About Uncertainty

The idea of a granular rough mereological logic (see [38, 48]) consists of measuring the meaning of a unary predicate in the model, which is a universe of an information system, against a granule defined to a certain degree by means of a rough inclusion. The result can be regarded as the degree of truth (the logical value) of the predicate with respect to the given granule. The logics thus obtained are intensional as they can be regarded as mappings from the set of granules (possible worlds) to the set of logical values in the interval  $[0, 1]$ , the value at a given granule regarded as the extension at that granule of the generally defined intension (see [60] for a general introduction to intensional logics).

For our purpose it is essential to extend rough inclusions to sets; we use the t-norm  $L$  along with the representation  $L(r, s) = g(f(r) + f(s))$  already mentioned. We denote these kind of inclusions with the generic symbol  $\nu$ .

For finite sets  $X, Y$ , we let,

$$\nu_L(X, Y, r) \text{ iff } g\left(\frac{|X \setminus Y|}{|X|}\right) \geq r; \quad (17)$$

as  $g(x) = 1 - x$ , see Polkowski [7], we have that  $\nu_L(X, Y, r)$  holds if and only if  $\frac{|X \cap Y|}{|X|} \geq r$ . Let us observe that  $\nu_L$  is *regular*, that is,  $\nu_L(X, Y, 1)$  if and only if  $X \subseteq Y$  and  $\nu_L(X, Y, r)$  only with  $r = 0$  if and only if  $X \cap Y = \emptyset$ .

Thus, the ingredient relation associated with a regular rough inclusion is the improper containment  $\subseteq$ , whereas the underlying part relation is the strict containment  $\subset$ .

Other rough inclusion on sets we exploit is the 3-valued rough inclusion  $\nu_3$  defined via the formula, see Polkowski [38],

$$\nu_3(X, Y, r) \text{ iff } \begin{cases} X \subseteq Y \text{ and } r = 1 \\ X \cap Y = \emptyset \text{ and } r = 0 \\ r = \frac{1}{2} \text{ otherwise.} \end{cases} \quad (18)$$

We assume that an information/decision system  $(U, A, d)$  is given, along with a rough inclusion  $\nu$  on the subsets of the universe  $U$ ; for a collection of predicates (unary)  $Pr$ , interpreted in the universe  $U$  (meaning that for each predicate  $\phi \in Pr$  the meaning  $[\phi]$  is a subset of  $U$ ), we define the intensional logic  $grm_\nu$  on  $Pr$  by assigning to each predicate  $\phi$  in  $Pr$  its intension  $I_\nu(\phi)$  defined by its extension  $I_\nu^\vee(g)$  at particular granules  $g$ , as,

$$I_\nu^\vee(g)(\phi) \geq r \text{ iff } \nu(g, [\phi], r). \quad (19)$$

With respect to the rough inclusion  $\nu_L$ , the formula (19) becomes,

$$I_{\nu_L}^\vee(g)(\phi) \geq r \text{ iff } \frac{|g \cap [\phi]|}{|g|} \geq r. \quad (20)$$

The counterpart for  $\nu_3$  is specified by definition (18).

We say that a formula  $\phi$  interpreted in the universe  $U$  of an information system  $(U, A)$  is *true* at a granule  $g$  with respect to a rough inclusion  $\nu$  if and only if  $I_\nu^\vee(g)(\phi) = 1$ .

Thus, for every regular rough inclusion  $\nu$ , a formula  $\phi$  interpreted in the universe  $U$ , with meaning  $[\phi]$ , is true at a granule  $g$  with respect to  $\nu$  if and only if  $g \subseteq [\phi]$ . In particular, for a decision rule  $r: p \Rightarrow q$  in the descriptor logic, the rule  $r$  is true at a granule  $g$  with respect to a regular rough inclusion  $\nu$  if and only if  $g \cap [p] \subseteq [q]$ .

The formula  $\nu(g, [\phi], r) = 1$  stating the truth of  $\phi$  at  $g, \nu$  with  $\nu$  regular can be regarded as a condition of orthogonality type, with the usual consequences.

1. If  $\phi$  is true at granules  $g, h$  then it is true at  $g \cup h$ .
2. If  $\phi$  is true at granules  $g, h$  then it is true at  $g \cap h$ .
3. If  $\phi, \psi$  are true at a granule  $g$  then  $\phi \vee \psi$  is true at  $g$ .
4. If  $\phi, \psi$  are true at a granule  $g$  then  $\phi \wedge \psi$  is true at  $g$ .
5. If  $\psi$  is true at a granule  $g$  then  $\phi \Rightarrow \psi$  is true at  $g$  for every formula  $\phi$ .
6. If  $\phi$  is true at a granule  $g$  then  $\phi \Rightarrow \psi$  is true at  $g$  if and only if  $\psi$  is true at  $g$ .

The graded relaxation of truth is given obviously by the condition, a formula  $\phi$  is *true to a degree at least  $r$*  at  $g, \nu$  if and only if  $I_\nu^\vee(g)(\phi) \geq r$ , that is,  $\nu(g, [\phi], r)$  holds. In particular,  $\phi$  is *false* at  $g, \nu$  if and only if  $I_\nu^\vee(g)(\phi) \geq r$  implies  $r = 0$ , i. e.  $\nu(g, [\phi], r)$  implies  $r = 0$ . Properties 1–6 above suggest that the notion of truth in rough mereological logics has similar properties with respect to connectives of logical calculi to those of classical sentential calculus. Therefore, we introduce the following semantics of sentential connectives  $\neg, \vee, \wedge, \Rightarrow$ .

1.  $[\neg\alpha] = U \setminus [\alpha]$ .
2.  $[\alpha \vee \beta] = [\alpha] \cup [\beta]$ .
3.  $[\alpha \wedge \beta] = [\alpha] \cap [\beta]$ .
4.  $[\alpha \Rightarrow \beta] = (U \setminus [\alpha]) \cup [\beta]$ .

With respect to these semantics, the following properties hold.

1. For each regular  $\nu$ , a formula  $\alpha$  is true at  $g, \nu$  if and only if  $\neg\alpha$  is false at  $g, \nu$ .
2. For  $\nu = \nu_L, \nu_3$ ,  $I_\nu^\vee(g)(\neg\alpha) \geq r$  if and only if  $I_\nu^\vee(g)(\alpha) \geq s$  implies  $s \leq 1 - r$ .
3. For  $\nu = \nu_L, \nu_3$ , the implication  $\alpha \Rightarrow \beta$  is true at  $g$  if and only if  $g \cap [\alpha] \subseteq [\beta]$  and  $\alpha \Rightarrow \beta$  is false at  $g$  if and only if  $g \subseteq [\alpha] \setminus [\beta]$ .
4. For  $\nu = \nu_L$ , if  $I_\nu^\vee(g)(\alpha \Rightarrow \beta) \geq r$  then  $\Rightarrow_L(t, s) \geq r$  where  $I_\nu^\vee(g)(\alpha) \geq t$  and  $I_\nu^\vee(g)(\beta) \geq s$ .

The functor  $\Rightarrow_L$  in 4. is the Łukasiewicz implication of many-valued logic:  $\Rightarrow_L(t, s) = \min\{1, 1 - t + s\}$ .

Further analysis should be split into the case of  $v_L$  and the case of  $v_3$  as the two differ essentially with respect to the form of reasoning they imply.

**Reasoning with  $v_L$**  The last property 4. shows in principle that the value of  $I_v^\vee(g)(\alpha \Rightarrow \beta)$  is bounded from above by the value of  $\Rightarrow_L(I_v^\vee(g)(\alpha), I_v^\vee(g)(\beta))$ .

This suggests that the idea of collapse due to S. Lesniewski can be applied to formulas of rough mereological logic in the following form: for a formula  $q(x)$  we denote by the symbol  $q^*$  the formula  $q$  regarded as a sentential formula (that is, with variable symbols removed) subject to relations  $\neg q(x)^*$  is  $\neg(q(x)^*)$  and  $p(x) \Rightarrow q(x)^*$  is  $p(x)^* \Rightarrow q(x)^*$ . As the value  $[q^*]_g$  of the formula  $q(x)^*$  we admit the value of  $\frac{|g \cap [q(x)]|}{|g|}$ . Thus, the item 4 can be rewritten in the form.

$$I_v^\vee(g)(\alpha \Rightarrow \beta) \leq \Rightarrow_L([q^*]_g, [\beta^*]_g). \quad (21)$$

The following statement is then obvious: if  $\alpha \Rightarrow \beta$  is true at  $g$  then the collapsed formula has a truth value of 1 at the granule  $g$ . This gives a necessity condition for verification of implications of rough mereological logic: if  $\Rightarrow_L([\alpha^*]_g, [\beta^*]_g) < 1$  then the implication  $\alpha \Rightarrow \beta$  is not true at  $g$ . This concerns decision rules in particular: for a decision rule  $p(v) \Rightarrow q(v)$ , it follows that the decision is true on a granule  $g$  if and only if  $[p^*]_g \leq [q^*]_g$ .

**Possibility and Necessity** Possibility and necessity are introduced in rough set theory by means of upper and the lower approximations, respectively. A logical rendering of these modalities in rough mereological logic exploits the approximations. Denoting the lower approximation as  $\underline{X} = \{u \in U : [u]_A \subseteq X\}$  and the upper approximation as  $\overline{X} = \{u \in U : [u]_A \cap X \neq \emptyset\}$ , we define two modal operators: M (possibility) and L (necessity) by means of their semantics.

To this end, we let

$$\begin{aligned} I_v^\vee(g)(M\alpha) &\geq r \text{ iff } v_L(g, \overline{[\alpha]}, r) \\ I_v^\vee(g)(L\alpha) &\geq r \text{ iff } v_L(g, \underline{[\alpha]}, r). \end{aligned} \quad (22)$$

Then we have the following criteria for necessarily or possibly true formulas.

A formula  $\alpha$  is *necessarily true at a granule  $g$*  if and only if  $g \subseteq [\alpha]$ ;  $\alpha$  is *possibly true at  $g$*  if and only if  $g \subseteq \overline{[\alpha]}$ .

This semantics of modal operators  $M, L$  can be applied to show that rough set structures carry the semantics of S5 modal logic, that is, the following relations hold at each granule  $g$ .

#### Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Table 3

Truth values for implication in  $L_3$

$\Rightarrow$	0	1	$\frac{1}{2}$
0	1	1	1
1	0	1	$\frac{1}{2}$
$\frac{1}{2}$	$\frac{1}{2}$	1	1

#### Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Table 4

Truth values for implication  $\alpha \Rightarrow \beta$  in logic based on  $v_3$

$\Rightarrow$	$I_v^\vee(g)(\beta) = 0$	$I_v^\vee(g)(\beta) = 1$	$I_v^\vee(g)(\beta) = \frac{1}{2}$
$I_v^\vee(g)(\alpha) = 0$	1	1	1
$I_v^\vee(g)(\alpha) = 1$	0	1	$\frac{1}{2}$
$I_v^\vee(g)(\alpha) = \frac{1}{2}$	$\frac{1}{2}$	1	1 when $g \cap [\alpha] \subseteq [\beta]$ ; $\frac{1}{2}$ otherwise

1.  $L(\alpha \Rightarrow \beta) \Rightarrow [(L\alpha \Rightarrow L(\beta))]$ .
2.  $L\alpha \Rightarrow \alpha$ .
3.  $L\alpha \Rightarrow LL\alpha$ .
4.  $M\alpha \Rightarrow LM\alpha$ .

Proofs can be found in Polkowski [38].

**Reasoning with  $v_3$**  In case of  $v = v_3$ , one can check on the basis of definitions that  $I_v^\vee(g)(\neg\alpha) \geq r$  if and only if  $I_v^\vee(g)(\alpha) \leq 1-r$ ; thus the negation functor in rough mereological logic based on  $v_3$  is the same as the negation functor in the 3-valued Łukasiewicz logic. For implication, the relations between rough mereological and 3-valued logics  $L_3$  follow from tables of values of truth.

Table 3 shows truth values of implications for the 3-valued logic  $L_3$  and Table 4 shows truth values of implications for rough mereological logic based on  $v_3$ .

To express the relation between the two implications, we introduce a new notion: we say that a formula  $\phi$  is *acceptable* in either logic ( $L_3$  or  $grm_{v_3}$ ) at a granule  $g$  if and only if  $[\phi^*]_g \geq \frac{1}{2}$ , respectively,  $I_v^\vee(g)(\phi) \geq \frac{1}{2}$ .

From Tables 3, 4 one infers that  $I_v^\vee(g)(\phi) \geq [\phi^*]_g$ . This crucial relationship implies that: if  $\phi^*$  is acceptable at  $g$  then  $\phi$  is acceptable at  $g$ ; if  $\phi^*$  is true at  $g$  then  $\phi$  is true at  $g$ . Also, if  $\phi$  at false at  $g$  then  $\phi^*$  is false at  $g$ .

This fact allows for a sufficiency criterion: if  $\phi^*$  is a theorem of the logic  $L_3$  than  $\phi$  is a true formula of the logic  $grm_{v_3}$ . This fact supplies a characterization of decision rules as well.

#### Reasoning in Distributed Systems

We now begin with the first of the principal applications of the apparatus developed so far (cf., [49]). Rough inclu-

sions and granular intensional logics based on them can be applied to describe workings of a collection of intelligent agents, which we will call granular agents, as they are endowed with granular logic.

A granular agent  $ag$  in its simplest form is a tuple

$$ag^* = (U_{ag}, A_{ag}, \mu_{ag}, \text{Pred}_{ag}, \text{UncProp}_{ag}, \text{GSynt}_{ag}, \text{LSynt}_{ag}),$$

where,

1.  $(U_{ag}, A_{ag}) = I_{ag}$  is an information system of the agent  $ag$ ;
2.  $\mu_{ag}$  is a rough inclusion induced from  $I_{ag}$ ;
3.  $\text{Pred}_{ag}$  is a set of first-order predicates interpreted in  $U_{ag}$ ;
4.  $\text{UncProp}_{ag}$  is the function that describes how uncertainty measured by rough inclusions at agents connected to  $ag$  propagates to  $ag$ ;
5. The operator  $\text{GSynt}_{ag}$ , the granular synthesizer at  $ag$ , takes granules sent to the agent from agents connected to it, and makes those granules into a granule at  $ag$ ;
6.  $\text{LSynt}_{ag}$ , the logic synthesizer at  $ag$ , takes formulas sent to the agent  $ag$  by its connecting neighbors and makes them into a formula describing objects at  $ag$ .

A network of granular agents is a directed acyclic graph  $N = (Ag, C)$ , where  $Ag$  is its set of vertices, that is, granular agents, and  $C$  is the set of edges, that is, connections among agents, along with disjoint subsets  $In, Out \subset Ag$  of, respectively, input and output agents.

We assume for simplicity that  $N$  consists of three agents connected into a tree, and we show a simple analysis of the direct fusion of knowledge; clearly, more complex schemes will require a deeper and more complex analysis, but will follow the lines indicated in the example that follows.

**Fusion of Knowledge: An Example** We consider an agent  $ag \in Ag$  and – for simplicity reasons – we assume that  $ag$  has two incoming connections from agents  $ag_1, ag_2$ ; the number of outgoing connections is of no importance as  $ag$  sends the same information along each of them.

We assume that each agent  $i$  applies the rough inclusion  $\mu_L$  induced by the Łukasiewicz t-norm  $L$  (see Subsect. “[Granulation: Tools](#)” A). Each agent also applies rough inclusion on sets of the form (17) in evaluations related to extensions of formulae intensions.

Clearly, there exists a fusion operator  $o_{ag}$  that assembles from objects  $x \in U_{ag_1}, y \in U_{ag_2}$  the object  $o(x, y) \in U_{ag}$ ; we assume that  $o_{ag} = id_{ag_1} \times id_{ag_2}$ , that is,  $o_{ag}(x, y) = (x, y)$ .

Similarly, we assume that the set of attributes at  $ag$ , equals:  $A_{ag} = A_{ag_1} \times A_{ag_2}$ , that is, attributes in  $A_{ag}$  are pairs  $(a_1, a_2)$  with  $a_i \in A_{ag_i}$  ( $i = 1, 2$ ) and that the value of this attribute is defined as:

$$(a_1, a_2)(x, y) = (a_1(x), a_2(y)).$$

It follows that the condition holds:

$$\begin{aligned} o_{ag}(x, y) \text{ IND}_{ag} o_{ag}(x', y') \\ \text{iff } x \text{ IND}_{ag_1} x' \text{ and } y \text{ IND}_{ag_2} y'. \end{aligned}$$

Concerning the function  $\text{UncProp}_{ag}$ , we consider objects  $x, x', y, y'$ ; clearly,

$$\begin{aligned} \text{DIS}_{ag}(o_{ag}(x, y), o_{ag}(x', y')) \\ \subseteq \text{DIS}_{ag_1}(x, x') \times A_{ag_2} \cup \\ A_{ag_1} \times \text{DIS}_{ag_2}(y, y'), \end{aligned} \quad (23)$$

and hence,

$$\begin{aligned} |\text{DIS}_{ag}(o_{ag}(x, y), o_{ag}(x', y'))| \\ \leq |\text{DIS}_{ag_1}(x, x')| \cdot |A_{ag_2}| \\ + |A_{ag_1}| \cdot |\text{DIS}_{ag_2}(y, y')|. \end{aligned} \quad (24)$$

By (24),

$$\begin{aligned} &\mu_{ag}(o_{ag}(x, y), o_{ag}(x', y'), t) \\ &= 1 - \frac{|\text{DIS}_{ag}(o_{ag}(x, y), o_{ag}(x', y'))|}{|A_{ag_1}| \cdot |A_{ag_2}|} \\ &\geq 1 - \frac{|\text{DIS}_{ag_1}(x, x')| \cdot |A_{ag_2}| + |A_{ag_1}| \cdot |\text{DIS}_{ag_2}(y, y')|}{|A_{ag_1}| \cdot |A_{ag_2}|} \\ &= 1 - \frac{|\text{DIS}_{ag_1}(x, x')|}{|A_{ag_1}|} + 1 - \frac{|\text{DIS}_{ag_2}(y, y')|}{|A_{ag_2}|} - 1. \end{aligned} \quad (25)$$

It follows that,

$$\begin{aligned} &\text{if } \mu_{ag_1}(x, x', r), \mu_{ag_2}(y, y', s) \\ &\text{then } \mu_{ag}(o_{ag}(x, y), o_{ag}(x', y'), L(r, s)). \end{aligned} \quad (26)$$

Hence,  $\text{UncProp}(r, s) = L(r, s)$ , the value of the Łukasiewicz t-norm  $L$  on the pair  $(r, s)$ .

In consequence, the granule synthesizer  $\text{GSynt}_{ag}$  can be defined in our example as,

$$\text{GSynt}_{ag}(g_{ag_1}(x, r), g_{ag_2}(y, s)) = (g_{ag}(o_{ag}(x, y), L(r, s))). \quad (27)$$

The definition of the logic synthesizer  $\text{LSynt}_{ag}$  follows directly from our assumptions,

$$\text{LSynt}_{ag}(\phi_1, \phi_2) = \phi_1 \wedge \phi_2. \quad (28)$$

Finally, we consider extensions of the operators of our intensional logic.

We have for the extension  $I(\mu_{ag})_{\text{GSynt}_{ag}(g_1, g_2)}^{\vee}$  ( $\text{LSynt}_{ag}(\phi_1, \phi_2)$ ):

$$\begin{aligned} I(\mu_{ag})_{\text{GSynt}_{ag}(g_1, g_2)}^{\vee}(\text{LSynt}_{ag}(\phi_1, \phi_2)) \\ = I(\mu_{ag_1})_{g_1}^{\vee}(\phi_1) \cdot I(\mu_{ag_2})_{g_2}^{\vee}(\phi_2), \end{aligned} \quad (29)$$

which follows directly from (27), (28).

Let us note that

$$\begin{aligned} I(\mu_{ag_1})_{g_1}^{\vee}(\phi_1) \cdot I(\mu_{ag_2})_{g_2}^{\vee}(\phi_2) \\ = P(I(\mu_{ag_1})_{g_1}^{\vee}(\phi_1), I(\mu_{ag_2})_{g_2}^{\vee}(\phi_2)), \end{aligned}$$

where  $P$  is the Product Archimedean t-norm.

Thus, in the case of parallel fusion, where each agent works according to the Łukasiewicz t-norm, uncertainty propagation and granule synthesis are described by the Łukasiewicz t-norm  $L$  and extensions of logical intensions propagate according to the Product t-norm  $P$ .

### Reasoning in Cognitive Schemes: Rough Neural Reasoning

In neural models of computation (see Bishop [5]), an essential feature of neurons is differentiability of transfer functions; hence, we introduce a special type of rough inclusions, called *gaussian* in Polkowski [40] because of their form, by letting,

$$\mu_G(x, y, r) \text{ iff } e^{-|\sum_{a \in \text{DIS}(x, y)} w_a|^2} \geq r, \quad (30)$$

where  $w_a \in (0, +\infty)$  is a weight associated with the attribute  $a$  for each attribute  $a \in A$ ; clearly, we retain the notation of previous sections. One may notice that the gaussian rough inclusion is a modification of rough inclusion  $\mu_P$  obtained from the Product t-norm.

Let us observe in passing that  $\mu_G$  can be factored through the indiscernibility relation  $\text{IND}(A)$ , and thus its arguments can be objects as well as indiscernibility classes; we will freely use this fact.

Properties of gaussian rough inclusions are the following (cf., [41]):

- $x \text{ ing } y \text{ iff } \text{DIS}(x, y) = \emptyset$ ;
- There exists a function  $\eta(r, s)$  such that  $\mu_G(x, y, r), \mu_G(y, z, s) \text{ imply } \mu_G(x, z, \eta(r, s))$ ;
- If  $x \text{ ing } g_r^{\mu_G} y, x \text{ ing } g_s^{\mu_G} z$ , then  $g_t^{\mu_G} x \text{ ing } g_r^{\mu_G} y, g_t^{\mu_G} x \text{ ing } g_s^{\mu_G} z$  for  $t \geq \max\{r^4, s^4\}$ .

Property 1 follows by definition, and Property 2 may be verified with  $\eta(r, s) = r \cdot s \cdot e^{2 \cdot (\log r \cdot \log s)^{1/2}}$  (op.cit.). Property 3 can be verified by observing that  $t$  should satisfy conditions  $\eta(r, t) \geq r, \eta(s, t) \geq s$  because of Property 2, see Polkowski [41].

**Rough Mereological Perceptron** The rough mereological perceptron is modeled on the perceptron (see, for example, [5]). It consists of an intelligent agent  $ag$ , endowed with a gaussian rough inclusion  $\mu_{ag}$  on the information system  $I_{ag} = (U_{ag}, A_{ag})$  of the agent  $ag$ .

The input to  $ag$  is in the form of a finite tuple  $\bar{x} = (x_1, \dots, x_k)$  of objects, and the input  $\bar{x}$  is converted at  $ag$  into an object  $x = O_{ag}(\bar{x}) \in U_{ag}$  by means of an operator  $O_{ag}$ .

The rough mereological perceptron is endowed with a set of *target concepts*  $T_{ag} \subseteq U_{ag}/\text{IND}(A_{ag})$ , each target concept a class of the indiscernibility  $\text{IND}_{ag}$ .

Formally, a rough mereological perceptron is thus a tuple

$$\text{RMP} = (ag, I_{ag}, \mu_{ag}, O_{ag}, T_{ag}).$$

The output  $res_{ag}(\bar{x})$  to RMP, at the input  $\bar{x}$ , is a granule of knowledge  $gr(res)x$  with

$$r(res) = \max\{r : \text{there is } y \in T_{ag} \mu_{ag}(x, y, r)\}. \quad (31)$$

Formula (31) shows that the input  $\bar{x}$  is classified at  $ag$  as the collection of indiscernibility classes that are as close to  $x$  as the closest target class (closeness with respect to  $\mu_{ag}$ ).

### Applications in Data Mining: Granular Classifiers

Granulation of knowledge by means of  $\mu_h$  consists of forming, for each  $r \in [0, 1]$  and each  $u \in U$ , a granule  $g_h(u, r) = \{v \in U : \mu_h(v, u, r)\}$ . As, clearly,  $\mu_h$  is symmetric, from  $v \in g_h(u, r)$  it follows that  $u \in g_h(v, r)$ .

Granular data sets were proposed in Polkowski [40, 42, 43] by means of the following constructions: Given  $r \in [0, 1]$ , the set of all granules  $G_r^h = \{g_h(u, r) : u \in U\}$  is defined. From this set, a covering  $\text{Cov}_r^h(\mathcal{G})$  is chosen according to a strategy  $\mathcal{G}$ . Granules in  $\text{Cov}_r^h(\mathcal{G})$  form a new universe of objects. For each  $g \in \text{Cov}_r^h(\mathcal{G})$ , and each attribute  $a \in A$ , a factored attribute  $a_h$  is defined as  $a_h(g) = S(\{a(u) : u \in g\})$ .

The new information system  $I_r^h = (\text{Cov}_r^h(\mathcal{G}), \{a_h : a \in A\})$  is a *granular reflection* of the original information system  $I$ . The same procedure is applied to a decision system  $D = (U, A, d)$  to form the reflection  $D_r^h = (\text{Cov}_r^h(\mathcal{G}), \{a_h : a \in A\}, d_h)$ . The object  $o(g)$  defined for

**Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Table 5**

Best results for Australian credit by some rough set based algorithms

Source	Method	Accuracy	Coverage
Bazan [3]	SNAPM(0.9)	error = 0.130	—
Nguyen SH [22]	simple.templates	0.929	0.623
Nguyen SH [22]	general.templates	0.886	0.905
Nguyen SH [22]	tolerance.gen.templ.	0.875	1.0
Wroblewski [57]	adaptive.classifier	0.863	—

**Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Table 6**

Best results for Australian credit by granular approach

Source	Method	Accuracy	Coverage
[45]	granular radius 0.642857	0.867	1.0
[45]	granular radius 0.714826	0.875	1.0
[45]	granular.concept dependent radius 0.785	0.9970	0.9995

a granule  $g$  by means of  $\inf(o(g)) = \{(a_h, a_h(g)): a \in A\}$  according to a strategy  $S$  is called an *S-reflection of the granule g*; clearly,  $o(g)$  need not be a real object in the training or test sets.

The idea of data classification by means of granular reflections consists in splitting a given data set into training and test parts, forming a granular reflection of the training set, inducing classification rules from this new data set, and applying the induced rules in classifying data in the original test set.

We include some examples showing the high efficiency of this approach (see [1,43,45]).

In tests with  $\mu_h$ , the exhaustive classifier (for its public version, see [51]) was applied (cf., [1,43]). The strategy  $G$  was a random choice of an irreducible covering from the set of all granules of a given radius and the strategy  $S$  was selected as majority voting with random tie resolution.

We show results of tests with the Australian credit data set (see UCI Repository [59]), well studied in rough set literature. For comparison, we include some best results, obtained by means of rough set – based methods, in Table 5. Classification quality is expressed by means of two factors: accuracy and coverage. Accuracy (also called *total accuracy*) is the ratio of the number of correctly classified objects to the number of recognized test objects, and *total coverage*,  $\frac{\text{rec}}{\text{test}}$ , where  $\text{rec}$  is the number of recognized test cases and  $\text{test}$  is the number of test cases.

Tests on these data with the granular approach described above were carried out by splitting the Australian credit data set into training and test sets in a 1:1 ratio. The training sample was granulated and a granular reflection was formed from which, by means of the RSES exhaustive algorithm, a classifier was produced which was applied to the test part of the data to find the quality of classification.

Granules were calculated in a twofold way: first as indicated above and, second, by a modified procedure of *concept dependent granulation* (see [1]). In the latter procedure, the granule  $g_h^c(u, r) = g_h(u, r) \cap [u]_d$  was computed relative to the *concept*, that is, the decision class, to which  $u$  belonged. The results of the tests are given in Table 6 in which the best results obtained with various granulation radii are shown, taken from [45].

Results in Table 6 attest that the granular approach gives results fully comparable with other results for satisfactorily large radii of granulation, and that concept-dependent granulation gives better results than any other existing approach.

In order to test the impact of the choice of granular covering on classification, we have carried out 10 experiments with the RSES exhaustive algorithm, with random coverings on the Heart dataset (Cleveland) (see UCI Repository [59]). Results are given in Table 7. Total accuracy was found to be 0.807, and total coverage 1.0 with an exhaustive algorithm on the full data. These values are achieved here with radius of at least 0.538462, and beginning with the radius of 0.384615, the error in total accuracy is at most 0.07, and the error in total coverage is at most 0.007.

This attests to a very high stability of the granular approach showing the essential independence of results of a choice of a granular covering for inducing a granular reflection of data.

**Granules from Parameterized Variants of Rough Inclusions  $\mu_h^\varepsilon$  in Classification of Data** Given an  $\varepsilon \in [0, 1]$ , we let  $\mu_h^\varepsilon(v, u, r)$  if and only if  $|\{a \in A: \rho(a(v), a(u)) < \varepsilon\}| \geq r \cdot |A|$ . It is manifest that  $\mu^\varepsilon$  is a rough inclusion if  $\rho$  is a non-archimedean metric, that is,  $\rho(u, w) \leq$

**Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Table 7**

Effect of a choice of a granular covering on classification

Radius	Total accuracy	Total coverage
0.0	0.0	0.0
0.0769231	0.0	0.0
0.153846	0.0	0.0
0.230769	0.0–0.789	0.0–1.0
0.307692	0.01.0	0.0–1.0
0.384615	0.737–0.799	0.993–1.00
0.461538	0.778–0.822	0.996–1.0
0.538462	0.881–0.911	1.0
0.615385	0.874–0.907	1.0
0.692308	0.963–0.974	1.0
0.769231	1.0	1.0
0.846154	1.0	1.0
0.923077	1.0	1.0
1.0	1.0	1.0

$\max\{\rho(u, v), \rho(v, w)\}$ ; otherwise the monotonicity condition (RM2) of Sect. “Introduction” need not be satisfied and this takes place with most popular metrics (Euclidean, Manhattan, etc.).

In this case, a rough inclusion  $\mu^*$  is defined as follows:  $\mu_h^*(v, u, r)$  if and only if there exists an  $\varepsilon$  such that  $\mu_h^\varepsilon(v, u, r)$ . It is easy, then, to check that  $\mu^*$  is a rough inclusion. The parameter  $r$  is called the catch radius.

Granules induced by the rough inclusion  $\mu_h^*$  with  $r = 1$  have a simple structure: a granule  $g_h^\varepsilon(u, 1)$  consists of all  $v \in U$  such that  $\rho(a(u), a(v)) \leq \varepsilon$ .

The idea poses itself to use granules defined in this way to assign a decision class to an object  $u$  in the test set. First on the training set, rules are induced by an exhaustive algorithm. Then, given a set  $Rul$  of these rules, and an object  $u$  in the test set, a granule  $g_h^\varepsilon(u, 1)$  is formed in the set  $Rul$ . In this manner, the duality between objects and rules is exploited, as rules and objects can be written down in the same format, that of information sets. This also allows the use of training objects instead of rules in forming granules and permits majority voting for a decision.

Thus,  $g_h^\varepsilon(u, 1) = \{r \in Rul : \rho(a(u), a(r)) \leq \varepsilon \text{ for each attribute } a \in A(r)\}$  where  $a(r)$  is the value of the attribute  $a$  in the premise of the rule and  $A(r)$  is the set of attributes in the premise of  $r$ .

Rules in the granule  $g_h^\varepsilon(u, 1)$  take part in a voting process: for each value  $c$  of a decision class, the following factor is computed,

$$\text{param}(c) = \frac{\text{sum of supports of rules pointing to } c}{\text{cardinality of } c \text{ in the training set}}, \quad (32)$$

cf., Bazan [3], Michalski et al. [25] for a discussion of various strategies of voting for decision values.

The class  $c_u$  assigned to  $u$  is decided by

$$\text{param}(c_u) = \max_c \text{param}(c), \quad (33)$$

with random resolution of ties.

In computing granules, the parameter  $\varepsilon$  is normalized to the interval  $[0, 1]$  as follows: first, for each attribute  $a \in A$ , the value  $\text{train}(a) = \max_{\text{training set}} a - \min_{\text{training set}} a$  is computed and the real line  $(-\infty, +\infty)$  is contracted to the interval  $[\min_{\text{training set}} a, \max_{\text{training set}} a]$  by the mapping  $f_a$ ,

$$f_a(x) = \begin{cases} \min_{\text{training set}} a \text{ in case } x \leq \min_{\text{training set}} a \\ x \text{ in case } x \in [\min_{\text{training set}} a, \max_{\text{training set}} a] \\ \max_{\text{training set}} a \text{ in case } x \geq \max_{\text{training set}} a . \end{cases} \quad (34)$$

When the value  $a(u)$  for a test object  $u$  is off the range  $[\min_{\text{training set}} a, \max_{\text{training set}} a]$ , it is replaced with the value  $f_a(a(u))$  in the range. For an object  $v$ , or a rule  $r$  with the value  $a(v)$ , respectively,  $a(r)$  of  $a$  denoted  $a(v, r)$ , the parameter  $\varepsilon$  is computed as  $\frac{|a(v, r) - f_a(a(u))|}{\text{train}(a)}$ . The metric  $\rho$  was chosen as the metric  $|x - y|$  in the real line. We show results of experiments with rough inclusions discussed in this work. Our data set was a subset of Australian credit data in which the training set had 100 objects from class 1 and 150 objects from class 0 (which approximately yields the distribution of classes in the whole data set). The test set had 100 objects, 50 from each class. The RSES exhaustive classifier (see Skowron et al. [51]) applied to this data set gave accuracy of 0.79 and coverage of 1.0.

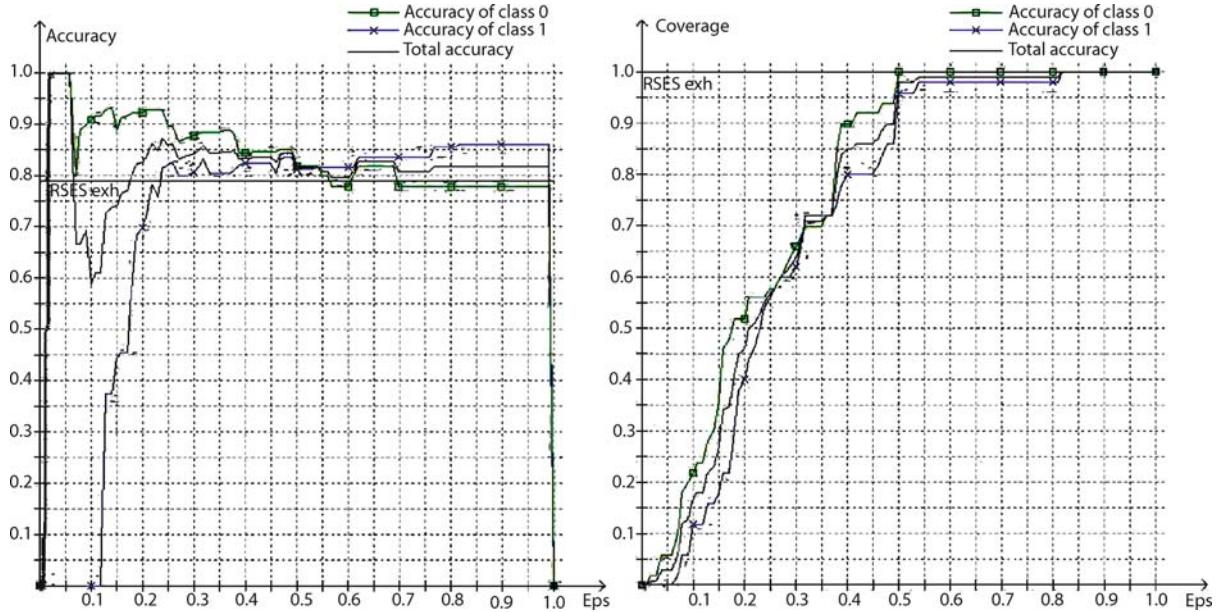
In Fig. 1 results of classification are given in function of  $\varepsilon$  for accuracy as well as for coverage.

We return to the rough inclusion  $\mu_h^*(v, u, r)$  with general radius  $r$ . The procedure applied in case of  $\mu_h^\varepsilon(v, u, 1)$  can be repeated in the general setting. The resulting classifier is a function of two parameters  $\varepsilon, r$ . In Table 8 results are included where against values of the catch radius  $r$  the best value for  $\varepsilon$ 's marked by the optimal value *optimal eps* is given for accuracy and coverage.

As shown in Sect. “Granulation of Knowledge: General Ideas and Technique of Rough Inclusions”, residual implications of continuous t-norms can supply rough inclusions according to a general formula, see (9),

$$\mu_\phi(v, u, r) \text{ iff } \phi(u) \Rightarrow_t \phi(v) \geq r, \quad (35)$$

where  $\phi$  maps the set  $U$  of objects into  $[0, 1]$  and  $\phi(u) \leq \phi(v)$  if and only if  $u$  is ing  $v$  (ing is an ingredient relation of the underlying mereology;  $\Rightarrow_t$  is the residual implication induced by the t-norm).



**Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Figure 1**  
Results for algorithm 1\_v1, best result for  $\epsilon = 0.62$ : accuracy = 0.828283, coverage = 0.99

**Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Table 8**

(40%-60%)(1-0); Australian credit; Algorithm 1\_v2. r\_catch= catch radius, optimal\_eps=Best  $\epsilon$ , acc= accuracy, cov= coverage

r_catch	optimal eps	acc	cov
nil	nil	0.79	1.0
0.071428	0	0.06	1.0
0.142857	0	0.66	1.0
0.214286	0.01	0.74	1.0
0.285714	0.02	0.83	1.0
0.357143	0.07	0.82	1.0
0.428571	0.05	0.82	1.0
0.500000	0	0.82	1.0
0.571429	0.08	0.84	1.0
0.642857	0.09	0.84	1.0
0.714286	0.16	0.85	1.0
0.785714	0.22	0.86	1.0
0.857143	0.39	0.84	1.0
0.928571	0.41	0.828283	0.99
1.000000	0.62	0.828283	0.99

A weak interesting variant of this class of rough inclusions is indicated. This variant uses sets

$$\text{dis}_\epsilon(u, v) = \frac{|\{a \in A: \rho(a(u), a(v)) \geq \epsilon\}|}{|A|},$$

and

$$\text{ind}_\epsilon(u, v) = \frac{|\{a \in A: \rho(a(u), a(v)) < \epsilon\}|}{|A|},$$

for  $u, v \in U, \epsilon \in [0, 1]$ , where  $\rho$  is a metric  $|x - y|$  on attribute value sets.

The resulting weak variant of the rough inclusion  $\mu_\phi$  is,

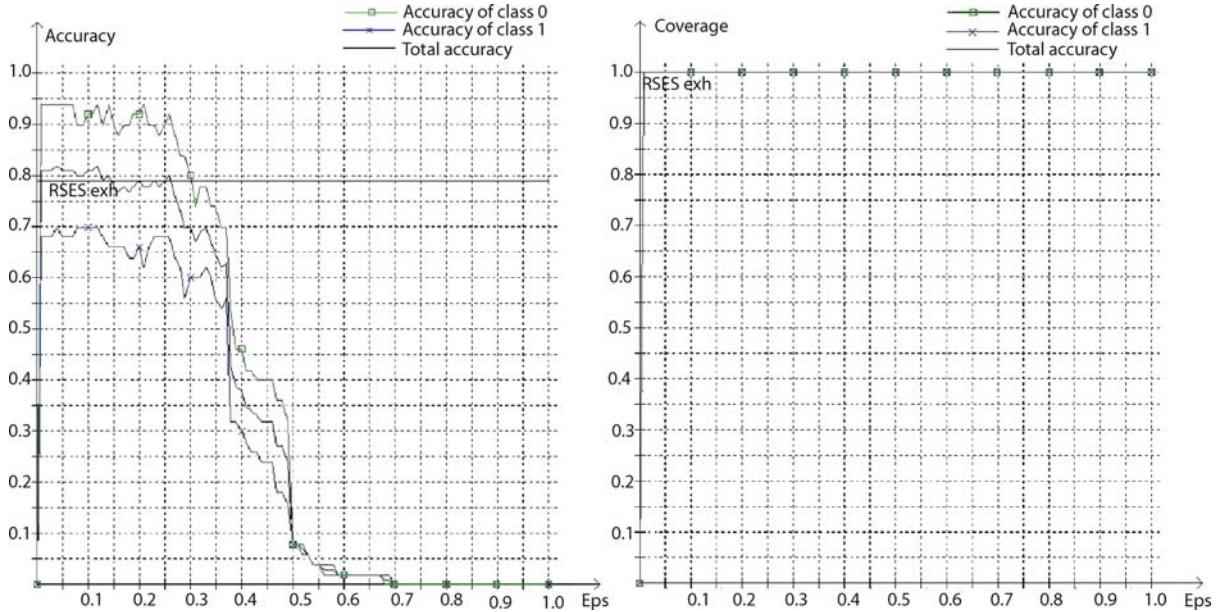
$$\mu_t(u, v, r) \text{ iff } \text{dis}_\epsilon(u, v) \rightarrow_t \text{ind}_\epsilon(u, v) \geq r. \quad (36)$$

Basic variants for three principal t-norms: the Łukasiewicz t-norm  $L = \max\{0, x + y - 1\}$ , the product t-norm  $P(x, y) = x \cdot y$ , and  $\min\{x, y\}$  are: (the value in all variants is 1 if and only if  $x \leq y$  so we give values only in the contrary case)

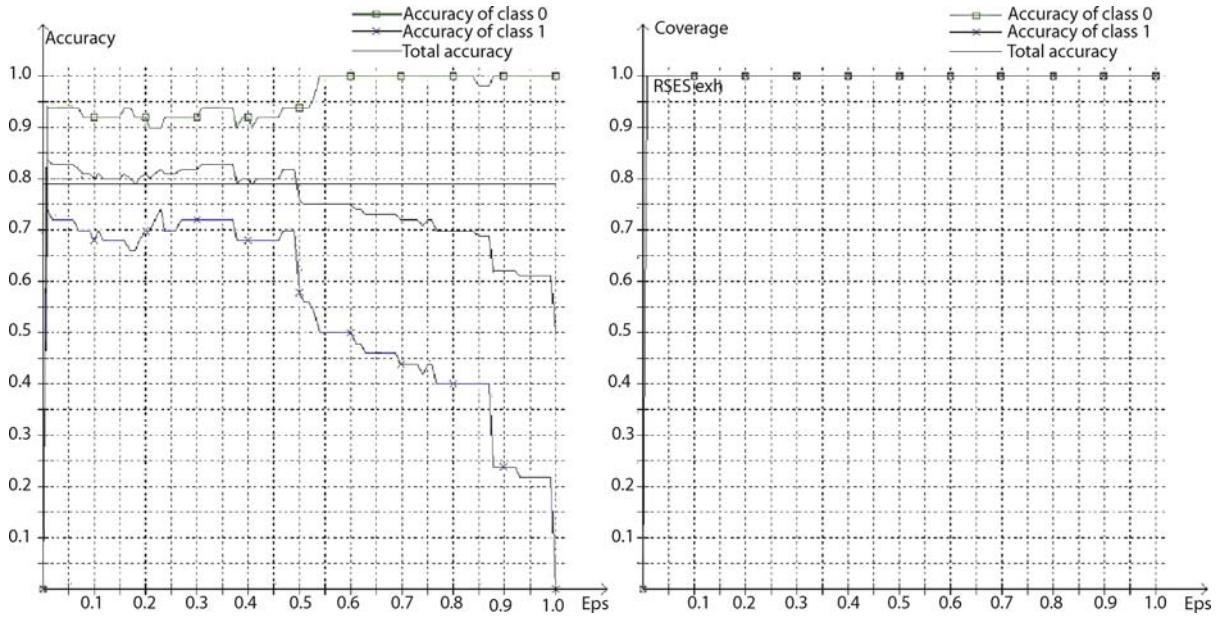
$$\mu_t(u, v, r) \text{ iff } \begin{cases} 1 - \text{dis}_\epsilon(u, v) + \text{ind}_\epsilon(u, v) \geq r & \text{for L} \\ \frac{\text{ind}_\epsilon(u, v)}{\text{dis}_\epsilon(u, v)} \geq r & \text{for P} \\ \text{ind}_\epsilon(u, v) \geq r & \text{for min.} \end{cases} \quad (37)$$

Objects in the class  $c$  in the training set vote for decision at the test object  $u$  according to the formula:

$$p(c) = \frac{\sum_{v \in c} w(v, t)}{|c| \text{ in the training set}}$$



Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Figure 2

Results for algorithm 5\_v1, best result for  $\epsilon = 0.04$ , accuracy = 0.82, coverage = 1

Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Figure 3

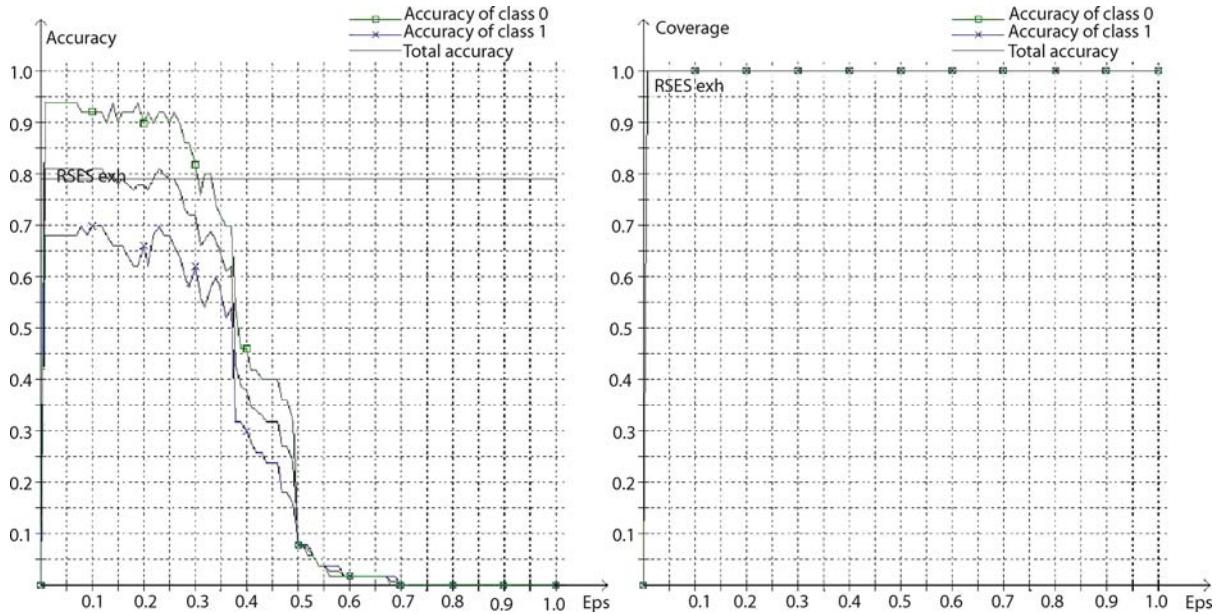
Results for algorithm 5\_v2, best result for  $\epsilon = 0.01$ , accuracy = 0.84, coverage = 1

where weight  $w(v, t)$  is  $\text{dis}_\epsilon(u, v) \rightarrow_t \text{ind}_\epsilon(u, v)$ ; rules induced from the training set pointing to the class  $c$  vote according to the formula:

$$p(c) = \frac{\sum_r w(r, t) \cdot \text{support}(r)}{|c| \text{ in the training set}}.$$

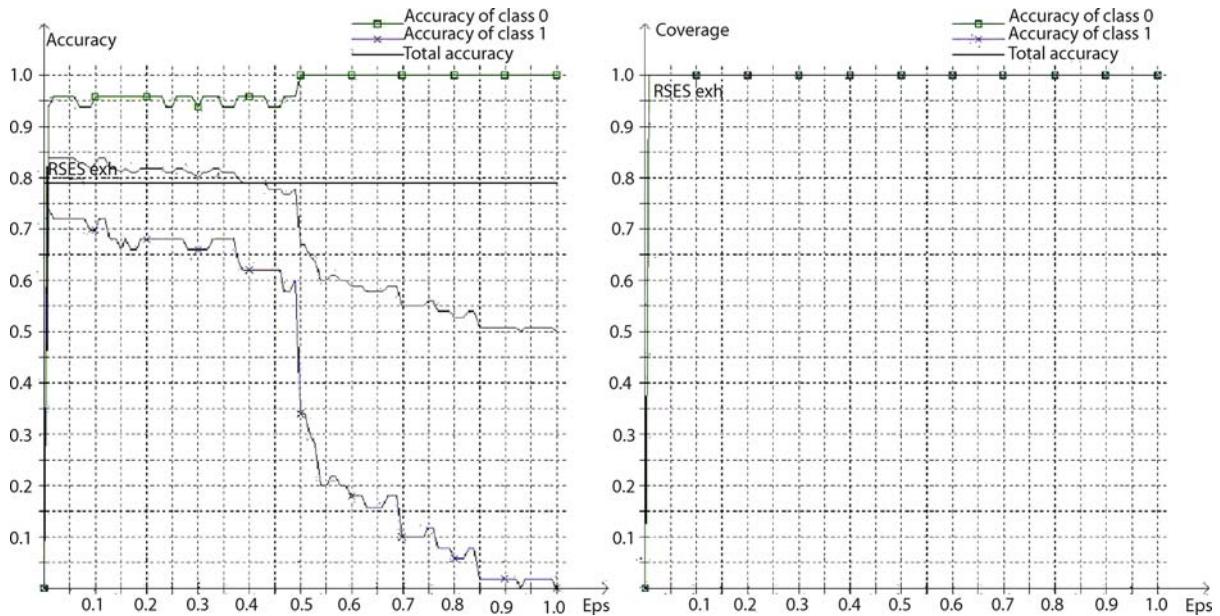
In either case, the class  $c^*$  with  $p(c^*) = \max p(c)$  is chosen. We include here results of tests with training objects and  $t = \min$  (Fig. 2) and rules and  $t = \min$  (Fig. 3).

Similarly, we include in Figs. 4, 5 results of tests with granules of training objects and rules for  $t = P$ , the product t-norm.



**Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Figure 4**

Results for algorithm 6\_v1, best result for  $\epsilon = 0.01$ , accuracy = 0.81, coverage = 1



**Granulation of Knowledge: Similarity Based Approach in Information and Decision Systems, Figure 5**

Results for algorithm 6\_v2, Best result for  $varepsilon = 0.01$ , accuracy = 0.84, coverage = 1

The results of tests in best cases for optimal values of  $\epsilon$  exceed results obtained with the standard exhaustive algorithm (the RSES exhaustive algorithm in this data set gives accuracy of 0.79 and coverage of 1.0).

### Complexity Issues

Concerning granulation in information systems, it can be proved that finding a granule about an object  $u$  is at most

$O(|U| \cdot |A|)$ . Therefore, finding all granules of the given radius is  $O(|U|^2 \cdot |A|)$ .

Selecting a coverage by granules of the given radius is therefore  $O(|U|^2 \cdot |A|)$ .

## Future Directions

As results of the last section attest, granulation is a very important ingredient in such diverse areas as reasoning about uncertainty, reasoning in distributed and cognitive schemes, and data mining. Especially in this last area, it is possible to objectively judge the contribution of granulation by comparing the quality of classifiers using granulation with those which dispense with granulation. These results show clearly that granulation produces results better, as a rule, than the standard approach of rule induction by exhaustive algorithm, and the best results are as good as any obtained by methods based on rule building by descriptors with optimization.

Further research should be aimed at better understanding the role of granulation, of the trade-off between the size of granules and quality of reasoning/classification, of the dynamics of granule formation and at finding better similarity measures adapted to particular data sets.

## Bibliography

1. Artiemjew P (2007) Classifiers from granulated data sets: Concept dependent and layered granulation. In: Proceedings RSKD'07, Workshop at ECML/PKDD'07. Warsaw Univ Press, Warsaw, pp 1–9
2. Baader F, Calvanese D, McGuinness D, Nardi D, Patel-Schneider P (eds) (2004) The description logic handbook: theory, implementation and applications. Cambridge U Press, Cambridge
3. Bazan JG (1998) A comparison of dynamic and non-dynamic rough set methods for extracting laws from decision tables. In: Polkowski L, Skowron A (eds) Rough sets in knowledge discovery 1. Physica, Heidelberg, pp 321–365
4. Bazan JG, Nguyen HS, Nguyen SH, Synak P, Wróblewski J (2000) Rough set algorithms in classification problems. In: Polkowski L, Tsumoto S, Lin TY (eds) Rough set methods and applications. New developments in knowledge discovery in information systems. Physica, Heidelberg, pp 49–88
5. Bishop CM (1995) Neural networks for pattern recognition, 2nd edn. Clarendon Press, Oxford
6. Bocheński JM (1954) Die zeitgenössischen Denkmethoden. A Francke AG, Bern
7. Duda RO, Hart PE, Stork DG (2001) Pattern classification. Wiley, New York
8. Greco S, Matarazzo B, Słowiński R (1999) On joint use of indiscernibility, similarity and dominance in rough approximation of decision classes. In: Proceedings of the 5th International Conference of the Decision Sciences Institute, Athens, Greece. pp 1380–82
9. Grzymala-Busse JW (1992) LERS – a system for learning from examples based on rough sets. In: Słowiński R (ed) Intelligent decision support: Handbook of advances and applications of the rough sets theory. Kluwer, Dordrecht, pp 3–18
10. Grzymala-Busse JW (2004) Data with missing attribute values: Generalization of indiscernibility relation and rule induction. In: Transactions on rough sets, vol I. Lecture Notes in Computer Science, vol 3100. Springer, Berlin, pp 78–95
11. Grzymala-Busse JW, Hu M (2000) A comparison of several approaches to missing attribute values in data mining. In: Rough sets and current trends in computing. Lecture Notes in CS, vol 2005. Springer, Berlin, pp 378–385
12. Klösgen W, Żytkow J (eds) (2002) Handbook of data mining and knowledge discovery. Oxford University Press, Oxford
13. Krawiec et al (1998) Learning decision rules from similarity based rough approximations. In: Polkowski L, Skowron A (eds) Rough sets in knowledge discovery 2. Physica, Heidelberg, pp 37–54
14. Kuratowski K, Mostowski A (1966) Set theory. Polish Scientific Publishers, Warsaw
15. Leśniewski S (1916) Podstawy ogólnej teorii mnogości (On the foundations of set theory), in Polish. The Polish Scientific Circle, Moscow. see also a later digest: Topoi, vol 2, 1982, pp 7–52
16. Lin TY (1988) Neighborhood systems and relational database. Abstract. In: Proceedings of CSC'88, 1988. p 725
17. Lin TY (1989) Neighborhood systems and approximation in database and knowledge based systems. In: Proceedings of the 4th International Symposium on Methodologies of Intelligent Systems, Poster Session, 1989. pp 75–86
18. Lin TY (1992) Topological and fuzzy rough sets. In: Słowiński R (ed) Intelligent decision support-handbook of applications and advances of the rough sets theory. Kluwer, Dordrecht, pp 287–304
19. Lin TY (1997) From rough sets and neighborhood systems to information granulation and computing with words. In: Proceedings of the European Congress on Intelligent Techniques and Soft Computing. pp 1602–1606
20. Lin TY (1999) Granular computing: fuzzy logic and rough sets. In: Zadeh LA, Kacprzyk J (eds) Computing with words in information/intelligent systems 1. Physica, Heidelberg, pp 183–200
21. Lin TY (2003) Granular computing. In: Lecture Notes in Computer Science, vol 2639. Springer, Berlin, pp 16–24
22. Lin TY (2005) Granular computing: Examples, intuitions, and modeling. In: Proceedings of IEEE 2005 Conference on Granular Computing GrC05, Beijing, China, 2005. IEEE Press, New York, pp 40–44
23. Lin TY (2006) A roadmap from rough set theory to granular computing. In: Proceedings RSKT 2006, 1st International Conference on Rough Sets and Knowledge Technology, Chongqing, China, 2006. Lecture Notes in Artificial Intelligence, vol 4062. Springer, Berlin, pp 33–41
24. Ling C-H (1965) Representation of associative functions. Publ Math Debrecen 12:189–212
25. Michalski RS et al (1986) The multi-purpose incremental learning system AQ15 and its testing to three medical domains. In: Proceedings of AAAI-86. Morgan Kaufmann, San Mateo CA, pp 1041–1045
26. Nieminen J (1988) Rough tolerance equality and tolerance black boxes. Fundamenta Informaticae 11:289–296
27. Nguyen SH (2000) Regularity analysis and its applications in Data Mining. In: Polkowski L, Tsumoto S, Lin TY (eds) Rough set methods and applications. New developments in knowl-

- edge discovery in information systems. Physica, Heidelberg, pp 289–378
28. Novotny M, Pawlak Z (1988) Partial dependency of attributes. *Bulletin Polish Academy. Ser Sci Math* 36:453–458
  29. Novotny M, Pawlak Z (1992) On a problem concerning dependence spaces. *Fundamenta Informaticae* 16:275–287
  30. Pawlak Z (1982) Rough sets. *Intern J Comput Inf Sci* 11:341–356
  31. Pawlak Z (1985) On rough dependency of attributes in information systems. *Bulletin Polish Academy. Ser Sci Tech* 33: 551–559
  32. Pawlak Z (1991) Rough sets: theoretical aspects of reasoning about data. Kluwer, Dordrecht (Holland)
  33. Pawlak Z, Skowron A (1993) A rough set approach for decision rules generation. In: Proceedings of IJCAI'93 Workshop W12. The Management of Uncertainty in AI, France, 1993. also ICS Research Report 23/93, Warsaw University of Technology, Institute of Computer Science
  34. Pawlak Z, Skowron A (1994) Rough membership functions. In: Yaeger RR, Fedrizzi M, Kasprzyk J (eds) Advances in the Dempster–Schafer Theory of evidence. Wiley, New York, pp 251–271
  35. Poincaré H (1905) Science and hypothesis. Walter Scott Publishing, London
  36. Polkowski L (2002) Rough sets. mathematical foundations. Physica, Heidelberg
  37. Polkowski L (2003) A rough set paradigm for unifying rough set theory and fuzzy set theory. In: Proceedings RSFDGrC03, Chongqing, China, 2003. Lecture Notes in AI, vol 2639. Springer, Berlin, pp 70–78. also in: *Fundamenta Informaticae* 54:67–88
  38. Polkowski L (2004) A note on 3-valued rough logic accepting decision rules. *Fundamenta Informaticae* 61:37–45
  39. Polkowski L (2004) Toward rough set foundations. Mereological approach. In: Proceedings RSCTC04, Uppsala, Sweden. Lecture Notes in AI, vol 3066. Springer, Berlin, pp 8–25
  40. Polkowski L (2005) Formal granular calculi based on rough inclusions. In: Proceedings of IEEE 2005 Conference on Granular Computing GrC05, Beijing, China. IEEE Press, New York, pp 57–62
  41. Polkowski L (2005) Rough-fuzzy-neurocomputing based on rough mereological calculus of granules. *Intern J Hybrid Intell Syst* 2:91–108
  42. Polkowski L (2006) A model of granular computing with applications. In: Proceedings of IEEE 2006 Conference on Granular Computing GrC06, Atlanta, USA. IEEE Press, New York, pp 9–16
  43. Polkowski L (2007) Granulation of knowledge in decision systems: The approach based on rough inclusions. The method and its applications. In: Proceedings RSEISP'07 (in memory of Z. Pawlak). Lecture Notes in AI, vol 4585. pp 69–79
  44. Polkowski L (2007) The paradigm of granular rough computing. In: Proceedings ICCI'07. 6th IEEE Intern. Conf. on Cognitive Informatics. IEEE Computer Society, Los Alamitos, pp 145–163
  45. Polkowski L, Artiemjew P (2007) On granular rough computing: Factoring classifiers through granular structures. In: Proceedings RSEISP'07, Warsaw. Lecture Notes in AI, vol 4585. Springer, Berlin, pp 280–289
  46. Polkowski L, Semeniuk-Polkowska M (2005) On rough set logics based on similarity relations. *Fundamenta Informaticae* 64:379–390
  47. Polkowski L, Skowron A (1994) Rough mereology. In: Proceedings of ISMIS'94. Lecture notes in AI, vol 869. Springer, Berlin, pp 85–94
  48. Polkowski L, Skowron A (1997) Rough mereology: a new paradigm for approximate reasoning. *Intern J Approx Reason* 15(4):333–365
  49. Polkowski L, Skowron A (1999) Towards an adaptive calculus of granules. In: Zadeh LA, Kacprzyk J (eds) Computing with words in information/intelligent systems,1. Physica, Heidelberg, pp 201–228
  50. Polkowski L, Skowron A, Żytkow J (1994) Tolerance based rough sets. In: Lin TY, Wildberger M (eds) Soft computing: rough sets, fuzzy logic, neural networks, uncertainty management, knowledge discovery. Simulation Councils Inc, San Diego CA, pp 55–58
  51. RSES: Skowron A et al (1994) A system for data analysis. <http://logic.mimuw.edu.pl/~rses/>. Accessed 19 Sept 2008
  52. Russell SJ, Norvig P (2003) Artificial intelligence. A modern approach, 2nd edn. Prentice Hall (Pearson Education Inc), Upper Saddle River NJ
  53. Skowron A (1993) Boolean reasoning for decision rules generation. In: Komorowski J, Ras Z (eds) Proceedings of ISMIS'93. Lecture Notes in AI, vol 689. Springer, Berlin, pp 295–305
  54. Skowron A, Rauszer C (1992) The discernibility matrices and functions in decision systems. In: Ślowiński R (ed) Intelligent decision support. handbook of applications and advances of the rough sets theory. Kluwer, Dordrecht, pp 311–362
  55. Skowron A, Stepaniuk J (2001) Information granules: towards foundations of granular computing. *Intern J Intell Syst* 16: 57–85
  56. Skowron A, Swiniarski RW (2004) Information granulation and pattern recognition. In: Pal SK, Polkowski L, Skowron A (eds) Rough – neural computing. Techniques for computing with words. Springer, Berlin, pp 599–636
  57. Stefanowski J (1998) On rough set based approaches to induction of decision rules. In: Polkowski L, Skowron A (eds) Rough sets in knowledge discovery 1. Physica, Heidelberg, pp 500–529
  58. Stefanowski J (2007) On combined classifiers, rule induction and rough sets. In: Transactions on rough sets, vol VI. Lecture Notes in Computer Science, vol 4374. Springer, Berlin, pp 329–350
  59. UCI Repository: available at <http://archive.ics.uci.edu/ml/>. Accessed 19 Sept 2008
  60. van Benthem J (1988) A manual of intensional logic. CSLI Stanford University, Stanford CA (US)
  61. Wille R (1982) Restructuring lattice theory: An approach based on hierarchies of concepts. In: Rival I (ed) Ordered sets. Reidel, Dordrecht, pp 445–470
  62. Wojna A (2005) Analogy-based reasoning in classifier construction. In: Transactions on Rough Sets vol IV. Lecture Notes in Computer Science, vol 3700. Springer, Berlin, pp 277–374
  63. Wróblewski J (2004) Adaptive aspects of combining approximation spaces. In: Pal SK, Polkowski L, Skowron A (eds) Rough – neural computing. Techniques for computing with words. Springer, Berlin, pp 139–156
  64. Yao YY (2000) Granular computing: Basic issues and possible solutions. In: Proceedings of the 5th Joint Conference on Information Sciences I. Assoc Intell Machinery, Atlantic NJ, pp 186–189
  65. Yao YY (2005) Perspectives of granular computing. In: Proceedings of IEEE 2005 Conference on Granular Computing GrC05, Beijing, China. IEEE Press, New York, pp 85–90
  66. Zadeh LA (1979) Fuzzy sets and information granularity. In:

- Gupta M, Ragade R, Yaeger RR (eds) Advances in fuzzy set theory and applications. North-Holland, Amsterdam, pp 3–18
67. Zeeman EC (1965) The topology of the brain and the visual perception. In: Fort MK (ed) Topology of 3-manifolds and selected topics. Prentice Hall, Englewood Cliffs NJ, pp 240–256

## Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures

P. MARTIN MAI

Swiss Seismological Service, Institute of Geophysics,  
ETH, Zürich, Switzerland

### Article Outline

- Glossary
- Definition of the Subject
- Introduction
- Characterizing Earthquake Source Complexity
- Wave Propagation in Complex Media:

  - Path and Site Effects
  - Ground-Motion Scaling Relations
  - Future Directions
  - Acknowledgments
  - Bibliography

### Glossary

**Attenuation relation (ground-motion prediction equation)** The term “attenuation relation” is a former shorthand notation in earthquake engineering for “empirical ground-motion attenuation relationship”, now referred to as “ground-motion prediction equation” (GMPE). Attenuation relations represent empirical scaling equations that relate observed **ground-motion intensity measures** to parameters of the earthquake source, the wave propagation from the source to the observer and the site response at the observer location.

**Dynamic rupture model** Dynamic rupture models build a physical understanding of the earthquake rupture based on the material properties around the source volume, and the initial and boundary conditions for the forces/stresses acting on the fault plane. The distribution of on-fault **slip-rate** vectors and the temporal rupture evolution is obtained by solving the elastodynamic equations of motion under an assumed constitutive law (friction model), considering the energy balance at the crack tip (Chap. 11 in [6]). See also **kinematic rupture model**.

**Ground motion intensity measures** Earthquake shaking due to seismic waves, observed at recording sites or experienced by people and structures, is commonly reported in terms of various scalar intensity measures that capture parts of the transient wave-field. Seismogram-based ground-motion intensity measures are, for instance, peak ground acceleration (*PGA*) and peak ground velocity (*PGV*), while the modified Mercalli intensity (*MMI*) is a damage-related measure. In earthquake engineering, ground-motion intensities are often reported as *response spectra*: the response of an idealized building (modeled as a single-degree-of-freedom oscillator) of given eigenperiod *T* and damping  $\zeta$  (usually 5%) to a given ground-motion time series. Spectral acceleration ( $S_A$ ), spectral velocity ( $S_V$ ) and spectral displacement ( $S_D$ ) are analyzed considering the period of the structure.

**Ground motion uncertainty** In ground-motion prediction for engineering purposes, random (aleatory) variability and scientific (epistemic) uncertainty are distinguished. The latter is due to incomplete knowledge and/or limited data, and is captured by alternative empirical attenuation relations or different ground-motion simulation strategies. Aleatory variability is quantified in terms of a standard deviation of an attenuation relation or by a large number of model realizations within a particular simulation method. The distinction between aleatory variability and epistemic uncertainty is particularly useful in probabilistic seismic hazard analysis (PSHA).

**Kinematic rupture model** A kinematic rupture model characterizes the time-dependent displacement field on the rupture plane without considering the forces or stresses acting on the fault and causing its motion. The rupture process is completely specified by the spatio-temporal distribution of the slip vector, the local **slip-velocity** function on the fault, and the **rupture velocity** with which the rupture propagates over the fault plane. See also **dynamic rupture model**.

**Path effects** Seismic waves propagating through the Earth are sensitive to the detailed geologic structure along the wave path, generating pronounced waveform complexities. Considering crust and upper-mantle structure (relevant for near-field ground motions) three major elements to path effects are distinguished in practice: (a) waves in a flat-layered attenuating Earth; (b) basins and other deterministic deviations from a flat-layered model; (c) random heterogeneities in the three-dimensional velocity-density structure. The distinction between path effects and **site effects** is often ambiguous.

**Rise time** The rise time (or slip duration)  $\tau_r$  measures how long each point on the fault moves during the rupture process, and must not be confused with the *rupture duration*. The rise time is related to *the slip-velocity function*, and is usually measured as the time it takes to attain 5–95% of the final slip at each point. For simple parametric slip-functions (e.g. boxcar, isosceles triangle or combinations thereof), the rise time is generally given by the width of this function.

**Rupture duration** The rupture duration characterizes the total time for the earthquake rupture process to complete, starting at the nucleation point (hypocenter) and lasting until the last point on the rupture plane stops sliding. Rupture duration therefore depends on *rupture velocity* and scales with the size (source dimension) of the earthquake.

**Rupture velocity** Earthquake ruptures, either modeled as propagating cracks or slip-pulses, expand over the fault plane at rupture speeds ( $v_r$ ) close to the local shear-wave velocity ( $v_s$ ), typically in the range  $0.5 \cdot v_s \leq v_r \leq 0.9 \cdot v_s$ , or about 1.0–3.5 km/sec for crustal earthquakes. However, the crack tip, the transition region from unbroken, intact rock to the currently slipping zone, does not necessarily travel at constant rupture speed. Rupture velocity may locally slow down or accelerate, even to super-shear velocities (in which case the crack front travels at speeds faster than the local shear-wave velocity), depending on the initial and boundary conditions that govern the dynamic rupture process.

**Site effects** Site effects refer to wave-propagation effects in the immediate proximity to the observation point; they are distinguished from *path effects* which comprise the complete path from the source to the receiver (although the boundary between these two is often ill defined). The local sedimentary cover, topography, strong geologic contrasts or water-table variations may contribute to site effects that modify the incoming “bedrock” seismic motions.

**Slip distribution** The slip distribution represents the cumulative slip on each point on the fault acquired during the co-seismic rupture process (i.e. small contributions from post-seismic slip episodes are ignored). A slip distribution for an earthquake is computed from the space-time integration of *slip-velocity functions* on the rupture plane.

**Slip-velocity function (Slip-rate function)** Each point participating in the rupture process experiences a time-dependent slip history during which the two sides of the fault go through a stage of acceleration, stable sliding, deceleration, and final stopping. This local

displacement trajectory is often represented in terms of a slip-velocity function (or slip function) whose details depend on the dynamic rupture process and the constitutive behavior of the host rock. Slip-rate functions are often modeled using simple parametric functions.

**Source effects** Amplitudes and waveform character of seismic waves are strongly affected by source effects, i.e. by the details of the earthquake rupture process. Far-field signals carry the signature of the overall “point-source” earthquake source mechanism; near-field recordings are very sensitive to the spatio-temporal details of the rupture process, characterized in a finite-fault source model either as *kinematic* or *dynamic rupture model*.

**Static stress drop** The static stress drop  $\Delta\sigma$  represents the difference between the initial and final stress across the fault before and after an earthquake, and is related to slip on the fault. It is defined, based on a shear crack with uniform stress drop, as  $\Delta\sigma = C \cdot \mu \cdot D/L_c$ , where  $\mu$  is the shear-modulus,  $D$  the mean slip over the fault,  $L_c$  a characteristic length scale (usually the smallest dimension of the rupturing fault), and  $C$  a constant of order unity which depends on the source geometry. Using the fault width  $W$  as characteristic length, static stress drop is related to the seismic moment,  $M_o = \mu \cdot L \cdot W \cdot D = C \cdot \Delta\sigma \cdot A^{3/2}$ , where  $A$  is fault area, and  $L$  is fault length. Inferred values of static stress drop are in the range of 0.1–10 MPa, independent of seismic moment, leading to the generally assumed self-similar constant stress-drop scaling (see ▶ [Earthquake Scaling Laws](#)). The static stress drop must not be confused with the dynamic stress drop [164] which captures the time-dependent stress change on a point of the fault during the dynamic faulting event, and may be significantly higher or lower than the static stress drop.

## Definition of the Subject

The accurate prediction of the level and variability of (potentially damaging) near-source strong-ground motions in future earthquakes is one of the key challenges for seismologists and earthquake engineers. The increasing number of near-source recordings collected by dense strong-motion networks exemplifies the inherent complexity of near-field ground shaking, governed by a number of (partially interacting) physical processes. Characterizing, quantifying, and modeling (either by means of empirical scaling relations or by numerical simulations) ground-motion complexity requires the joint investigation

of three dominant ingredients: (I) the physics of earthquake rupture; (II) the details of wave-propagation in heterogeneous media; (III) the effects of local site conditions.

This article discusses briefly the beginnings of strong-motion seismology and the recognition of ground-motion complexity. Using two well recorded recent earthquakes, I introduce observational aspects of near-field ground shaking and the basic mathematical description for computing ground motion. The article proceeds by describing each of the three “ground-motion ingredients” in some detail, but does not attempt to provide an in-depth review of all the scientific advancements in these fields. Rather, I explain the key elements for characterizing and modeling ground-motion complexity, supplemented with a concise overview of the underlying physical processes. Current research increasingly incorporates advanced physical concepts into standard practice, therefore leading to improved strong-motion simulation approaches to accurately predict intensity and variability of near-source shaking.

## Introduction

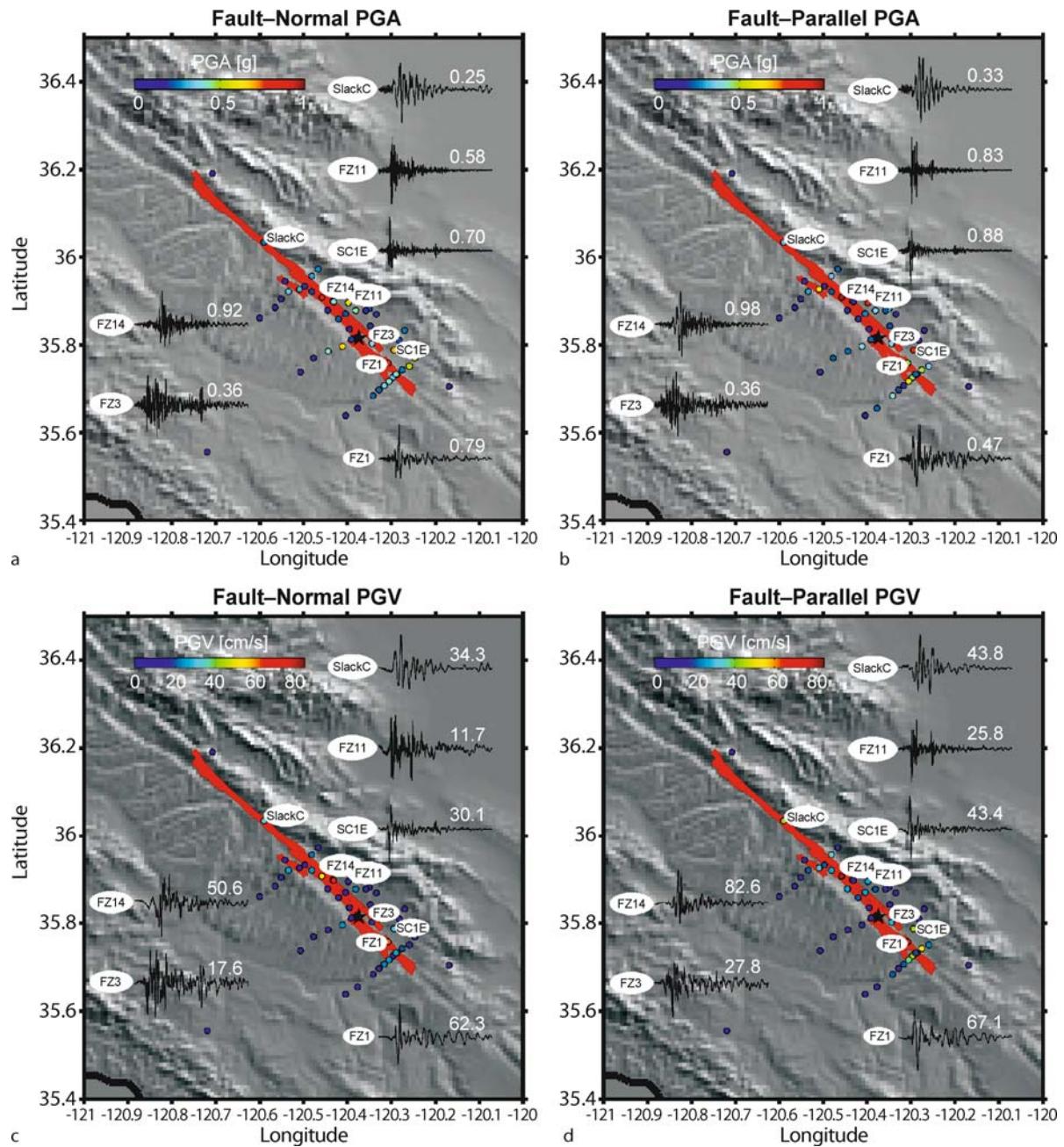
As early as 1910 Reid [148] identified strong spatial variations of seismic shaking due to the 1906 San Francisco earthquake, which he correctly attributed to the specific geologic conditions at the locations from which shaking intensities were reported. The space-time complexity of the seismic wavefield in the vicinity of the causative fault became more widely recognized since the beginnings of strong-motion seismology in the 1930s and the establishment of strong-motion networks in the 1960s [9]. Observed peak accelerations frequently reach 1 g, occasionally even exceed 2 g, where nearby stations often show not only much lower peak amplitudes but also very different waveform character. Since the advent of modern digital instruments in the 1990s and corresponding online near-real-time databases, seismologists and earthquake engineers have access to high-quality recordings which irrefutably show the complexity of earthquake shaking. Given a particular location, near-source ground motions vary strongly for different earthquakes, and so do the motions for a single event when recorded at different sites.

While hundreds of earthquakes happen annually in the magnitude range  $6.0 \leq M \leq 6.9$ , only about 15 events occur in the range  $7.0 \leq M \leq 7.9$ , most of which happening in remote areas and do not cause large damage. Those events are typically recorded at rather large distances (so called teleseismic events) and have weak (low amplitude) motions. Their ground-motion variability is due to differences in the large-scale wave paths through the Earth and due to the properties of the point-source representa-

tion for far-distant earthquakes. Weak-motion recordings are also obtained from small nearby earthquakes. In this article, however, we will be concerned with strong-motion observation in the near-field of (large) earthquakes. Such observations are much harder to obtain than teleseismic recordings because we cannot anticipate in detail where future earthquakes may happen (in order to optimize the instrumentation) and also because large crustal earthquakes are rather infrequent.

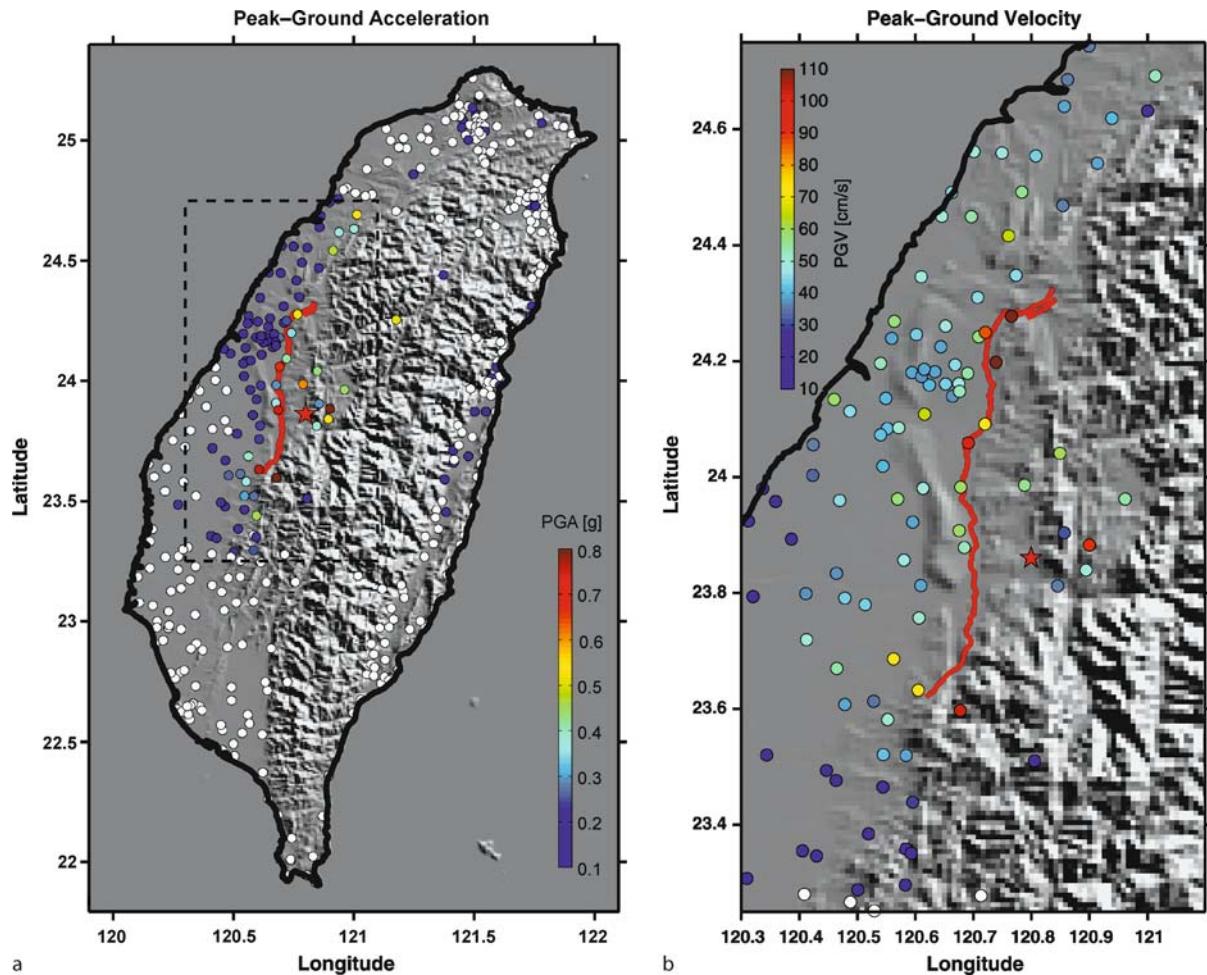
Investigating the characteristics of near-source ground-motions is of great importance for earthquake engineers who are concerned with seismically safe design, and for seismologists who study the physical processes leading to ground-motion complexity. To illustrate the large ground-motion variability in a single earthquake, I plot horizontal peak ground accelerations and peak ground velocities recorded at 47 sites in the Sept 28, 2004,  $M$  6.0 Parkfield earthquake (Fig. 1), contrasted with ground-motion intensities ( $PGA$ ,  $PGV$ ,  $PGD$ ) for the Sept 20, 1999,  $M$  7.6 Taiwan (Chi-Chi) earthquake, shown at 441 locations (Figs. 2–4) (waveforms from the COSMOS database, <http://db.cosmos-eq.org>). First-order observations from these data are: (i) the maximum  $PGA$  is larger for the smaller earthquake while the highest  $PGV$  values are roughly identical; (ii) ground-motions tend to be larger close to the fault trace and decay in amplitude with increasing distance from the fault. However, closer inspection reveals strong variability in the near-field motions for each earthquake. The Parkfield data exhibit ground-motion differences between the two horizontal components of motion (fault-parallel and fault-normal, both for  $PGA$  and  $PGV$ ) and large variability between neighboring sites. This site dependence is even more pronounced for the Chi-Chi event.  $PGA$ -values are highest for sites very close to the fault, on the hanging-wall (east of the fault trace) and at a few locations in the northward and southward extension of the fault-trace. This pattern changes, however, for the recorded  $PGV$  and  $PGD$ -values which are largest on the foot-wall (west side of the fault trace) and towards the northern end of the rupture.

In addition to the scalar measures of shaking amplitude (intensity), Figs. 1–4 display recorded waveforms for selected sites which further illustrate the complexity and variability of near-field motions. Acceleration time series accentuate the high-frequency content of ground-motion and display an almost “shaped random noise” character, with amplitudes that increase rapidly in the beginning and then taper off gradually in the seismic coda (the incoherent wavefield after the arrival of prominent seismic phases). In contrast, velocity waveforms usually show very distinct arrivals and energetic pulses (directivity



Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 1

Ground-motion intensities (colored circles) for the Sept 28, 2004 Parkfield (M 6.0) earthquake, recorded at near-field stations, and selected waveforms (data from COSMOS database). The black star denotes the epicenter, ( $\text{Lon} = -120.374$ ,  $\text{Lat} = 35.815$ ) red-lines show the mapped fault trace of the San Andreas Fault in the Parkfield area, the gray background displays a shaded relief map. **a** Peak ground accelerations (PGA) of fault-normal component; **b** PGA of fault-parallel component; **c** Peak ground velocity (PGV) of fault-normal component; **d** PGV of fault-parallel component. Seismic traces are shown for 30s, the small number indicates the corresponding PGA or PGV value of that record



#### Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 2

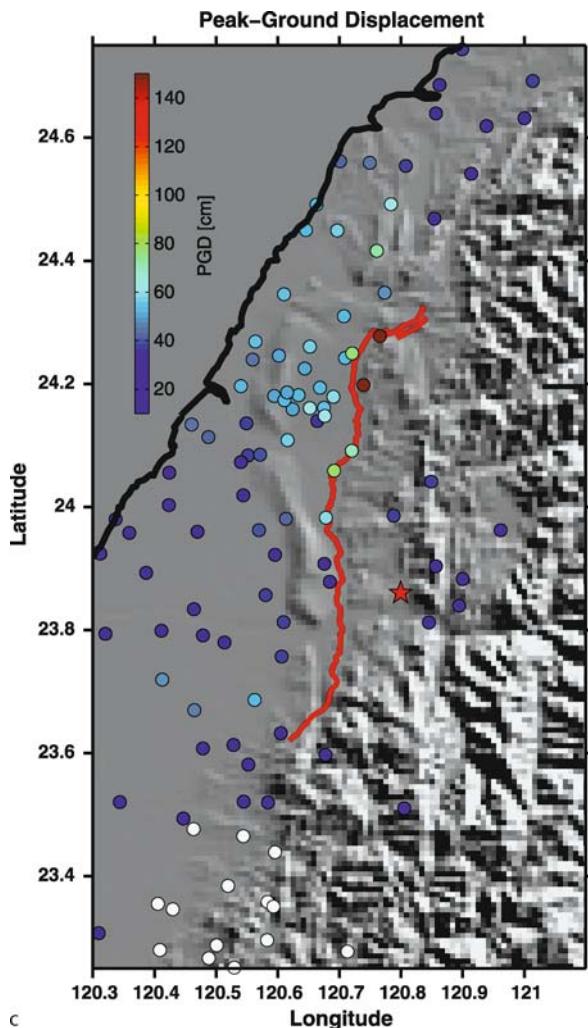
Ground-motion intensities (colored circles) for the Sept 20, 1999 Taiwan (M 7.6) earthquake, recorded at 441 locations (data from COSMOS database). The red star denotes the epicenter (Lon = 120.7995, Lat = 23.860), red-lines show the mapped fault trace of the Chelungpu fault, heavy black lines trace the coast line, the gray background displays a shaded relief map. **a** Large-scale view of PGA, white dots are sites with PGA < 0.1 g. **b** Zoomed view for PGV (white dots mark PGV < 10 cm/s). **c** Zoomed view for PGD (white dots mark PGD < 10 cm)

pulses) whose frequency-dependent amplitude and waveform character change with source-site geometry and rupture propagation direction. Their waveform character also varies strongly between closely spaced stations for a given event, illustrating the importance of localized source properties and site effects. Note also the large differences in ground-motion amplitudes and waveform shapes between the two events for stations at similar distance to the fault, indicating the dependence on earthquake source properties of near-field ground motions. Section “Future Directions” compares the near-field recordings for these two earthquakes against empirical predictions for several published attenuation relations to further illustrate ground-motion complexity.

The variability in shaking intensity and the complexity in near-fault seismograms results from three physical processes: (I) the complex dynamics of earthquake rupture and the associated radiation of seismic waves; (II) the propagation of these seismic waves through the heterogeneous Earth; (III) the interaction of the seismic wavefield with local geology/morphology, referred to as the site conditions at each observation points. Mathematically, the time-dependent ground displacement  $u_k(t)$  at a particular location  $k$  is described as

$$u_k(t) = s(t) * g_k(t) * l_k(t) \quad (1)$$

where  $*$  denotes the convolution operator,  $s(t)$  represents the **source effects** due to the earthquake rupture pro-



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 2**  
(continued)

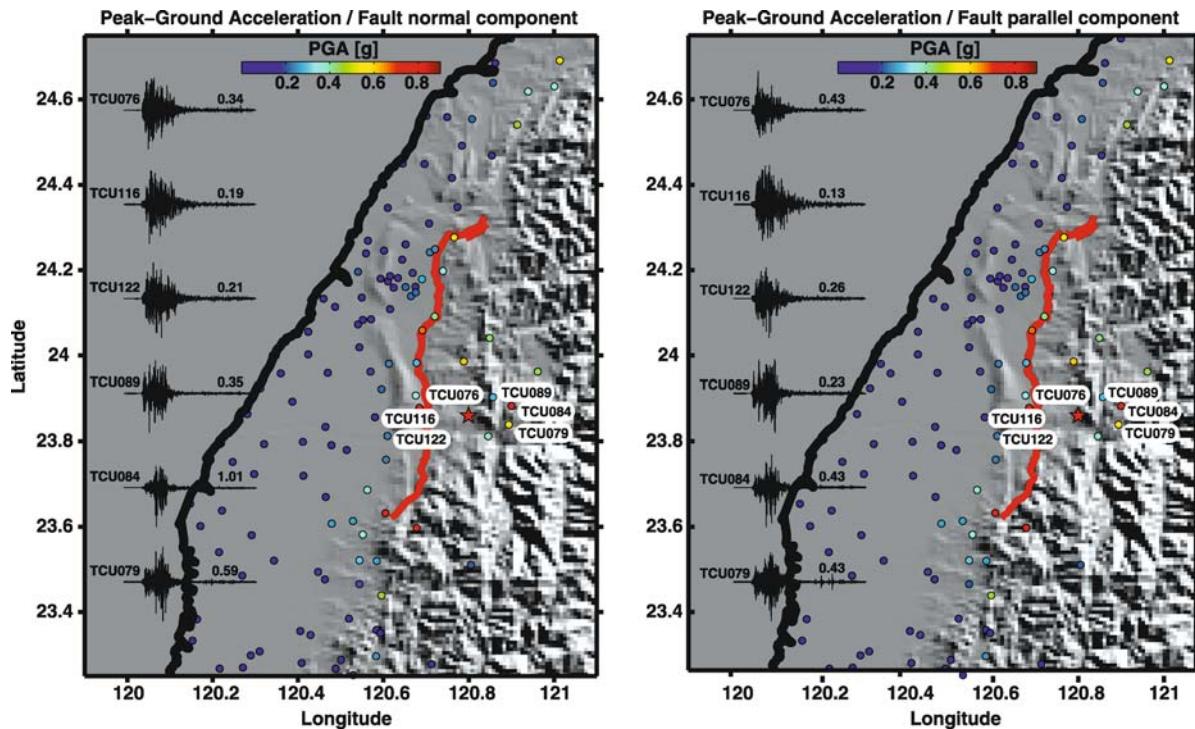
cess,  $g_k(t)$  describes the **path effects** due to wave propagation from the source to site  $k$ , and comprises  $l_k(t)$  the **local site effects** due to the small-scale geological conditions at the  $k$ th observation point. Equation (1) quantifies ground-motion generation, omitting for simplicity an additional instrument response  $i_k(t)$  that modulates the seismic recording. The earthquake-source contribution  $s(t)$  can be further subdivided into effects originating from the local time-dependent particle motion on the fault,  $p_{ij}(t)$ , and from rupture finiteness,  $f(t)$ . The term  $g_k(t)$  represents the Earth transfer function and contains contributions from layered Earth structure and seismic-wave attenuation, but may also comprise effects due to random heterogeneities in the Earth and/or basin and topographic structures. The factor  $l_k(t)$  describes the effects

of soil structure and may also account for non-linear soil behavior.

In the following, I will use Eq. (1) as a “roadmap” for this article to illustrate the various factors of ground-motion complexity. In Sect. “[Characterizing Earthquake Source Complexity](#)”, I present methods and relevant parameters to quantify the earthquake rupture process. Wave propagation in complex media, both deterministic and stochastic, are described in Sect. “[Wave Propagation in Complex Media: Path and Site Effects](#)”. A large body of literature exists for each of these topics, both from an observational/experimental view and from theoretical work; it is beyond the scope of this article to provide an in-depth review of all relevant material. Instead, I will focus on some of the key aspects of earthquake ruptures and waves in inhomogeneous media that are most relevant for understanding the complexity of near-field ground motions. Note also that a number of specialized articles in this encyclopedia provide more detailed information on earthquake source physics (see ▶ [Earthquake Nucleation Process](#), ▶ [Earthquake Scaling Laws](#)) and wave-propagation phenomena in complex media (see ▶ [Seismic Wave Propagation in Media with Complex Geometries, Simulation of](#), ▶ [Seismic Waves in Heterogeneous Earth, Scattering of](#)). Section “[Wave Propagation in Complex Media: Path and Site Effects](#)” also includes various aspects of local site conditions that lead to pronounced site effects, in particular non-linear site phenomena. Section “[Ground-Motion Scaling Relations](#)” focuses on empirical ground-motion prediction equations (GMPE’s). Since many studies have been published on non-linear site effects and ground-motion prediction equations, both from the seismology and the earthquake-engineering communities, an extensive review of the developments in these fields is not attempted. Instead, I concentrate on non-linear soil behavior directly beneath a site of interest that substantially affects the ground-shaking levels, and summarize some of the latest findings related to ground-motion attenuation relations. The article concludes with an outlook onto future tasks and challenges for characterizing, quantifying and predicting ground-motion complexity.

### [Characterizing Earthquake Source Complexity](#)

Geologic faults are generally geometrically complex multi-scale structures, characterized by one or more main fault strands, with associated subsidiary branches (fault segments) that form three-dimensional (3D) fault networks. However, in many cases and for almost all practical purposes, faults are approximated as planar surfaces. Using this plane-fault approximation, Fig. 5 illustrates the three



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 3**

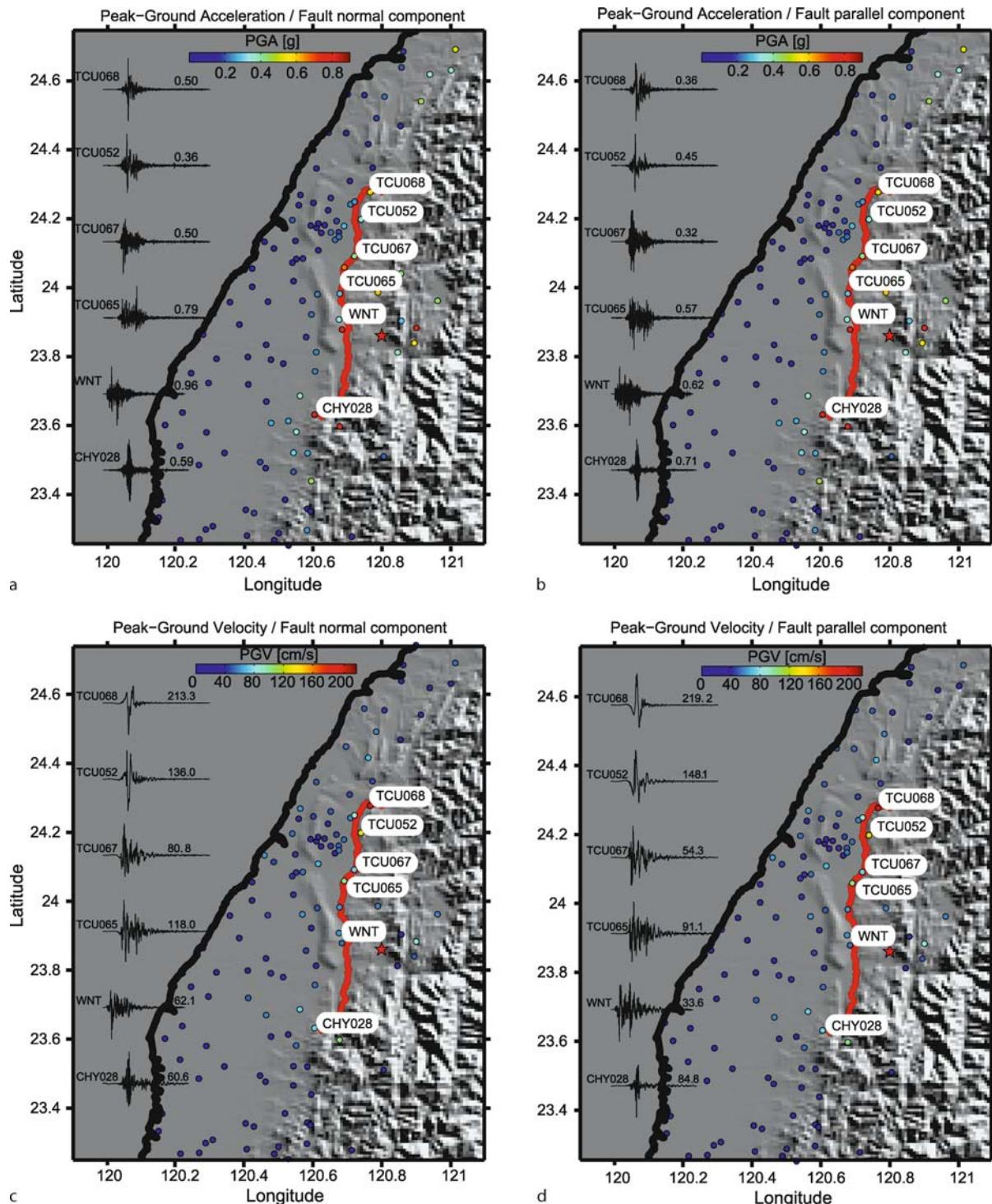
Near-field ground accelerations for the fault-normal (left) and fault-parallel (right) motion during the 1999 Chi-Chi earthquake; waveforms are plotted for selected stations in the vicinity of the epicenter (red star). Sites located on the footwall, i.e. west of the fault-trace, generally show lower ground-motions than sites located on the hanging-wall (east of the fault trace) despite similar epicentral distance

main factors affecting near-source ground-motion complexity – source, path and site effects. This Section focuses on the earthquake rupture process, i. e. the source, and its properties important for ground-motion generation. A fault (shown as a planar surface with color-coded slip distribution) is embedded in a rock volume, and has a specific orientation in space described by the strike-angle  $\Phi$  (representing the azimuth of the fault's projection onto the surface, measured clockwise from North) and the dip-angle  $\delta$  (measured downward from the surface to the fault in the vertical plane perpendicular to the strike). The strike direction is defined such that, using the right-hand rule, the dip-angle is smaller than  $90^\circ$  (for a vertically dipping fault,  $\delta = 90^\circ$ , the strike direction is arbitrarily either direction).

Given the overall source geometry, the slip-vector on the fault plane defines the relative motion between the two blocks. The angle of slip, or rake angle  $\lambda$ , measured in the fault plane from the strike direction, shows the movement of the hanging wall relative to the foot wall (see inset in Fig. 5). The following definitions apply:  $\lambda = 0^\circ$  – *left-lateral strike-slip*, i. e. the hanging wall (or near-side of a ver-

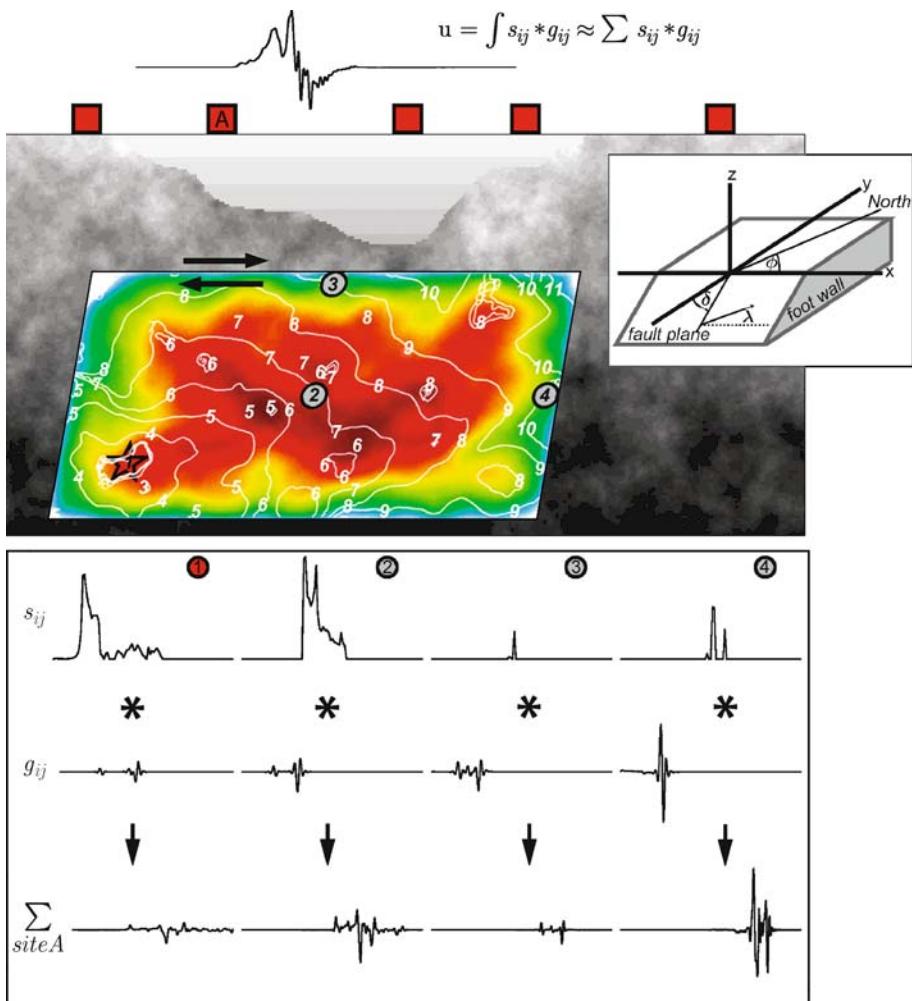
tical fault) moves horizontally to the right, so the opposite side moves to the left;  $\lambda = 180^\circ$  – *right-lateral strike-slip*, i. e. the hanging wall (or near-side of a vertical fault) moves horizontally to the left, so the opposite side moves to the right. For  $\lambda = 90^\circ$ , the hanging wall moves upward (*thrust faulting*), for  $\lambda = 270^\circ$  the hanging wall moves downward (*normal faulting*). Figure 5 also displays the rupture model for a right-lateral strike-slip earthquake with  $\lambda = 180^\circ$  (indicated by the black arrows) on a fault-plane that dips  $80^\circ$ . The strike in this hypothetical case is undefined. The amount of displacement (slip) on each point of the fault is color-coded, white contours (at  $\Delta t = 1$  s spacing) show the expanding rupture front. Examples of past earthquakes show that the rake angle may also vary over the fault plane (Fig. 6c, d). Geometrical complexity, manifested in several fault segments, is often present (Fig. 9); in these cases the faulting-style is characterized by the predominant fault direction and slip angle.

The strength of seismic radiation is typically quantified by an earthquake magnitude, whereby a variety of magnitude scales exist. The seismic moment, defined as  $M_0 = \mu \cdot A \cdot D$ , ( $\mu$ : shear modulus in the source region;



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 4**

Near-field ground accelerations (top) for the fault-normal **a** and fault-parallel **b** motions during the 1999 Chi-Chi earthquake; waveforms are plotted for selected stations along the fault trace (red line). **c** and **d** show the corresponding ground-velocities. Sites located at either end of the fault trace recorded waveforms with shorter duration of the dominant wave-energy. Velocity records in the North show strong rupture-directivity effects as high-amplitude short-duration velocity pulses

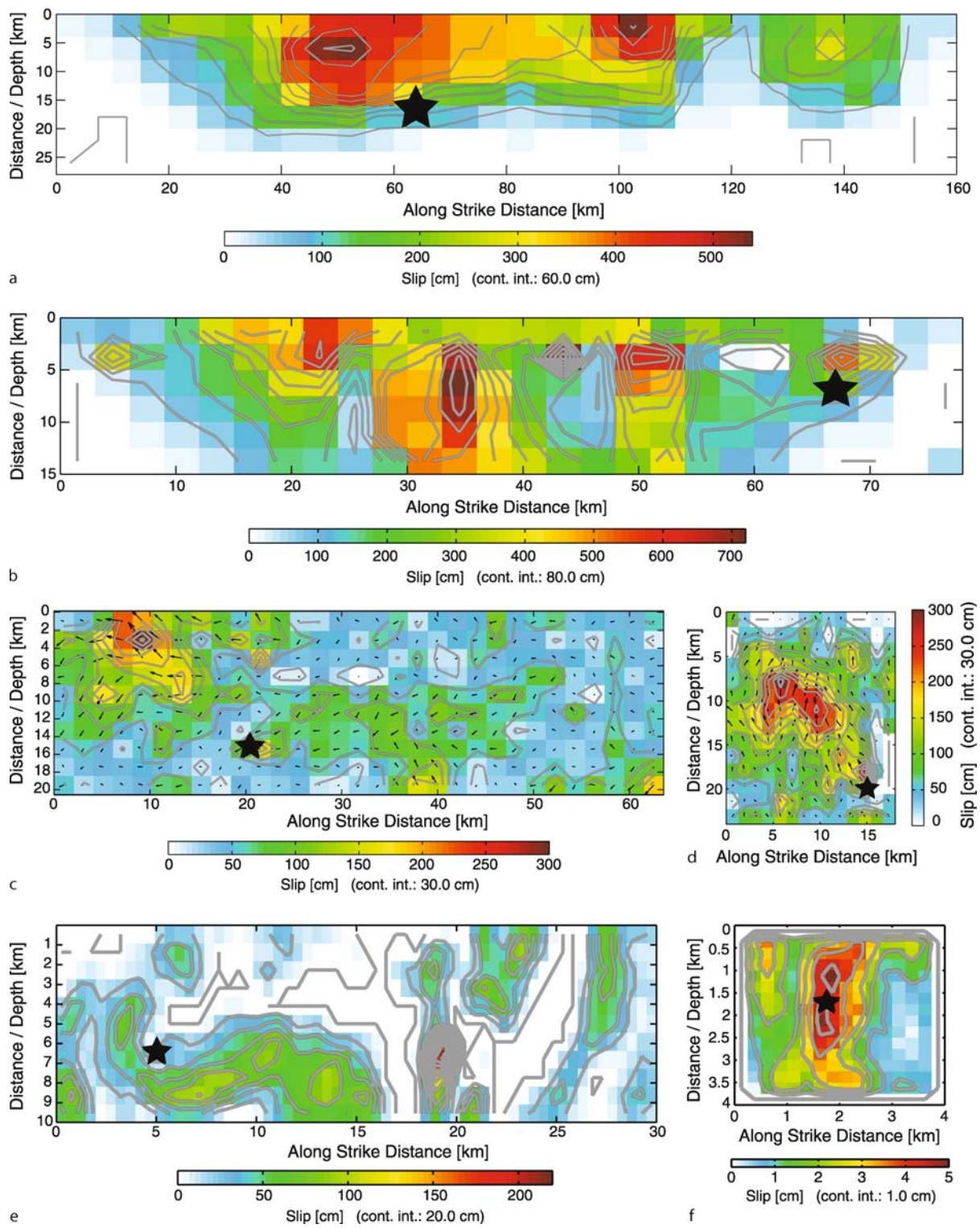


**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 5**

Sketch of the physical components of near-source ground motions (top-right inset: illustrates the source-geometry definition). A simulated earthquake rupture occurs on an assumed planar fault, with color-coded slip amplitudes (0–2.5 m) and white contours of the propagating rupture front (emanating from the hypocenter shown by the red star). The rupture is embedded in a gray-scale model of the Earth' crust. Darker gray-tones denote higher  $v_p$ ,  $v_s$ ,  $\rho$  than lighter gray, reflecting the overall trend of higher velocities at greater depth, locally disturbed by random heterogeneities. A basin-structure with less compliant layered sedimentary rocks exhibits complex sub-surface topography. A near-field seismogram (for an arbitrary horizontal component) is shown at a representative site A. Ground-motions are composed as the summation of the slip- functions  $s_{ij}$  at grid-points  $i, j$  on the fault, convolved with the corresponding Green's functions  $g_{ij}$  for this observer and grid points (the site term  $I_k$  in Eq. (1) is neglected for simplicity)

$A = L \cdot W$ : fault area, given by fault-length  $L$  and fault-width  $W$ ;  $D$ : average displacement on the fault) is considered the best scalar quantity characterizing earthquake size. Ben-Zion [20,23] proposes to use the more basic concept of seismic potency,  $P_0 = A \cdot D$ , to avoid a source quantification that includes ambiguously defined or poorly known material properties; however, seismic potency is rarely used in the earthquake engineering. The moment magnitude follows from  $M_w = 2/3 \cdot \log M_0 - 6.07$  (with  $M_0$  given in Nm). The static stress drop  $\Delta\sigma$  is related to the ra-

tio of average slip to a characteristic fault length, therefore seismic moment is proportional to static stress drop:  $M_0 \propto \Delta\sigma \cdot A^{3/2}$ . Observational evidences [106,186,187] confirm this scaling for moderate to large earthquakes, suggesting – constant (i. e. scale-invariant) static stress drop and self-similar earthquake source scaling: large earthquakes are only scaled-up versions of small earthquakes. Other scaling relations between observable fault parameters (fault area, length, and width; average slip) and magnitude (seismic moment) [79,122,169,187] partially confirm the self-



◀ **Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 6**

Selection of finite-source rupture models, obtained from inversion of seismic recordings and geodetic data, for six earthquakes in the magnitude range  $4.5 \leq M_w \leq 7.5$ , illustrating the heterogeneous distribution of earthquake slip (color-coded using different color scale for each model). The *black star* denotes the hypocenter, generally located in the vicinity of a high-slip patch (also called “asperity”). **a** 1999 Izmit earthquake ( $M7.5$ ) [51]; **b** 1992 Landers earthquake ( $M7.3$ ) [184]; **c** 1995 Kobe event ( $M7.0$ ) [165]; **d** 1994 Northridge earthquake ( $M6.7$ ) [185]; **e** 1984 Morgan Hill event ( $M6.1$ ) [29]; **f** Hida Mountains event #5 ( $M4.5$ ) [96]. The rupture length of **b** is about half that of **a**; models **b**, **c** and **d** are drawn to scale with respect to each other. Model **e** has fault length  $L = 30$  km, a factor of two shorter than **c**, while **f** has source dimensions  $4 \times 4$  km<sup>2</sup>

similarity of earthquake rupture, but also provide evidence for the break-down of self-similarity for very small earthquakes [23] and very long strike-slip earthquakes [154]. The exact physics leading to these source-scaling properties is a current topic of active research [90,91].

The overall faulting-style, defined by the dip and rake angle, affects the radiation pattern of P- and S-waves (see Chap. 4 in [6] for more details), and thus the resulting ground-motion characteristics. Together with the earthquake magnitude, the style-of-faulting is used as a predictor variable in empirical ground-motion prediction equations (see Sect. “Future Directions”). However, ground-motion complexity arises largely from the details of the rupture process (Fig. 5). The earthquake nucleates at the hypocenter (indicated by the red star), and propagates over the fault plane (white contour lines) with a rupture velocity that may exhibit local variations due to initial stress conditions and frictional properties on the fault. Stress state and friction also determine the slip-time history  $s_{ij}$  for all points  $ij$  on the fault (which start sliding once reached by the propagating rupture front). To first order, the shape of the local slip-velocity function depends on the overall fault dimensions (i. e. fault aspect ratio), the relative position of each point with respect to the rupture nucleation point and its distance to the closest fault edge [52,58,59,78] (variations in  $s_{ij}$  at four different locations on the fault plane of lower panel). The duration of slip (rise time) is related to the length of these local slip-functions. Once sliding has stopped at all points on the fault, the rupture has attained its final slip distribution (color-coded in Fig. 5) within a characteristic time (called rupture duration) which depends on fault dimensions and rupture velocity (and marginally on rise time). As depicted in Fig. 5, the slip distribution, the rupture velocity and the rise times are highly heterogeneous over the fault plane, illustrating earthquake source complexity in terms of kinematic source parameters.

The governing equation that relates ground displacement to the motion on the fault is given by a representation theorem (Chap. 3 in [6]), which we use in the following notation:

$$\mathbf{u}(\mathbf{x}, t) = \int d\tau \int_S \Delta u(\xi, \tau) c_{ijkl} v_j G_{nk,l}(\xi, \tau; \mathbf{x}, t) dS. \quad (2)$$

Equation (2) states that the time-dependent ground displacement at observer location  $\mathbf{x}$  and time  $t$  depends on the space-time integral over the space-time-dependent slip function  $\Delta u(\xi, \tau)$  on the fault plane ( $\xi$  defines the position on the fault,  $\tau$  is time), the elasticity tensor  $c_{ijkl}$ , the fault-normal vector  $v_j$ , and the Green's tensor  $G_{nk,l}$  (subscripted comma indicate the derivative with respect to the subsequent variable). Equation (2) is related to Eq. (1), but contains the functional dependency on the source explicitly; in this context the Green's function typically neglects the site effects  $l_k(t)$ . The term  $\int_S \Delta u(\xi, \tau) dS$  contains the time-dependent local particle motion on the fault,  $p_{ij}(t)$ , and the effects of fault finiteness,  $f(t)$ .

Equation (2) can be applied to retrieve the spatio-temporal slip distribution on the fault plane from seismic recordings (and other data) by either forward-modeling or a formal inversion procedure. Earthquake-source inversions have been carried out since the early 1980ies, manifesting the complexity of the rupture process as depicted in Figs. 6 and 9 (see also database of finite-source rupture models [121]). These earthquake source images represent kinematic rupture models that quantify  $\Delta u(\xi, \tau)$  based on observations, but do not explicitly derive  $\Delta u(\xi, \tau)$  from physical principles as attempted in dynamic rupture models e. g. [14,15,28,78].

For examining earthquake source complexity in more detail, it is mandatory we need to distinguish kinematic from dynamic rupture models, and we need to quantify slip heterogeneity (and its associated stress-change distribution) on the fault plane. It is also important to note that the position of the rupture nucleation point is not arbitrary on the fault plane for a given slip (stress) distribution but adheres to fundamental concepts of energy balance during the rupture process. This Section characterizes source complexity in space and time before introducing to *isochrone theory*, a powerful tool to visualize how the details of the rupture process determines near-fault ground-motions.

### Kinematic Rupture Models

Kinematic rupture models characterize the space-time evolution of earthquake rupture in terms of a time-dependent displacement field (distribution of slip vectors)

on a predefined fault plane without considering the forces and stresses that cause these motions on the fault. The local slip-rate (or slip-velocity) function is specified along with the rupture propagation properties. Ground motions can then be computed using Eq. (2) with any kinematic rupture model  $\Delta u(\xi, \tau)$ . Current ground-motion simulation approaches are largely based on kinematic source models since they can be efficiently generated; recent advancements try to capture at least the basic principles of source dynamics by proposing *pseudo-dynamic* source models [77,78,120]. Earthquake source inversions parameterized in terms of Eq. (2) retrieve only a kinematic rupture model which, in principle, does not need to obey any physical laws. The database of finite-source rupture models [121] provides a compilation of kinematic source models for past earthquakes, obtained by applying Eq. (2) to a variety of data and using different methods to solve the inverse problem.

### Dynamic Rupture Models

Dynamic rupture models build a physical understanding of the earthquake rupture process based on the material properties around the source volume, and the initial and boundary conditions for the forces/stresses acting on the fault plane [80]. The distribution of slip vectors  $\Delta u(\xi, \tau)$  on the fault plane is obtained by solving the elasto-dynamic equations of motion under the assumption of a constitutive law (i. e. a friction model), considering essentially the energy balance at the crack tip during rupture growth (for details see Chap. 11 in [6]). Dynamic rupture models have been developed for (i) canonical model to study general feature of dynamic rupture e. g. [10,58,59] (ii) for existing kinematic source models to infer their specific dynamic rupture process e. g. [14,15,126,181] (iii) for heterogeneous initial conditions in stress and/or friction and material distribution to investigate rupture details for classes of events e. g. [11,13,63,68,78,83,133,151]. Due to the high computational demands, dynamic rupture models are not (yet) developed routinely for ground-motion simulations, but rather to investigate source physics for general cases, earthquakes of special interest, and to study rupture behavior for certain classes of initial conditions.

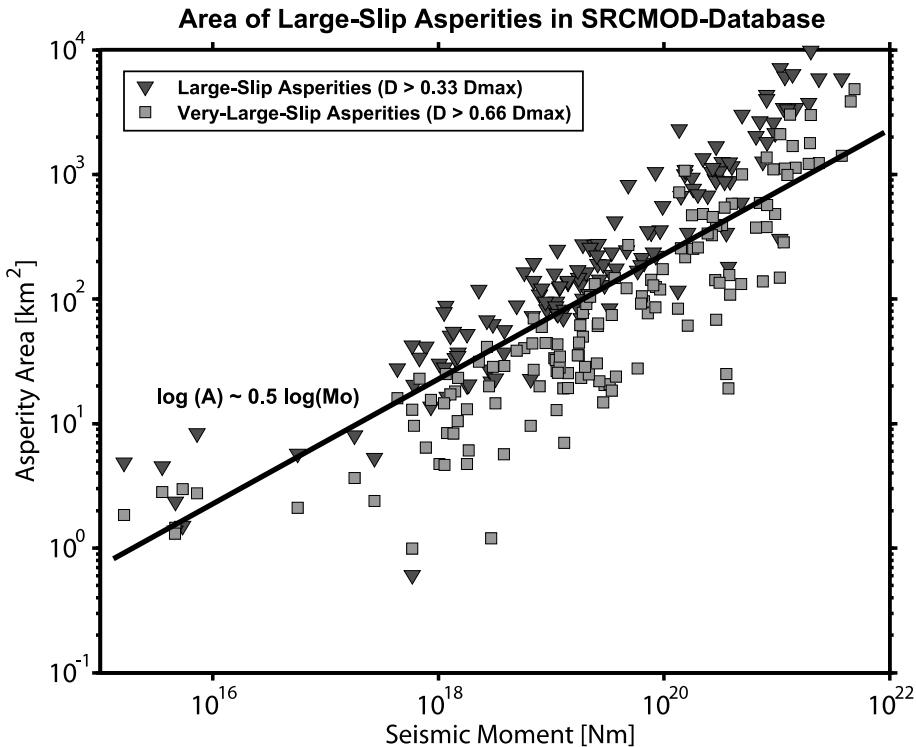
### Quantifying Slip Heterogeneity

Recent approaches to characterize and quantify slip heterogeneity demonstrate that slip distributions exhibit statistical properties and empirical laws grounded in physical principles. First-order observations for these slip distributions (obtained in finite-source inversions based on Eq. (2)), can be made in Fig. 6, a compilation of

slip models for six earthquakes in the magnitude range  $4.5 \leq M_w \leq 7.5$ . The source dimensions increase from  $4 \times 4 \text{ km}^2$  to  $25 \times 160 \text{ km}^2$  while the corresponding maximum fault displacements grow by two orders of magnitude (from  $\sim 5 \text{ cm}$  to over  $500 \text{ cm}$ ). Mai and Beroza [122] have shown that the overall source-scaling relations of such rupture models is roughly consistent with global earthquake scaling laws (i. e.  $M_o \propto A^{3/2}$ ), but that there is evidence that the commonly assumed self-similar constant stress-drop scaling may not hold, because slip on the fault does not saturate but keeps increasing for growing fault dimensions (albeit at a progressively lower rate). The topic of general earthquake scaling laws is still hotly debated e. g. (see ▶ [Earthquake Scaling Laws](#)) [90,91], and directly affects ground-motion prediction [188,189] (see Sect. “[Future Directions](#)”).

More fundamentally, earthquake slip is heterogeneous on the rupture plane (Fig. 6), i. e. regions of little displacement are separated from areas of high slip (often called “asperities”). The distribution and properties of these high-slip patches strongly influences seismic radiation and hence near-fault ground motions. Characterizing and quantifying slip heterogeneity is thus important for accurately predicting and simulating ground-motions for future earthquakes, but also to better understand the physics of earthquake rupture. Two basic avenues have been pursued in the recent past to quantify slip complexity: (i) a deterministic approach that counts the number of high-slip patches and extracts their properties in terms of size, displacement and other quantities [127,169]; (ii) a stochastic approach that characterizes slip heterogeneity in terms of a random-field model [113,114,123]. Results of both methods can be used for simulating stochastic slip distributions for ground-motion calculation to investigate how source complexity affects near-field ground shaking.

Before interpreting and utilizing inferred source-rupture models for future science or engineering applications, the reliability and resolution of these source-inversion images needs to be addressed. Examining the intra-event variability (source models for the same earthquake obtained by different research groups) often reveals large differences in slip distributions and temporal rupture parameters for the set of source models (online-database [121] for examples and [25] for a critical review). The degree of (dis-)similarity between these slip histories is controlled by differences in the Green’s functions (due to computational methods and the choice of the Earth model), variations in the fault parameterization, different inversion schemes and their control parameters (tuning parameters in non-linear inversion; damping/smoothing constraints in linearized multi-time-window inversion). Source-rup-



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 7**

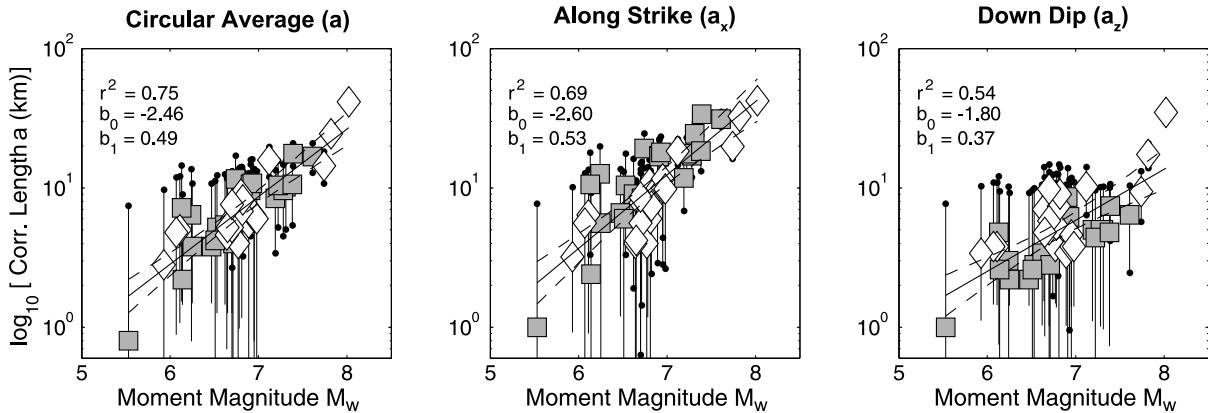
Asperity area, measured for ~90 rupture models (<http://www.seismo.ethz.ch/srcmod>), plotted against seismic moment. A scaling of  $\log_{10}(A) \propto 0.5 \cdot \log_{10}(M_0)$  is shown for reference. Asperity size is measured based on slip values being a certain fraction of the maximum slip on the fault [127]:  $1/3 \cdot D_{\max} \leq D < 2/3 \cdot D_{\max}$  for "large-slip asperities",  $D \geq 2/3 \cdot D_{\max}$  for "very-large-slip" asperities

ture models are also affected by the type of data (seismic, geodetic, or both; additional geologic data), their selection and processing, and their weighting in the inversion.

Consequently, variations in inverted source models are expected, but the actual uncertainties are rarely quantified. Only recently, more efforts are devoted to perform rigorous uncertainty estimation by testing different inversion algorithms, model parameterizations, and data selection criteria for the same earthquake [56,86,116,117]. The work by Monelli and Mai [128] even estimates posteriori probability density functions for the model parameters of interest, using a non-linear inversion strategy coupled to a Bayesian inference technique. Despite the variability in imaged slip distributions for a given earthquake, several source quantities are stably estimated: rupture dimension, seismic moment and average displacement (both within a factor of 2 generally), and also the slip near the hypocenter. Estimates for the average rupture velocity may vary strongly (up to ~ 30%) between models for the same event because rupture speed estimates trade off with the slip-rate function. The locations of high-slip patches on the fault ("asperities") are relatively well located if sufficient data

are used, but exceptions exist (i. e. the 1999 Izmit earthquake). The intra-event variability of maximum fault slip is generally quite high, and depends on the particular inversion strategy. However, estimates of correlation lengths or dominant wave-lengths of heterogeneous slip maps are robust for the different models of a given earthquake [123], indicating that the statistical properties of earthquake rupture are well imaged. Also the hypocenter location with respect to the regions of high-slip is estimated reliably in finite-source inversions [127].

Measuring slip heterogeneity deterministically by defining large-slip ( $\frac{1}{3} \cdot D_{\max} \leq D < \frac{2}{3} \cdot D_{\max}$ ) and very-large-slip ( $D \geq \frac{2}{3} \cdot D_{\max}$ ) asperities ( $D = D(x, z)$  is the local slip on the rupture plane,  $D_{\max}$  the corresponding maximum slip), [127] examine the scaling of total asperity area ( $A_{ta}$ ) with respect to seismic moment for 90 finite-source rupture models (Fig. 7). Without a formal regression, the measurements suggest the scaling  $\log_{10}(A_{ta}) \propto 0.5 \cdot \log_{10}(M_0)$ , meaning that the area occupied by high-slip patches grows slower with increasing magnitude than the total fault size (which scales as  $\log_{10}(A) \propto \frac{2}{3} \cdot \log_{10}(M_0)$ ). This scaling requires relatively larger maximum displace-



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 8**

Correlation lengths,  $a$ ,  $a_x$ ,  $a_z$ , versus moment magnitude  $M_w$  for 44 slip models (modified after [123]). Filled squares denote strike-slip earthquakes, open diamonds represent dip-slip events, vertical lines and black dots mark the  $1\sigma$ -error-estimates. The slope of  $b_1 \approx 0.5$  of the regression curves (solid lines) indicates that the correlation lengths scale self-similarly with moment magnitude

ments on the high-slip patches for large and very large earthquakes, which in turn affects seismic radiation and the scaling properties of near-fault ground motions. Large fault slip, concentrated on relatively small portions of the rupture plane, lead to large stress heterogeneity (with high static stress drop locally), thus implying large ground-motion variability for moderate to large earthquakes.

Alternatively, slip distributions (Fig. 6) can be treated as spatial random fields [113,123] to estimate the fractal dimension or the correlation lengths for an auto-correlation function (ACF). Such measurements are typically based on the two-dimensional power spectrum  $P(\mathbf{k})$  of the slip map. Assuming a simple fractal model for a random field, its power spectral density is given as

$$P(\mathbf{k}) \propto \mathbf{k}^{-\nu-1} \quad (3)$$

with wavenumber vector  $\mathbf{k}$  and scaling exponent  $\nu$ . [113] estimate one-dimensional scaling exponents for a small number of rupture models, using only horizontal slices of slip distributions, and find  $0.8 \leq \nu \leq 1.5$ , while [123] find for the two-dimensional exponents ( $\nu + 1$ ) values of about 1.7, implying a fractal dimension  $D = 2.3$ . The differences in these estimated scaling exponents can be reconciled when accounting for slightly different initial processing of the slip distributions and fitting strategies [114]. Testing a variety of auto-correlation functions to fit the spectral properties of a large number of slip distributions, [123] observe that a von Karman auto-correlation function, with its power-spectral density given by

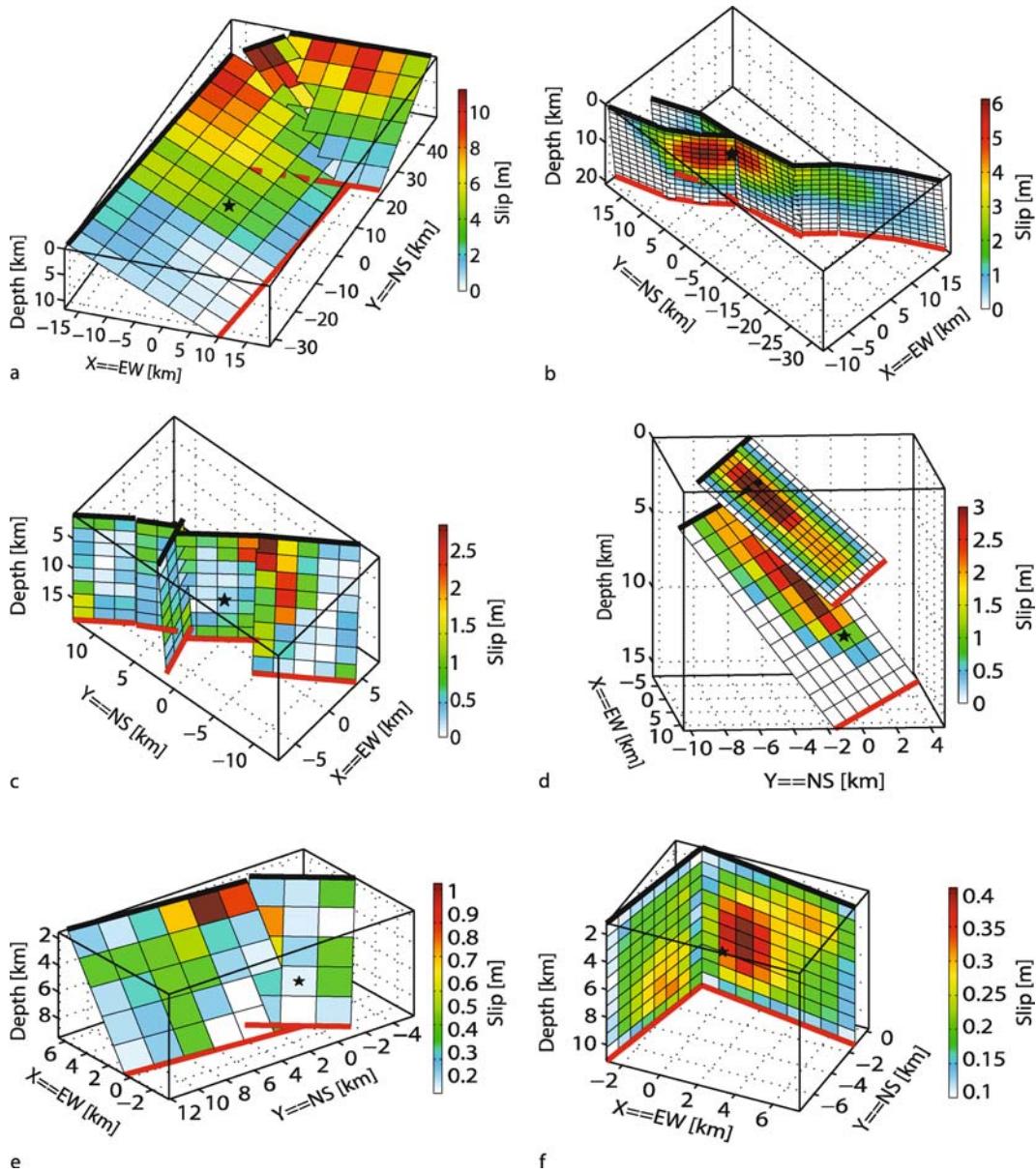
$$P(\mathbf{k}) = \frac{4\pi H}{K_H(0)} \cdot \frac{a_x \cdot a_z}{(1 + \mathbf{k}^2)^{H+1}}, \quad (4)$$

with magnitude-dependent correlation lengths  $a_x$ ,  $a_z$ , and a scale-invariant Hurst number  $H \approx 0.7$  best matches the slip heterogeneity spectra ( $K_H$ : modified Bessel function order  $H$ ;  $\mathbf{k}$  wavenumber). Figure 8 displays their estimated correlation lengths for 44 slip models [123], along with a least-squares regression that exhibits scaling relations of the form

$$\begin{aligned} \log_{10}(a_x) &\propto \frac{1}{2} M_w \\ \log_{10}(a_z) &\propto \frac{1}{3} M_w. \end{aligned} \quad (5)$$

Interestingly, the scaling law in Eq. (5) for correlation lengths of heterogeneous slip maps is similar to the relationship inferred for the deterministic measure of asperity size. This corroborates the previous argument that correlation lengths (asperity size) increases with increasing magnitude, though at a lower rate than the overall fault dimensions. Thus, to accommodate the corresponding seismic moment, the displacements (and associated stress drops) on these high-slip patches need to grow faster than a self-similar scaling would predict. This conclusion agrees with the conjecture of Heaton [88] that large earthquakes require a “very strong kick” (due to an area of large stress drop) in order to be able to grow into a very large rupture.

Treating earthquake slip as being distributed on a *planar* fault surface strongly simplifies geologic observations of geometrical fault complexity, where the degree of complexity depends on fault-zone maturity [20,90]. Earthquakes that break several fault segments can still be imaged with Eq. (2) using an appropriate parameterization, but planar-fault ruptures exhibit very different rupture dy-



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 9**

Three-dimensional views of heterogeneous slip on geometrically complex faults, imaged using seismic and/or geodetic data. The black star shows the hypocenter, thick black lines are the top of the fault segments, thick red lines the bottom. a The 1999 Chi-Chi earthquake ( $M_w 7.6$ ) [101]; b the 1999 Hector Mine earthquake ( $M_w 7.2$ ) [102]; c the 2000 Tottori earthquake ( $M_w 6.6$ ) [100]; d the 1971 San Fernando event ( $M_w 6.6$ ) [87]; e the 2003 Miyagi-hokubo event ( $M_w 6.1$ ) [89]; f the 1997 Kagoshima earthquake ( $M_w 6.0$ ) [94]

namics than earthquakes that have to overcome geometrical obstacles [82,84,131,134]. Figure 9 shows a collection of such geometrically complex fault models, involving two or more segments that form either a system of sub-parallel fault planes (Fig. 9d), or a branching fault (Fig. 9b), or fault planes oriented at arbitrary angles to

each other (Fig. 9a, c, e, f). The seismic radiation from earthquakes on geometrically complex faults is more complicated than from a single-plane rupture, a topic that is investigated in the earthquake source dynamics community e.g. [14,15,43,81,82,84,131,134]. Unfortunately, little attention has been devoted in the past to rigorously exam-

ine and quantify the degree of ground-motion complexity generated by such fault systems. Thus, future research, using advanced numerical techniques on high-performance computing architecture, needs to properly represent fault systems at sufficient spatial resolution to capture the intricacies of rupture dynamics and associated seismic radiation on geometrically complex faults.

### Rupture Nucleation and Directivity

Besides the slip heterogeneity and the temporal rupture evolution (discussed later), the location of the hypocenter (point of rupture nucleation) is a critical factor affecting near-source ground motions. Some metric of hypocentral distance enters empirical attenuation relations (Sect. “[Ground-Motion Scaling Relations](#)”), while the on-fault hypocenter location determines the directivity effect [170,171]. This global directivity effect due to fault-hypocenter-site geometry results in large velocity-pulses, in particular at sites close to the fault, because of the constructive interference of S-wave energy which is continuously radiated from the propagating crack front and arrives within a short time window in the forward-direction of rupture propagation (see velocity records in Fig. 1 at station SlackC, FZ1, SC1E for the Parkfield event; Fig. 4 at station TCU052, TCU068 for the Chi-Chi event). In contrast, for sites in the backward direction of rupture propagation the arriving seismic energy is spread over a longer time interval, generating lower-frequency motions with smaller amplitudes. The directivity effect is therefore not a purely source-related phenomenon, but also depends on the observer location.

Additionally, there is an “on-fault” directivity effect which operates on smaller scales and is most prominent for large earthquakes and very near-fault sites. The on-fault directivity effect can be efficiently quantified using isochrone theory (see Sect. “[Isochrone Theory](#)”). For an explanatory description, consider the slip distribution and hypocenter location in Fig. 6a, for which the large-scale “global” directivity effect was observed towards the right (i.e. East in this case of the 1999 Izmit earthquake). However, any site located at about  $30 \leq X \leq 50$  km in along-strike distance (above the left-most high-slip region) experiences strong “local” directivity effects as the rupture propagates from the hypocenter towards the site and across the high-slip patch. In this case, the integrated slip along this rupture trajectory generates a high-amplitude short-duration pulse. Considering a more westerly hypocenter (e.g. at  $X = 30$  km instead of  $X = 65$  km) the large-scale directivity effect remains essentially unchanged while the integrated slip along the rupture trajectory for

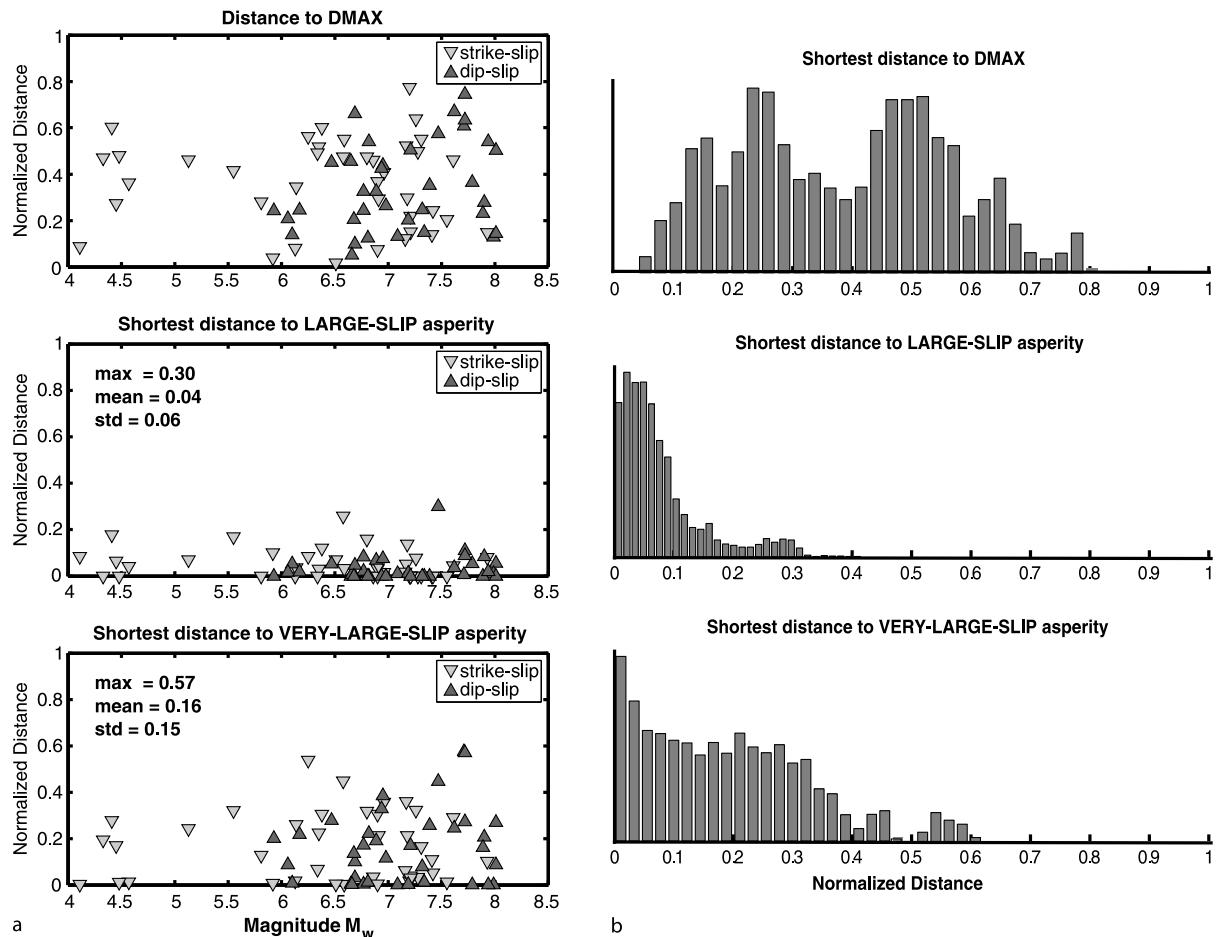
stations at  $30 \leq X \leq 50$  km is much smaller, thus diminishing on-fault directivity. The spatial relation between the hypocenter and large-slip regions plays an important role for ground-motion complexity, but is also fundamentally related to the dynamics of the rupture process.

Examining the hypocenter positions in Fig. 6 and Fig. 9, two observations are evident: (a) ruptures generally do not nucleate at the fault boundaries, but in the interior of the fault plane, but rarely exactly in the center; (b) the nucleation point is generally not located in regions of low slip (light colors,  $D \leq 0.1 \cdot D_{\max}$ ) but in areas where  $0.2 \cdot D_{\max} \leq D \leq 0.4 \cdot D_{\max}$  and the distance to a nearby large-slip zone (asperity) is small compared to the overall source dimensions [127]. This study investigated these first-order observations statistically using a database of  $\sim 80$  finite-source rupture models, and concluded that hypocenters are not randomly located on a fault but are located either within or close to regions of large slip (Fig. 10). More specifically, ruptures nucleate within or very close to large-slip asperities ( $\frac{1}{3} \cdot D_{\max} \leq D < \frac{2}{3} \cdot D_{\max}$ ), but rarely start on very-large-slip asperities ( $D \geq \frac{2}{3} \cdot D_{\max}$ ; often located far away from the hypocenter).

These observational constraints of slip heterogeneity and rupture nucleation are rooted in the energy balance of earthquake source physics, and have been confirmed by dynamic rupture simulations [78,133,151]. The essence of the physical mechanism is that a sustained large earthquake can only be generated if a sufficiently large amount of energy is furnished to the propagating crack tip to facilitate rupture growth. In a simplistic view, an earthquake can only grow in size if the energy absorbed to create new crack surface (fracture energy) balances the available elasto-static energy and the energy radiated by seismic waves (Chap. 11 in [6]). If the hypocenter is located in regions of low slip (low stress drop) and far away from any point of significant stress drop, the required fracture energy will become too large to allow further crack propagation, and the rupture will stop prematurely. The spatial correlation between hypocenter and asperity is thus consistent with this simplified energy budget of dynamic rupture. Moreover this hypocenter-asperity relation manifests the “on-fault” directivity and strongly effects seismic radiation and thus ground-motion complexity.

### Temporal Rupture Evolution

The characteristics of the slip distribution and the hypocenter position alone are insufficient to parameterize  $\Delta u(\xi, \tau)$  in Eq. (2) because the temporal rupture evolution is not yet specified. Rupture velocity and the local slip-velocity function with its associated rise time (slip duration)



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 10**

Hypocenter distance to “large-slip” and “very-large-slip” regions in finite-source rupture models. **a** Raw measurements of shortest distances between hypocenter and point of maximum displacement (top), hypocenter and closest large-slip asperity (middle), and hypocenter and closest very-large-slip asperity (bottom), separated into strike-slip and dip-slip earthquakes. **b** Distributions of shortest distances between the hypocenter and the point of peak-slip, the closest large-slip, and the closest very-large-slip asperity, generated using Monte Carlo simulations that include hypocenter uncertainties. The distributions indicate that 16% of all hypocenters occur on a very-large-slip asperity, 35% within a large-slip asperity, and 48% right outside an asperity (compiled from [127])

need to be given, either pre-defined by means of a kinematic rupture modeling approach or solved for from physical principles in a dynamic rupture model.

**Rupture Velocity** After earthquake nucleation the rupture propagates over the fault plane with a rupture velocity  $v_r$  which generally is close to the local shear-wave velocity (typically in the range  $0.6 \cdot v_s \leq v_r \leq 0.9 \cdot v_s$ ). Seismic radiation and thus near-field motions strongly depend on rupture velocity: slow earthquakes radiate little seismic energy while fast ruptures generate higher ground-motion amplitudes. Earthquake source inversions mostly assume a constant rupture velocity over the fault plane, although dynamic rupture modeling and evidence from par-

ticularly well recorded earthquakes indicates that rupture speed may vary significantly on the fault plane [16, 44]. Dynamic modeling shows that the initial rupture speed (during and right after the nucleation phase) may be small (e.g.  $v_r^{\text{init}} \leq 0.5 \cdot v_s$ ), but then rapidly increases during the rupture’s dynamic breakout and propagation phase (Chap. 11 in [6]). Depending on the stress conditions and the frictional parameters on the fault, and the geometrical properties of the rupturing fault, the rupture speed may actually exhibit very strong small-scale variations (see Fig. 5 for example). Variations in rupture velocity are a source of high-frequency seismic radiation [119, 173], and thus strongly contribute to near-field ground motion complexity.

While the rupture speed for most earthquakes is lower than the local S-wave velocity (i. e. sub-shear rupture propagation), there is evidence from laboratory measurements [155] and from seismic data and source modeling e. g. [16,44,134] that the crack front may travel at super-shear speeds at least over parts of the fault plane. Such super-shear ruptures generate rather distinct ground-motion characteristics at the few sites where they were recorded, but it is not yet clear what the “generic” ground-motion signature of super-sonic rupture speed would be. [1] address some aspects of this question by kinematic rupture modeling with various rupture speeds and computing the resulting near-field motions, indicating very peculiar “mode-switching” of the seismic-energy distribution on the two horizontal components as the rupture speed becomes very large. Furthermore dynamic rupture simulations [10,58,59,63,68,134] provide physical models for the occurrence of super-shear propagation, which depends on local stress and frictional conditions on the fault plane. The potential occurrence of super-shear rupture burst [62,63] and rupture speed variations in general are critical for understanding ground-motion complexity. To efficiently include such effects into ground-motion simulations, [78] developed an initial physics-based representation of rupture-velocity heterogeneity for kinematic source modeling.

**Slip-Velocity Function and Rise Time** As the propagating rupture sweeps over the fault plane points on the fault plane are “activated” and start slipping. Each point traces an individual slip-time history (slip-velocity function) whose shape and duration depends on the stress conditions and frictional properties on the fault, but also on overall fault size and the position of each point with respect to rupture nucleation [59]. Figure 5 illustrates the variability of local slip-velocity functions  $s_{ij}$  on the fault. The slip duration (rise time  $\tau_r$ ) is usually defined by integrating the slip-velocity function and measuring the time it takes to complete 5–95% of the total displacement at each point.

Several approximations of these complicated slip-velocity functions are in use: a boxcar or isosceles-triangle function with a rise time equal to the width of the function, or symmetric and asymmetric cosine-functions. The classical approximation uses the solution of a quasi-dynamic crack model [110], showing an  $1/\sqrt{t}$ -decay after a rapid onset. Recent modifications to this Kostrov-type slip function [129,180] are compatible with earthquake dynamics, while ground-motion simulations and source inversion often assume simple parameterization of overlapping triangles [78,92].

As Fig. 5 indicates, rise time varies over the fault plane. However many kinematic source inversions and simple source-model simulations assume constant rise time. A self-similarly expanding crack model predicts longer rise times in the center of the fault e. g. [58,110] where the rupture nucleates, but short rise times at the crack periphery due to the earlier arrival of the healing front from the crack rim. Analysis of source-inversion results [88] and dynamic rupture modeling [28] however suggest that rise times are in general short, i. e. the rupture propagates as a self-healing pulse over the fault plane with rise times determined by local healing of the rupture front.

For fixed rupture velocity and slip distribution, ground-motions are very sensitive to rise time variations. Shorter rise times lead to larger ground-motions as the seismic energy is released in a shorter time interval. The detailed shape of the slip-velocity function is less important for the seismic waveforms, because they are determined by the summation of slip-functions, convolved with the appropriate Green’s function, over the entire fault plane (effectively filtering out small-scale features of the slip-velocity parameterizations). However, the peak slip-velocity value that the slip-rate function may attain is crucial for the final ground-motion amplitudes.

### Isochrone Theory

The effects of various kinematic source parameters (slip and slip-velocity distribution, rupture propagation) on near-field ground-motions can be efficiently visualized in the framework of the isochrone theory, a high-frequency (ray-theory) approximation to calculate seismic radiation from earthquake ruptures [27,173]. By approximating the elastic wave Green’s functions with far-field body waves, surface-waves and low-frequency near-field terms are not considered in the isochrone method. It is correct, however, at higher frequencies, where far-field terms dominate, and in the distance range to the fault where surface waves have not yet developed but near-field terms are unimportant. The main aspect of this approach is that the space-time integration in Eq. (2) can be replaced by a series of line integrals over the fault for simple slip-function parameterizations. For each time  $t_i$  in the observer seismogram a line integral is computed for which the integration path comprises only those points on the fault which radiate seismic waves that arrive at the observer location at exactly the time  $t_i$ . Defining an arrival-time function, composed of the rupture-front arrival time at each point on the fault and the corresponding seismic-wave travel time to the observer, the integration path represents an isochrone of this arrival-time function. The corresponding isochrone veloc-

ity,  $\mathbf{c}$  (the spatial derivative of the arrival-time function) is related to rupture velocity  $v_r$ , and thus resembles the directivity function. Since ground velocity is proportional to isochrone velocity, the characteristics of high-frequency body-waves in near-field seismograms can easily be examined with respect to arbitrarily heterogeneous slip and rupture velocity distributions.

The omission of near-field terms and surface-waves restricts the isochrone method to cases in which seismic radiation from compact slip zones or sudden changes in rupture velocity dominate the ground-motions. Such source behavior has been reported for many earthquakes, consistent with [119] who showed that the high-frequency far-field radiation is emitted at the propagating crack tip, with additional radiation coming from stopping phases as the rupture heals. The isochrone theory has also been successfully applied to earthquake source inversions [29].

The detailed theoretical development of the isochrone theory is beyond the scope of this article, but a brief description will help to illustrate the concept. Spudich and Frazer [173] link the representation theorem, Eq. (2), to geometrical ray theory (without Fraunhofer approximation) for the far-field displacements of P- and S-waves. According to [176] a simple, yet versatile, parameterization of the slip function  $s(\mathbf{y}, t)$ , is given by

$$\mathbf{s}(\mathbf{y}, t) = \mathbf{s}_r(\mathbf{y}) f_r[t - t_r(\mathbf{y})], \quad (6)$$

with rupture time  $t_r$  at positions  $\mathbf{y}$  on the fault plane, position-independent shape-function  $f_r$  for the slip-velocity function, and position-dependent amplitude  $\mathbf{s}$ . Any heterogeneous rupture model can be approximated by Eq. (6). Inserting this expression into Eq. (2), shown here for S-waves only, one obtains

$$\mathbf{u}^S(\mathbf{x}, t) = f_r(t) \cdot \int_S \mathbf{s}_r \cdot G_a^S \delta(t - t_a^S) dS \quad (7)$$

where  $t_a^S(\mathbf{y}, \mathbf{x}) = t_r(\mathbf{y}) + t^S(\mathbf{y}, \mathbf{x})$  is the arrival time function for an observer at location  $\mathbf{x}$  due to an S-wave radiated at point  $\mathbf{y}$  on the fault  $G_a^S$  is the corresponding Green's function. The surface integral in Eq. (7) is non-zero only if the argument of the  $\delta$ -function is zero. Curves  $\mathbf{y}(t_a^S, \mathbf{x})$  define the contours of equal arrival time (isochrones) at observer  $\mathbf{x}$ ; the surface-integral in Eq. (7) thus reduces to a line integral

$$\mathbf{u}^S(\mathbf{x}, t) = f_r(t) \cdot \int_{\gamma(t, \mathbf{x})} \left( \mathbf{s}_r \cdot G_a^S \right) \cdot \mathbf{c}(\mathbf{y}, \mathbf{x}) dl \quad (8)$$

in which  $\mathbf{c}(\mathbf{y}, \mathbf{x}) = |\nabla t^S(\mathbf{y}, \mathbf{x})|^{-1}$  represents the “isochrone velocity” of these curves along the fault surface ( $\nabla$  denotes the gradient operator). Note that the isochrone velocity is

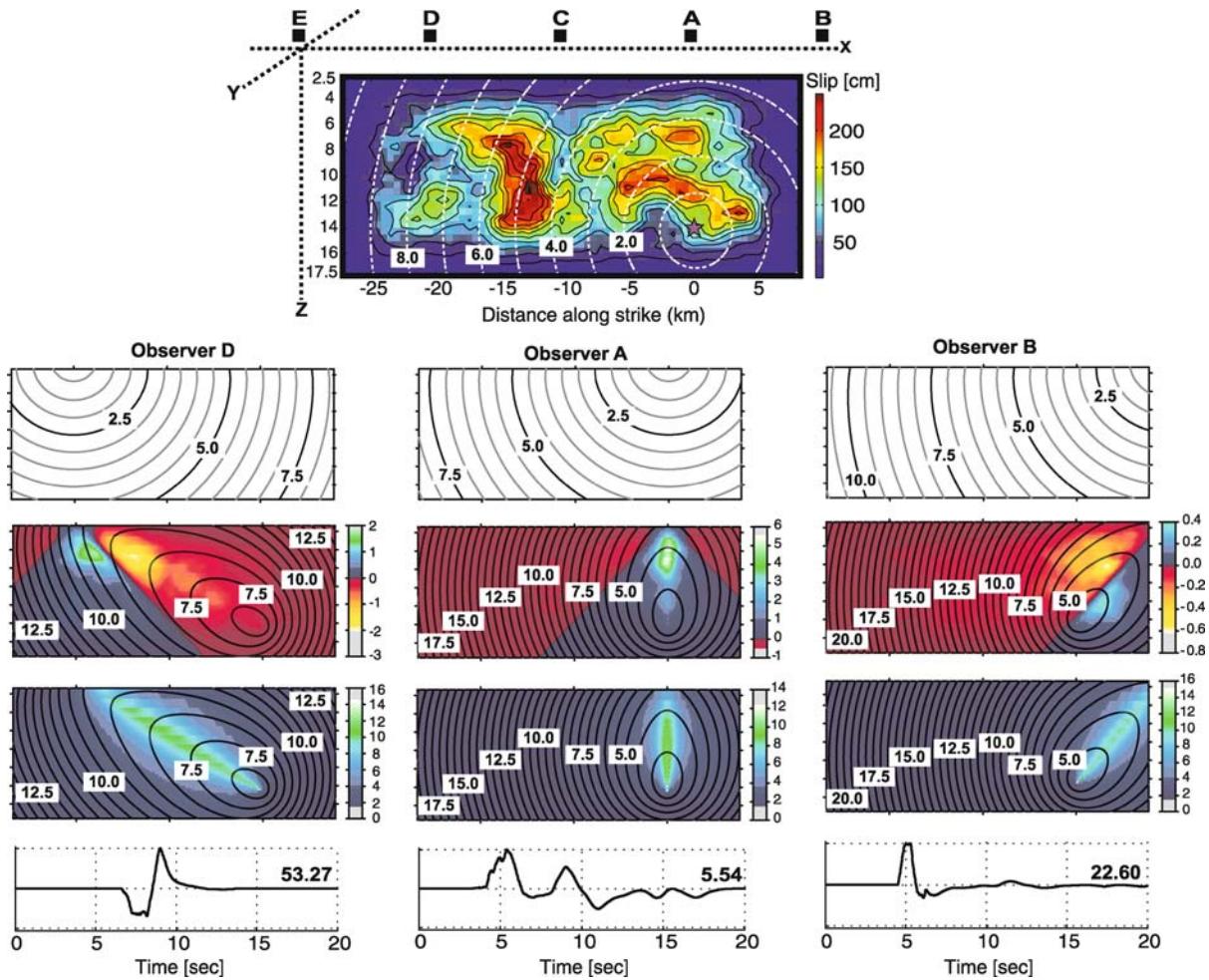
related to the directivity function; isochrone theory therefore helps to visualize on-fault directivity effects [172].

As an application of the isochrone method, we compute near-field seismograms for a hypothetical rupture ( $M_w 6.7$ ) with constant rupture velocity (radially spreading rupture front) at three sites, each at fault-perpendicular position  $y = 0.2$  km, but various positions along-strike (Fig. 11). A simple layered medium is assumed, and anelastic attenuation is excluded. For each observer, the S-wave travel time function (travel times from the site to each point on the fault) is shown (top panels), the color-coded integrand (Eq. (8)) with isochrone contours (spacing  $dt = 0.5$  sec) (2nd row of panels) and the isochrone velocity with isochrone contours (3rd row of panels). S-wave pulses in the fault-normal velocity seismograms (bottom) can then be attributed to isochrone properties. Consider for example observer D. The first isochrone appears at  $\sim 7$  sec, consistent with the S-wave arrival time in the seismogram. Isochrones up to 8 sec are entirely in areas where the integrand is negative (yellow to red colors), hence the down-ward initial motion. The peak velocity (53.2 cm/sec) occurs at  $t \approx 9$  sec when the corresponding isochrones hit the large-amplitude regions of the integrand (green areas). Observer D exhibits stronger directivity than observers A or B due to its position with respect to the hypocenter and the dominant asperity, generating large amplitudes due to high isochrone velocity over regions of large slip. At site D, the seismic energy radiated from the high-slip patch at  $x \approx -15$  km,  $y \approx 10$  km arrives within a short time window, whereas at site A, the contributions from this slip patch are spread over a longer time window. For site A, the isochrone velocity is smaller than for D, hence directivity is less pronounced. As a result, site A exhibits lower peak ground motions but a longer shaking duration.

The current implementation of the isochrone method [176] could be extended to include near-field terms [104] and to be applicable in three-dimensional velocity structures by adding 3D ray-tracing capabilities. Spatially variable slip-velocity functions and anelastic attenuation are further possible extensions. Finally, effects of wave-scattering in random media could be implemented using scattering operators [125,192,193,194] convolved with the local slip-rate function  $s(\mathbf{y}, t)$ .

### Wave Propagation in Complex Media: Path and Site Effects

Returning to Eq. (1), this Section discusses path effects,  $g_k(t)$ , due to source-to-site wave propagation, and local site effects,  $l_k(t)$ , due to the small-scale geological con-



Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 11

Ground-motion simulation using isochrone theory for a hypothetical  $M_w 6.7$  earthquake, buried at 2.5 km depth and embedded in a layered medium. **Top graph:** Rupture spreads from the hypocenter (red star) at constant rupture speed ( $v_r = 2.7$  km/s, white contours) over the fault with heterogeneous slip (color-coded). **Bottom graphs:** Isochrone quantities at three sites. **Top-most panels:** S-wave arrival time function; **2nd row of panels:** isochrones overlain over the real part of the integrand in Eq. (8) (in units of  $s_r \cdot G_a^S \cdot c(y, x)$ ); **3rd row of panels:** isochrones overlain over isochrone velocity (in km/s). **Lower-most panels:** Resulting fault-normal velocity seismograms dominated by large S-wave pulses (in cm/s; peak velocity indicated at the end of each trace)

ditions at each observation point. The wave propagation from the earthquake to the observation point depends on the complexity of Earth structure, which can be separated into two parts: (i) deterministic wave-propagation within a layered medium, three-dimensional basin effects, topographic features and large-scale geological structures; (ii) wave-propagation in a stochastic random medium with small-scale heterogeneity leading to incoherently scattered wave energy. Local site effects in the shallow near-surface structure (usually the top-most 100m) beneath the observation site further complicate near-fault

ground-motions, leading to either increased or decreased motions (compared to bed-rock level), increased shaking duration, shifts in the dominant frequency of ground-motion, and perhaps nonlinearity effect. These wave-propagation effects have to be incorporated into the characterization of ground-motion complexity, and should be applied in strong-motion simulations for predicting the intensity and variability of seismic shaking. However, the success in incorporating these effects hinges on the available knowledge of seismic properties of the Earth, which in turn depends on the dominant wavelengths (frequen-

cies) at which the Earth is sampled by observational data. Similarly, inversion for Earth structure and forward simulation of seismic waves require an appropriate spatial discretization of the modeling region for the numerical calculations, which essentially provides an upper limit on either the maximum resolvable frequency or the size of the computational domain for ground-motion simulations.

In this Section, I first I review some aspects of “deterministic” wave propagation in a flat-layered attenuating Earth, and then qualitatively describe effects of basin structures, Earth topography and local features (e.g. fault zones, narrow belts of low shear-wave velocity). I then examine aspects of wave propagation in random media and corresponding simulation methods. Site effects and nonlinearity conclude this section. Due to the large body of literature on these topics, a detailed quantitative review is not attempted in this Section. Instead I merely select a few key elements of particular interest for characterizing ground-motion complexity.

### Wave-Propagation in a Flat-Layered, Attenuating Earth

The first-order approximation of Earth’ internal structure is a radially layered medium, which for short source-site distances (< 600 km) can be treated as a flat-layered structure. Many research branches in seismology successfully apply this approximation of a one-dimensional depth-dependent velocity-density model: earthquake location, arrival-time measurements of various seismic wave types, focal-mechanism determination, moment-tensor inversion, finite-source inversion. A number of well tested methods exist to compute Greens functions for a flat-layered Earth model, although the exact frequency range in which this model is accurate is difficult to determine and depends on the particular application and region. At long periods ( $T > 20$  sec and corresponding long wavelengths ( $\lambda \approx 120$  km for P-waves traveling at speed of  $v_p \approx 6$  km/sec), used for instance in moment-tensor inversions), basin structures, mountain ranges and other geologic features with minor changes in physical properties (density, wave-speed) can be ignored at spatial scales on the order of tens of kilometers, thus allowing accurate modeling of the complete low-frequency seismogram. For small areas with relatively simple Earth structure, body-wave Greens functions can be accurately synthesized up to frequencies of  $f \approx 1$  Hz, a property that is often used in finite-source inversions. In these cases, ground-motion complexity is assumed to be largely due to the heterogeneity in the rupture process, since the local, observer-specific Greens functions (containing the impulse response in

a layered medium due to point-source excitation) are simple (examples are shown in Fig. 5).

In many cases, however, sites of interest may be located on top of basin structures, in zones that are geologically different on a small-scale, or even in a narrow fault-zone belt which may have distinctly different seismic properties. By examining observed ground-motions at many sites, site-specific one-dimensional Earth models can be used to compute corresponding Greens functions, a cumbersome and error prone approach since source and receiver are located at points with different seismic properties. In such cases, 2D or 3D-models are preferable for the Greens function calculation, as they allow to appropriately include basin effects, local geology and potentially also topography, usually at the expense of reduced frequency resolution and/or a smaller computational domain.

### Effects of Sedimentary Basins, Fault-Zones, Topography

Geological basins, containing layers of compliant sedimentary rock units and covered by potentially poorly consolidated, low shear-wave-velocity sediments, have a variety of effects on incoming seismic waves. Characterizing and quantifying the basin-related ground-motion complexity is particularly important for seismic hazard studies since many major urban areas are built in geological basins that are located within or close to a seismically active region (e.g. Mexico City, Tokyo, Los Angeles).

For instance, the 1985  $M_w$  8.0 Michoacan earthquake generated major damage in Mexico City, several hundreds of kilometers away from the epicenter. While the near-source strong-motion stations did not show unusual ground motions, the shaking level in Mexico City was unexpectedly large in those parts of the town constructed on lake bed sediments [9]. Ground-motions and associated damage exhibited high spatial variability; long-period surface waves, generated at the edge of the basin, had particularly severe consequences for tall high-rise building with eigenperiods of several seconds. Many of them suffered complete collapse. Strong basin effects have been simulated for earthquakes in the greater Los Angeles area, indicating localized basin-response effects caused by surface waves generated at the edge of the basin [74,75,135,136,138] and the focusing of seismic energy due to basin geometry [76,108,138]. A particular interesting effect has been found for the 1994  $M_w$  6.7 Northridge earthquake, attributed to an isolated high-damage area in Santa Monica to focusing effects caused by a small-scale localized 3D lens-like high-velocity structures [57].

Other prominent examples of strong basin response are given for the Wellington Basin (New Zealand) [24], for the Tokyo plain (Japan) [109,159,160,162], and the Kobe area (Japan), which was strongly damaged during the 1995 Kobe earthquake ( $M_w$  6.9). In the latter case, the largest damage did not occur in direct vicinity of the rupturing fault, but was concentrated within an elongated band offset to the southeast of the fault. Modeling studies showed that this “damage belt” was caused by constructive interference of seismic waves taking different paths: (i) directly through the low-velocity basin structure and (ii) through high-velocity rocks outside the basin and then reflected back into the basin [107,143,144]. Additionally, topographic effects may have contributed to the complexity of near-fault motions for the Kobe earthquake [145]. Another example for strong topographic effects on near-source motion is the Tarzana (California) site, located atop a small hill, which recorded a value of  $PGA = 1.78\text{ g}$  on the EW-component of motion during the 1994  $M_w$  6.7 Northridge earthquake [174]. However, it is difficult to separate topographic effects from subsurface-structure properties beneath the topographic features, and firm conclusions on the contribution of topographic effects are not well defined [108].

These basin-induced or topographic site effects strongly depend on the direction from which the incoming wave field arrives, as shown by strong-motion simulations for the Basel (Switzerland) area [139] or the greater Los Angeles region [118,138]. Strong variations in the basin response, and hence ground-motions at individual sites, depend on whether the incoming waves are predominantly polarized in the direction parallel or perpendicular to the main geologic structures (i. e. the large-scale basin shape and major faults inside and bounding the basin).

Low-frequency ground-motions, modulated by basin effects, topography, or narrow fault-zone-related regions of low shear-wave velocity, exhibit also longer shaking duration (due to surface-waves arriving after the dominant S-arrivals), and potentially larger amplitudes in case of constructive interference of wave packets (often occurring as trapped waves) [141]. Due to the dependency of the basin response on the direction of the incoming wavefield and the detailed small-scale structure of the basin, individual sites within a particular basin will experience very different ground-shaking. The ground-motion complexity is thus greatly increased by these effects, and general scaling relations for ground-motion (de-)amplification and prolongation due to basin and topographic structures are difficult to derive. Only numerical simulations for particular earthquakes or specific scenario events can help to understand the corresponding ground-motion variability which

can then be related to standard ground-motion attenuation relationships (see Sect. “[Ground-Motion Scaling Relations](#)”).

While there has been considerable progress in interpreting and modeling strong-motion waveforms for frequencies  $f < 1\text{ Hz}$ , one of the major challenges is to calculate reliable broadband near-source seismograms for the frequency range of engineering interest which extends to  $f \sim 10\text{ Hz}$ . The works described above examine and model the low-frequency wavefield contribution, using 1D-, 2D- or 3D-finite-element, finite-difference or spectral-element techniques (see ▶ [Seismic Wave Propagation in Media with Complex Geometries, Simulation of](#)). Seismic source properties, Earth structure, and site effects strongly affect high-frequency motions, but computational limitations still prohibit purely deterministic ground-motion simulations for frequencies above  $f \approx 1\text{ Hz}$ . Instead, at least some part of the simulation procedure needs to involve a stochastic component since Earth structure, and to some extent the earthquake source, are essentially unknown at short spatial scales required for accurate high-frequency simulations. The difficulty is thus to capture and correctly quantify the scattering properties of the Earth at the scale-lengths and frequencies that are of interest for seismic hazard purposes.

### Scattering in Inhomogeneous Media

The seismic coda, the energy in the seismogram after the prominent direct P- and S-wave arrivals, consist of P-, S- and surface-wave energy scattered in the inhomogeneous rock volume between the source and the recording site. Figures 1, 3, 4, display near-field waveforms for which the coda-characteristics show large variability. In some cases, the coda waves decay fast, in other cases they persist for a long time; the frequency content of coda waves also appears to be site dependent. These observations manifest that, scattering and attenuation of seismic waves in inhomogeneous media strongly contribute to ground-motion complexity.

The basic method of estimating attenuation properties of seismic waves uses the Fourier amplitude spectrum of observed ground motion  $\mathbf{u}(r; f)$ , for a spherical S-wave of frequency  $f$  given as

$$\mathbf{u}^S(r; f) \propto \frac{1}{r} \cdot \exp \left[ -\frac{\pi \cdot f \cdot r}{Q_S \cdot \beta} \right] \quad (9)$$

where  $r$  is distance,  $\beta$  the S-wave velocity, and  $Q_S$  is the attenuation coefficient for S-waves, containing both intrinsic anelastic attenuation and scattering attenuation. For S-waves,  $Q$ -values in the lithosphere range from

$20 \leq Q_s \leq 10\,000$ , for P-waves they are about a factor of two larger [158]. However, the frequency dependence of  $Q_{S,P}$  above about  $f \approx 1$  Hz indicates that attenuation increases for higher frequencies as  $Q_{S,P}^{-1} \propto f^{-n}$  (with  $0.5 \leq n \leq 1.0$ ) [161]. While the effects of seismic wave attenuation due to geometrical spreading are well understood and can be readily measured and incorporated into ground-motion simulation methods. The scattering losses of seismic waves are more difficult to estimate.

Based on seismic array measurements, recorded coda envelopes or sonic logs from borehole measurements, several studies estimated characteristic scale lengths of seismic scattering in the Earth [71,93,152,190]. Alternatively, numerical simulations for assumed random-media realizations have been used to assess scattering properties of the Earth e.g. [65,66]. These works show that the heterogeneity spectrum of velocity fluctuations in the Earth (i.e. the random variations of P- and S-wave velocities around a depth-dependent velocity gradient) can be adequately modeled as a fractal medium (e.g. Eq. (3)) or with a correlation function that does not decay too rapidly in wave-number domain. Smoothly varying Gaussian-type correlation functions are thus inappropriate whereas an exponential or a more general van Karman correlation function Eq. (4) well models seismic scattering in the Earth crust [93,159].

Due the tectonic history of the Earth and the multi-scale nature of geologic structures, scattering parameters depend on region and depth. Correlation lengths of  $a \approx 10$  km are inferred for wave-scattering in the lower crust e.g. [66], whereas [86] assumed  $a \approx 5$  km for upper-crustal wave-field simulations. Upper-mantle heterogeneities may be modeled with correlation length  $a \approx 20$  km. The actual velocity fluctuations are on the order of 2–10%. A concise review of seismic scattering and estimated scattering parameters is given in (see ▶ Seismic Waves in Heterogeneous Earth, Scattering of) [161], a thorough introduction is published by Sato and Fehler [158].

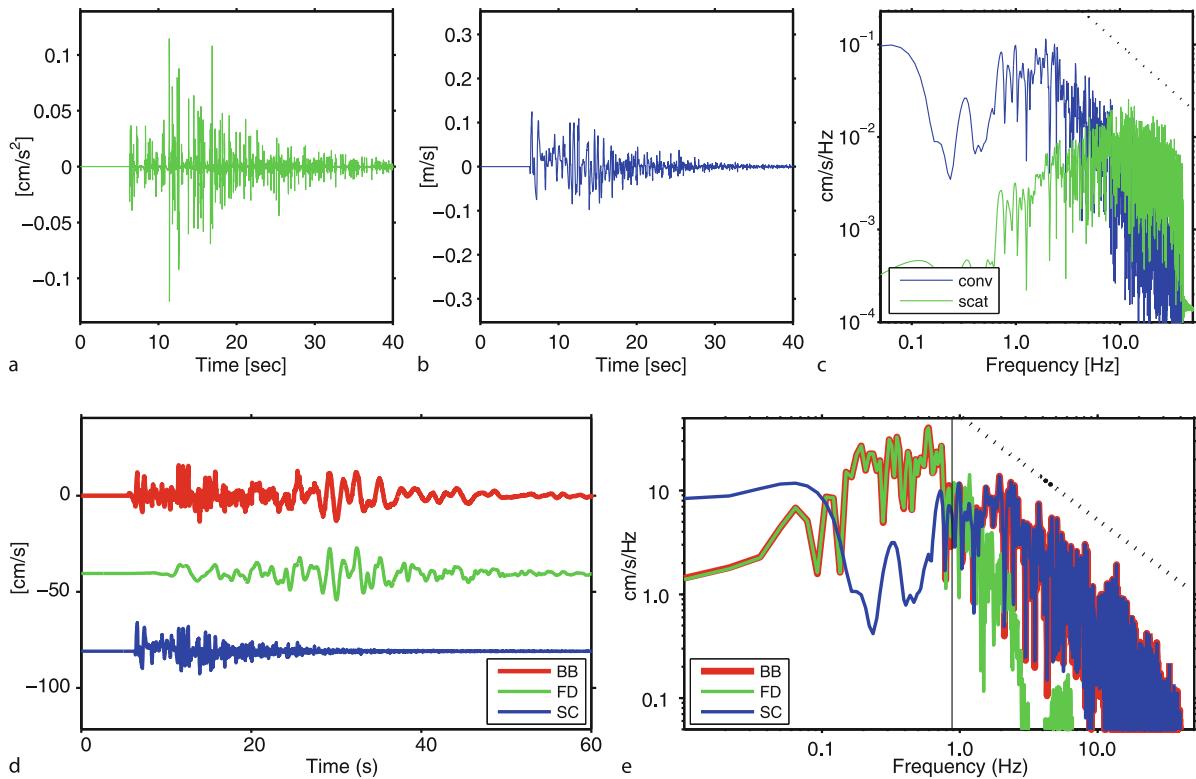
Ground-motion simulations for seismic hazard assessment or studies on the nature of ground-motion complexity should incorporate seismic scattering, i.e. or at least to some degree the stochastic nature of high-frequency seismograms. In earthquake engineering random-vibration theory has been used for this purpose e.g. [35,103], which oversimplifies the earthquake rupture process but appropriately accounts for the apparent randomness of high-frequency ground-motions. More advanced techniques combine deterministic low-frequency motions with stochastic high-frequency signals [99,146,147], in which the stochastic part reflects the short-scale variability in source prop-

erties and seismic-wave scattering. The stochastic signal used in these methods is generated as random white noise and does not contain any physical scattering mechanism.

More physical approaches calculate the 3D-seismic wavefield in a 3D-heterogeneous medium, a computationally expensive task which cannot be carried out for large-scale simulations or many scenario events for which high-frequency ground-motions are needed. To devise a simplified method, [86] have combined realistic small-scale heterogeneity in the source properties with scattering operators for a von-Karman random medium to compute broadband time histories. Their study indicates that the effects of scattering are masked by the heterogeneity in the kinematic source characterization. Note also that, near-field seismic wave scattering in a heterogeneous medium may also replicate apparent nonlinear site effects [130] (see Section “Site Effects and Nonlinearity”).

A computationally less demanding approach to include seismic scattering into ground-motion simulation uses radiative transfer theory to model the space-time distribution of the seismic-energy envelope due to scattered waves [192,194]. In this formulation, a time-dependent multiple S-to-S scattering process occurs due to a shear-dislocation point source embedded in a 3D-medium with background velocity  $v_0$  in which point-like isotropic scatterers of cross section  $\sigma_0$  are randomly distributed with density  $N$ . The total scattering coefficient is thus  $g_0 = N \cdot \sigma_0$ . The detailed theoretical developments are beyond the scope of this article, but it is worth noting that the calculation of the energy-density envelopes is computationally efficient and can be used for high-frequency ground-motion simulations [193].

The multiple S-to-S scattering theory has recently been applied to compute hybrid broadband near-field seismograms [125]. Their technique joins low-frequency 3D-finite-difference (FD) synthetics (which may contain the effects of a sedimentary basin) with site-dependent high-frequency scattering operators to synthesize broadband ground motions. The high-frequency scattering operators are convolved with an appropriate source-time function, and then these site-dependent “scatterograms” are combined in the Fourier domain with the corresponding low-frequency synthetics using a phase-matching optimization technique [124] (conceptually depicted in Fig. 12). Comparing data and simulations for a site that recorded the 1994 Northridge earthquake indicates a good agreement in terms of waveforms, amplitude spectra and spectral acceleration (Fig. 13). The model bias (logarithm of the ratio between the observed and simulated quantity) for spectral acceleration at 30 sites that recorded the 1994 Northridge event (Fig. 14a) shows only small deviations from zero, ex-



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 12**

Conceptual diagram for computing hybrid broadband seismograms. **a** site-specific scattering Greens function for a point-source at the hypocenter. **b** time series representing the local high-frequency scattering contribution (obtained for instance by convolving the scattering Greens function in **a** with an appropriate slip-rate function). **c** Fourier amplitude spectra for the time series in **a** and **b**; the velocity-scatterogram decays as  $1/\omega$  (dotted line) beyond the corner frequency. **d** broadband seismogram (top) computed by combining the LF-seismogram (center) with the site-specific HF-scatterogram (bottom). **e** amplitude spectra for the time series in **d**; the spectra of the broad-band synthetics exactly represent the LF-motions at low-frequencies and the HF-scattering contribution at high frequencies

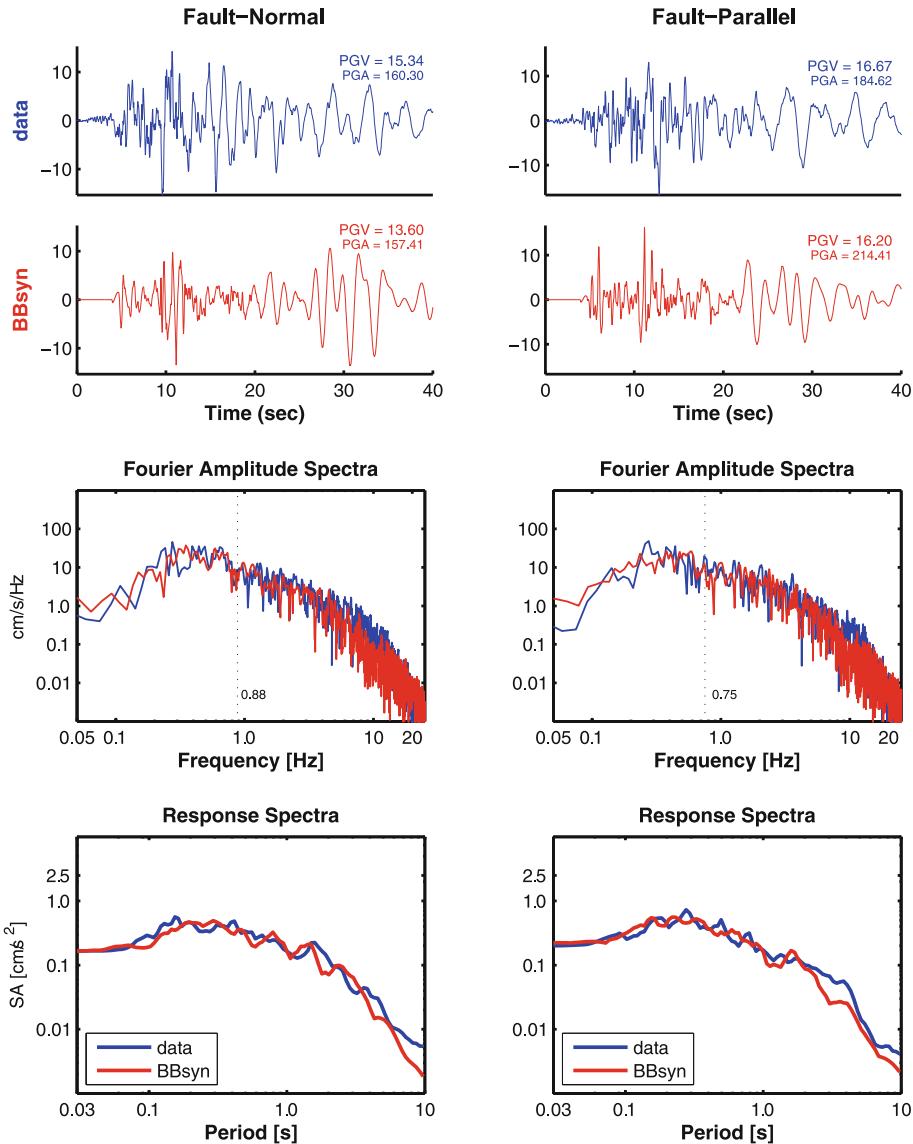
hibiting also narrow 90%-confidence limits. Considering also the general agreement between observed and simulated PGV and PGA values (Fig. 14b) indicates that such hybrid broadband wavefield simulations can reproduce ground-motion complexity, where differences at individual site remain due to unmodeled site and path effects.

Adding seismic scattering into ground-motion characterization and simulation helps to capture the large degree of complexity in the near-source seismic wavefield, and thus presents an important future research topic. Here I have only briefly discussed a few key elements of seismic scattering and its application to a specific case study. However, to fully explore the range of realistic scattering parameters, combined with the inherent complexity of the source-rupture process, requires extensive numerical simulations that need to be calibrated and validated against observational data.

## Site Effects and Nonlinearity

Site effects play a major role in characterizing and quantifying strong ground motion as they may amplify or deamplify the incoming “bedrock” motions in the uppermost velocity layers beneath the observer. Since site-amplification factors may reach two orders of magnitude [108], these effects cannot be neglected in earthquake engineering practice. A comprehensive review of various approaches to estimate and model site effects is given by Kawase [108], so I restrict the following discussion to a few key aspects.

There is no strict boundary between site effects and path effects, but site-effects usually refer to wave-field modifications in the immediate vicinity of the observer location (top 30 m of the local sedimentary cover). Usually, the time-averaged shear-wave velocity in the top 30–100 m



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 13**

Comparison of data (blue) and broadband simulations (red) at a site that recorded the 1994  $M_w$  6.7 Northridge earthquake (left: fault-normal; right: fault parallel motions). Note the consistency between the Fourier amplitude spectra (vertical line denotes the matching frequency) and between the response spectra ( $\zeta = 5\%$  damping). PGA- and PGV-values are also similar

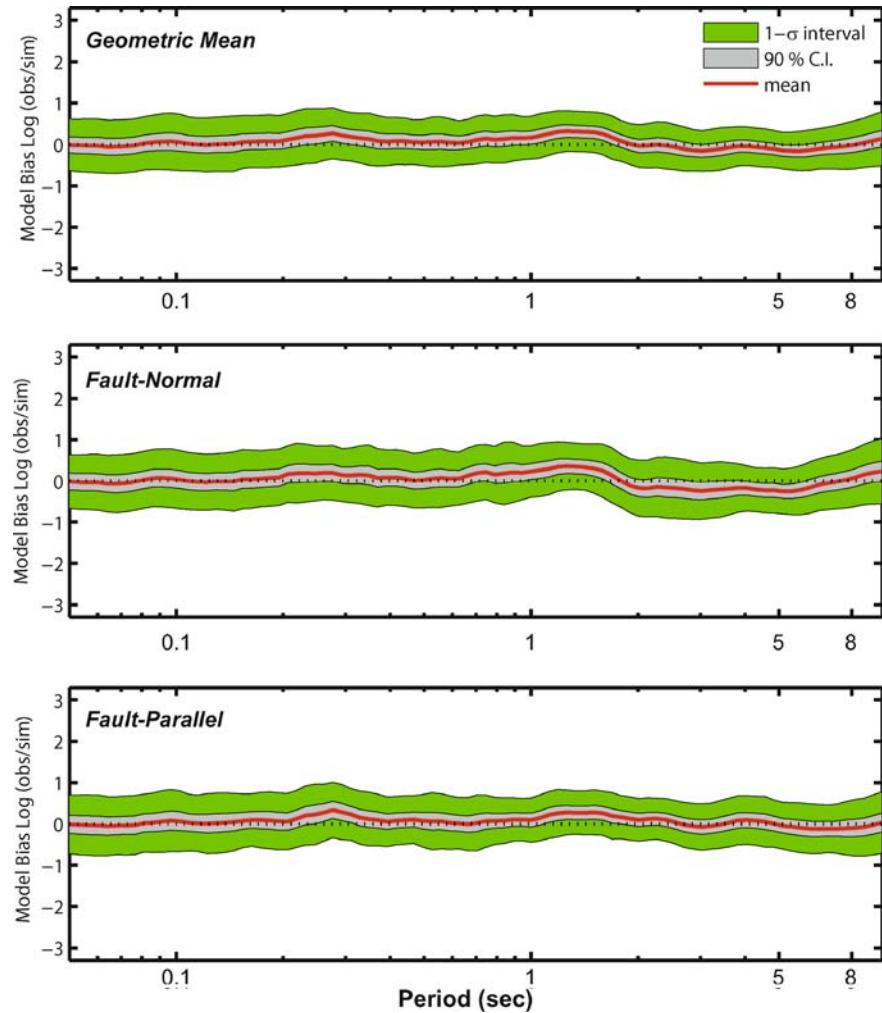
( $V_{S30}$ ) is used to define the station-specific soil classification. Site effects may be affected by the water table, and even soil-building interaction can be considered a site effect.

The classical method to estimate site effects from seismic observations is based on Eq. (1), and attempts to separate source and path from site effects. Expressing Eq. (1) in the Fourier domain and considering the ground motions due to earthquake  $j$  observed at site  $k$ , we obtain source-

site specific Fourier amplitude spectra as

$$A_{jk}(f) = S_j(f) \cdot G_{jk}(f) \cdot L_k(f) . \quad (10)$$

The Green's function term,  $G_{jk}(f)$  can be expressed in terms of geometrical spreading factors with respect to the source-site distances,  $r_{jk}$ ; combining intrinsic and scattering attenuation into a common  $Q(f)$  one obtains (for



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 14a**

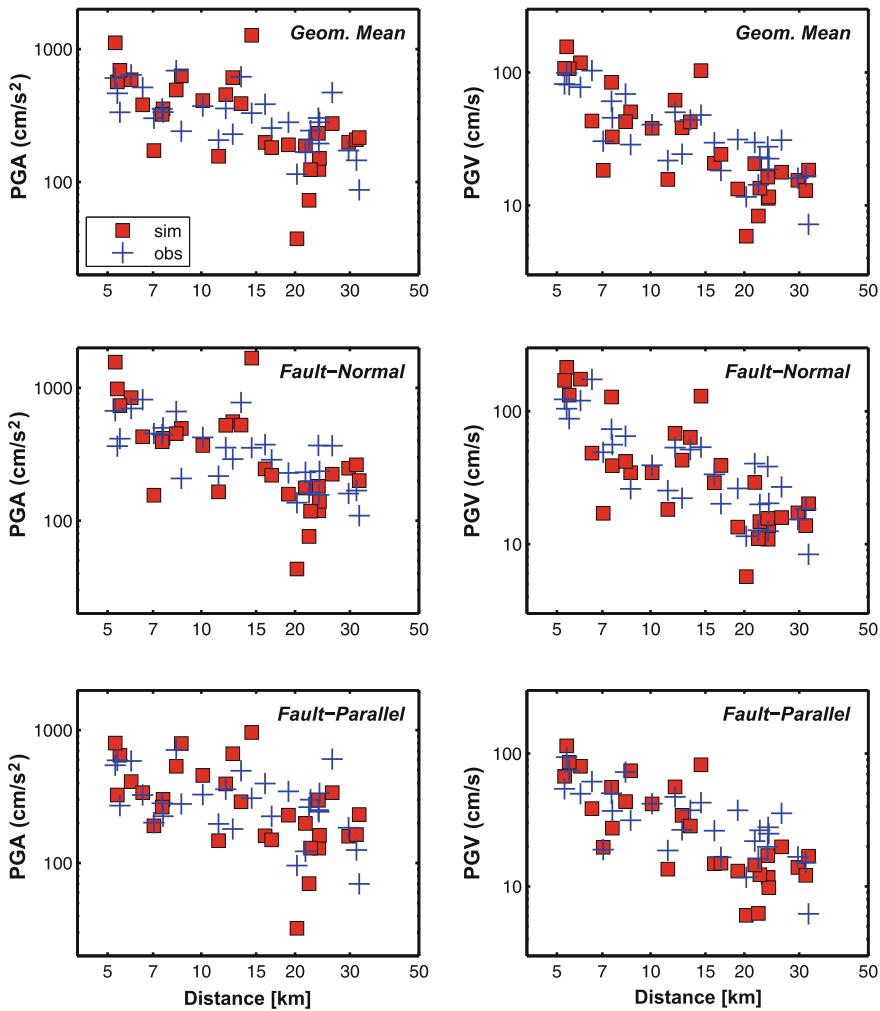
Hybrid broadband ground-motion calculations using low-frequency finite-difference synthetics and high-frequency scattering operators. The three panel show, for the geometric mean (top), the fault-normal (center), and fault-parallel (bottom) component, the model bias, computed as  $\log(\frac{\text{obs}}{\text{sim}})$ , for spectral acceleration at 30 sites that recorded the 1994 Northridge earthquake

S-waves)

$$G_{jk}(f) = \frac{1}{r_{jk}} \cdot \exp \left[ -\frac{\pi \cdot r_{jk}}{\beta \cdot Q(f)} \right] \quad (11)$$

where  $\beta$  is the representative shear-wave velocity over the entire path. Eq. (10) can then be solved by adding at least one independent constraint on either source, site or path. A common approach uses the average site factor as a reference, meaning that the logarithm of the sum of all site factors is equal to zero. Other methods select the site with the smallest site factor as reference, or make an a priori reference-site selection based on geologic or other information.

A recent method for site-effect estimation uses microtremor, ambient noise measurements, or aftershock recordings. Taking the spectral ratio of the horizontal component of motion with respect to the vertical component (so called H/V-ratios), these microtremor ratios often show similar site characteristic as inferred by independent methods [108]. However, the exact physical meaning of this H/V-ratio is not well understood, and since H/V-measurements do not directly represent the true soil amplification, those need to be supplemented and calibrated by numerical modeling [69,178]. Nonetheless, H/V-measurements are useful for microzonation studies which the detailed properties of the uppermost soil layers for a small



**Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 14b**

Comparison of observed (blue crosses) and simulated (red squares) peak-ground acceleration (left) and peak-ground velocity (right) for 30 strong-motion sites that recorded the 1994  $M_w$  6.7 Northridge earthquake. Note the overall agreement between simulated and recorded motions. Discrepancies at individual stations are due to unmodeled ground-motion complexity

area, thus characterizing the susceptibility of particular locations to strong site effects [70].

Site-correction factors have been derived also from the seismic coda. Because coda waves essentially consist of scattered S-waves, local site amplification of the coda itself should be an average of the S-wave amplification factor for S-waves that arrive from different azimuths under different incidence angles. However, the assumption that the seismic coda consists only of scattered S-waves is too simplistic, because surface waves and strong P-coda waves may contaminate the S-wave coda. Site-effect estimation based only on S-coda measurements should be therefore interpreted with care [108].

A common assumption in seismic wave propagation is that strong and weak motions are affected in an identical manner. Analysis of near-field ground-motions provides evidence for nonlinear effects (i.e. a non-linear stress-strain relationship) [26,64,156], due to lower shear-wave velocity and increased damping within the sedimentary cover [54]. Local (or temporal) modification of the seismic properties of the rock material underneath a site or along the wave path may shift resonant modes to lower frequencies and generate reduced amplitudes. Nonlinearity occurs once a certain ground-motion intensity threshold is exceeded. Previous studies suggest  $PGA > 0.3$  g, or  $PGV > 20$  cm/sec, or peak-strains in excess of 0.06% [9],

roughly consistent with values inferred for strong-motion recordings of the 1989  $M_w$  6.9 Loma Prieta earthquake for which the threshold for large S-wave travel-time delays in repeating earthquakes occurs at  $PGA > 0.4 \text{ g}$  or  $PGV > 40 \text{ cm/sec}$  [156].

Effects of nonlinearity greatly increases the complexity of ground-motions, but only on a very localized scale. Recently, detailed spectrogram analyses of seismic recordings of the 2003  $M_w$  8.3 Tokachi-Oki (Japan) earthquake demonstrate liquefaction, quantified by a dramatically reduced high-frequency contents of the waveforms [55]. However, it is very difficult to establish general scaling laws that account for potential nonlinearity effects. In case of water-saturated loose sands, strong shaking may result in nonlinear soil liquefaction due to a rapid, temporary increase of pore water pressure, resulting in a dramatic loss of soil stiffness. Liquefaction effects are extremely important for earthquake engineering, but have received little attention from seismology in the past.

### Ground-Motion Scaling Relations

Despite the fact that observed near-source ground-motions show complicated time histories and large variability in intensity measures, there is the need in earthquake engineering and seismic-hazard analysis to devise simple approaches for estimating expected ground-motions in future earthquakes. For this reason, ground-motion prediction equations (GMPE's) are developed, providing mathematical expressions that relate shaking intensities to seismological quantities, source-site geometry, and potentially site-specific parameters. Thus, the complexity of the earthquake rupture process, the wave-propagation effects from the source to the site, and the detailed site conditions are condensed into relative simple functional forms, comprising a few parameters that constitute empirical ground-motion scaling relations.

#### Development of Empirical Scaling Relations Based on Recordings of Past Events

Using strong-motion observations, ground-motion prediction equations are developed using a number of different parameterizations. The most fundamental form is given as [52]

$$Y = a_1 \cdot e^{a_2 M} \cdot R^{a_3} \cdot e^{a_4 r} \cdot e^{a_5 F} \cdot e^{a_6 S} \cdot e^{\sigma} \quad (12)$$

where  $Y$  is the ground-motion intensity measure of interest ( $PGA$ ,  $PGV$ ,  $S_A$ , ...).  $M$  is the earthquake's magnitude,  $r$  the source-to-site distance;  $F$  is a parameter that char-

acterizes the type of faulting, and  $S$  captures the local site conditions.  $R$  is an additional magnitude-dependent distance function which can take on alternative forms, for example [52]

$$R = \begin{cases} r + c_7 \cdot e^{c_8 M} \\ \sqrt{r^2 + [c_7 + e^{c_8 M}]^2} \end{cases}. \quad (13)$$

This model contains some of the basic physics of earthquakes (terms with  $M$  and  $F$ ), the attenuation due to geometrical spreading of seismic waves (terms related to distance  $r$  and  $R$ ) and the site conditions ( $S$ ). Due to the approximately log-normal distribution of ground-motion intensities, attenuation relations are parameterized using the natural logarithm of  $Y$ . Correspondingly, the standard deviation of the zero-mean random error term  $\sigma$  is estimated as the standard error of  $\ln Y$  ( $\sigma_{\ln Y}$ ).

Over the years, different models have been proposed to capture various aspects of the inherent complexity of near-fault shaking. One class of models assumes a functional form whose shape is magnitude-independent for all distances. A simple parametric form is given by [38]

$$\begin{aligned} \ln(Y)_{M,R,F} &= a_1 + a_2 \cdot M + a_3 \cdot M^2 + a_4 \cdot \ln(r + a_5) + a_6 \cdot F. \\ &\quad (14) \end{aligned}$$

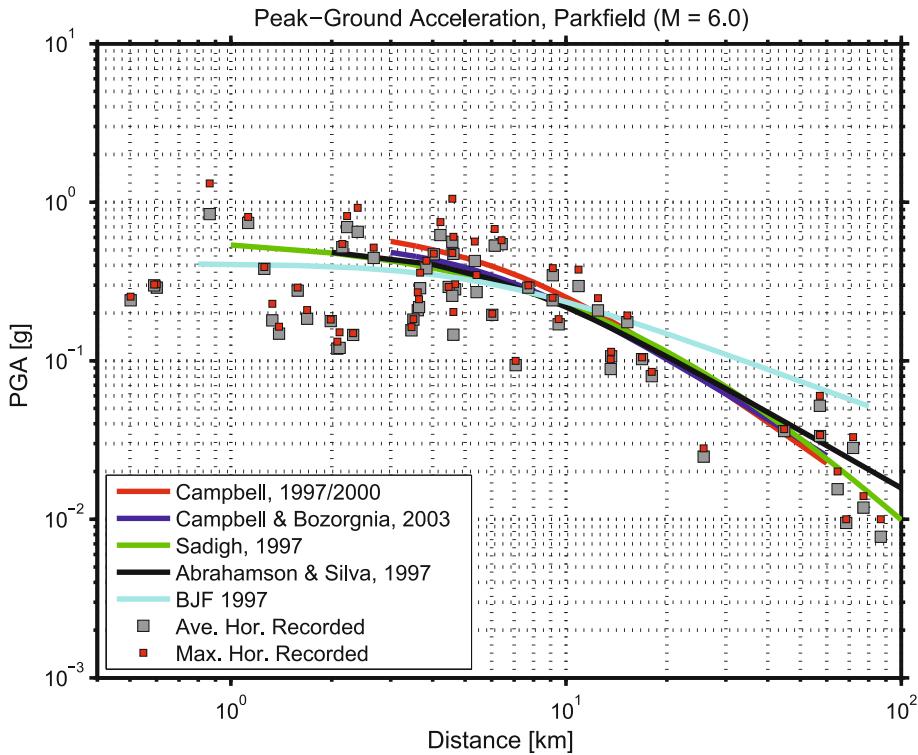
However, many observations show that ground-motion intensities saturate at close distances to the fault, i. e. moderate-magnitude earthquakes ( $5.0 \leq M \leq 6.5$ ) may generate about the same level of high-frequency shaking as large magnitude events ( $M > 6.5$ ). The second class of ground-motion scaling relations captures these observations. For instance, starting from Eq. (14) Abrahamson and Silva [5] propose the relation

$$\begin{aligned} \ln(Y)_{M,R,F} &= a_1 + a_2 \cdot M + a_3 \cdot M^2 \\ &\quad + (a_4 + f_1(M)) \cdot \ln(r + a_5) + a_6 \cdot F \quad (15) \end{aligned}$$

with  $f_1(M) = a_7 M$ , while Sadigh et al. [157] developed an attenuation model given by

$$\begin{aligned} \ln(Y)_{M,R,F} &= a_1 + a_2 \cdot M + a_3 \cdot M^2 + a_4 \cdot \ln(r + f_2(M)) + a_6 \cdot F \\ &\quad (16) \end{aligned}$$

where  $f_2(M) = c_7 \cdot e^{c_8 M}$ . The difference between these two scaling relations becomes evident at larger distances ( $r > 50 \text{ km}$ ), while for short distances they lead to about the same ground-motion intensities (Fig. 15).



#### Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 15

Observed peak ground accelerations for the Sept 28, 2004  $M_w$  6.0 Parkfield earthquake, compared against five different empirical attenuation relationships. Data are plotted using the geometrical mean of the two horizontal components (fault-normal and fault-parallel) ( $PGA = \sqrt{PGA_{FN} \cdot PGA_{FP}}$ ), and for the maximum of the two components. The median of the observations is generally in good agreement with the empirical predictions, but the data exhibit large variability with much higher (and lower) motions than empirically predicted, independent of the closest distance to the fault. Note that the recording at station "Fault Zone 16" is not shown as it was not available at the COSMOS database. The PGA-value for site "Fault Zone 16", located at  $\sim 0.6$  km distance from the fault, is estimated to be in the range 2.0–2.5 g [167]

Empirical ground-motion prediction like those presented in Eqs. (14), (15), (16) have been widely applied to global and regional ground-motion data sets, but a number of observations call for modification to these "generic" models. First of all, attenuation models for different tectonic provinces have been published accounting for variations in the regional geology [7,8,17,18,31,67,175]. While there is the need to distinguish between stable-cratonic areas (like Eastern North America), extensional regimes (like the Basin and Range province in the western US), zones of complex active tectonics (like California), or subduction zones (like in Mexico, Chile, Japan), recent work [7] suggests that data from different tectonically active areas with crustal seismicity can still be jointly analyzed. In certain regions the geologic conditions cause the presence of several of such tectonic regimes over a rather small area (e.g. Italy, Switzerland, Japan) which further complicates the development of adequate ground-motion prediction equations.

Besides the local geology, the source-site geometry greatly influences seismic shaking. In the near-fault regime, ground-motions are strongly affected by geometrical source-site effects, leading to large long-period pulses of motion termed "directivity pulses" (see Sect. "Isochrone Theory"). Recent work includes such directivity pulses in the attenuation equation [170,171]. Large permanent offset of the ground due to the rupture process will also generate long-period large motions; this so-called "fling-step" is distinguished from the directivity pulse [32]. For thrust-faulting earthquakes, the "hanging-wall" effect has been shown to generate very different ground motions [3] for sites located on the hanging-wall than on the footwall (Fig. 3). This effect is explicitly included in several recent attenuation models [3,53].

An intriguing observation has been made following the large surface-breaking earthquakes in 1999 in Turkey (the Aug. 17  $M_w$  7.5 Izmit and the Nov. 12  $M_w$  7.1 Duzce events) and Taiwan (the Sept. 20  $M_w$  7.6 Chi-Chi event):

near-field ground-motions were significantly lower than empirically predicted [105]. This counter-intuitive result can partly be explained by the general differences in stress-release patterns between buried-faulting and surface-rupturing earthquakes. The primary cause for the ground-motion variation between these two classes of earthquakes lies in general differences in the dynamic rupture process and the associated energy balance of the system (see ▶ [Earthquake Scaling Laws](#)). Ultimately, these effects could be related to depth-dependent fault-zone properties which exhibits highly damaged rocks with intense microcracking and lower seismic velocity close to the surface, but more compliant and less damaged material at depths [22]. Fault-zone structure affect the dynamics of the rupture process [13,21,22,83] and the resulting seismic radiation [141], and therefore strongly influence the resulting ground-motion intensities.

Besides the functional form and the given geological, geometrical, and physical parameters a number of technical issues affect ground-motion scaling equations (a comprehensive review is given in [60]): (i) the choice of the ground-motion intensity measure  $Y$  determines the values for the coefficients  $a_i$  (Eqs. (12)–(16)) derived by regression analysis; (ii) the specific approach for the regression (data fitting) matters (common techniques are weighted non-linear least-squares regression, two-step regression, or random-effects regression); (iii) the data selection criteria for assembling the individual strong-motion database; (iv) data corrections may be applied to the recorded strong-motions (e.g. uniform instrument response for all records, baseline correction, frequency filtering) and also to the meta-data (magnitudes, distances, site information etc.) which represent the independent parameters.

In developing ground-motion prediction equations, a uniform and well-calibrated magnitude definition is particularly important, but in practice often difficult to achieve when merging datasets from various institutions whose magnitude values may not be compatible. The style-of-faulting factor  $F$  in Eqs. (14), (15), (16) is generally well defined, but the site-effect factor  $S$  is often poorly known. One of the most critical parameters in ground-motion scaling is the distance  $r$  between the site and the source. In particular for extended rupture planes, the applied distance metric for the source-to-site geometry becomes crucial; different source-to-site distance definitions are in use [4] but the particular choice for  $r$  in turn will affect the seismic-hazard calculations [163].

In an effort to harmonize and calibrate ground-motion data, related meta-data, and the development of empirical ground-motion prediction equations, the Pacific

Earthquake Engineering Research Center (PEER) carried out the New Generation Attenuation of Ground Motions (NGA) Project (completed in Jan. 2008). In this context, [37] for instance refine Eq. (14) to accommodate the effects of anelastic attenuation when modeling far-distance recordings ( $R > 80$  km) and to include an “effective” magnitude-dependent geometrical spreading (allowing to predict ground-motion amplitudes out to distances  $R = 400$  km). Their data-driven equation includes only terms that are truly needed to adequately fit the data, involving (i) a complicated magnitude-scaling function,  $f_M(M)$ ; (ii) a versatile distance function,  $f_D(R_{JB}, M)$ ; (iii) a site-effect function that includes potential nonlinearity effects,  $f_S(V_{S30}, R_{JB}, M)$ :

$$\begin{aligned} \ln(Y)_{M,R,F} = & f_M(M) + f_D(R_{JB}, M) \\ & + f_S(V_{S30}, R_{JB}, M) + \epsilon \cdot \sigma_T . \end{aligned} \quad (17)$$

All terms in Eq. (17) are period-dependent;  $M$  is moment magnitude,  $R_{JB}$  is the Joyner–Boore distance (the closest distance to the surface projection of the fault)  $\epsilon$  is the fractional number of standard deviations of a single predicted value of  $\ln(Y)$ , and  $\sigma_T = \sqrt{\sigma_{\text{intra}}^2 + \sigma_{\text{inter}}^2}$  describes the uncertainty by combining the intra-event and inter-event aleatory uncertainty. In this context, the exact definition for the ground-motion parameter  $Y$ , usually taken as the geometric mean of the two horizontal components, has received increased attention. Because of misaligned recording instruments or complicated (multi-)pathing of radiated waves, simple measures of “mean ground-motion” appeared to be incorrect when taking the sensor orientation as installed in the field. Works by [30,39,150] discuss the usefulness of orientation-independent ground-motion measures based on data and numerical simulations.

### Relating Earthquake Source-Scaling to Ground-Motion Prediction

When using empirical ground-motion attenuation relations for predicting the expected shaking level at a given site the analyst needs quantitative information regarding the type of faults and their properties, the fault locations with respect to the site, and the site conditions. Probabilistic seismic hazard analysis (PSHA) requires additional information about the recurrence rate of earthquakes on the chosen faults [111]. Therefore, earthquake statistics (see ▶ [Geo-complexity and Earthquake Prediction](#)) [111] and earthquake physics (see ▶ [Earthquake Scaling Laws](#)) [6] are major components for accurate seismic hazard calculations. However, ground-motion prediction Eqs. (14),

(15), (16) only contain a faulting-style factor ( $F$ ) and the magnitude dependence as earthquake-source related parameters. A correct estimation of the potential magnitudes for earthquakes occurring on the selected faults is the most crucial step in any seismic hazard study. Earthquake physics (see ▶ [Earthquake Scaling Laws](#)) [6] provides the theoretical foundation for defining these magnitudes and to explain observational data and empirical relations e. g. [106,122,187].

In modern seismic hazard analysis, the magnitude  $M$  relates to moment magnitude,  $M_w$ , derived from the seismic moment as  $M_w = 2/3 \cdot \log M_0 - 6.07$ . Since  $M_0 = \mu \cdot L \cdot W \cdot D$ , source-scaling relations between observable fault dimensions (length  $L$ , width  $W$ , area  $A = L \cdot W$ , average displacement  $D$ ) and magnitude are used in empirical ground-motion prediction for obtaining self-consistent input parameters.

Based on geologic observations it is often possible to estimate the length of a fault, considering also cases with multiple segments that may rupture in individual events or jointly (these cases are then treated in PSHA using logic trees) [188]. Estimating fault width is more difficult; it may be inferred from the location of background seismicity in a crustal tectonic setting, or from modeling interseismic deformation which constrains the fault's locking depth, but often a generic fault width of  $W = [10 - 20]$  km is assumed. Several studies have published earthquake scaling relations between various fault parameters [79,120,122,169,186,187]; these source-scaling relations are not only useful for seismic hazard analysis but also provide important insight into earthquake mechanics (see ▶ [Earthquake Scaling Laws](#)) [164]. A generic relation between magnitude  $M_w$  and fault area  $A$  (in  $\text{km}^2$ ) is given by:

$$M_w = p + q \cdot \log_{10} A \quad (18)$$

For  $q \equiv 1$ , Eq. (18) is consistent with self-similar constant average-stress-drop scaling (see ▶ [Earthquake Scaling Laws](#)). Rule-of-thumb values for the coefficients in Eq. (18) are  $p = 4$  and  $q = 1$  (adapted from [187] who find  $p = 3.98$ ,  $q = 1.02$  for strike-slip earthquakes, and  $p = 4.07$ ,  $q = 0.98$  for all faulting styles), leading to a magnitude  $M_w 7$  earthquake for a fault area  $A = 1000 \text{ km}^2$ . However, a number of studies have found considerable deviations from self-similar earthquake scaling [79,122,169], and reported values in the range  $3.97 \leq p \leq 4.39$  and  $0.97 \leq q \leq 1.33$  [120]; these differences strongly affect empirically predicted ground-motion intensities and seismic hazard calculations [188,189]. For instance, [188] proposes  $p = 4.2$ ,  $q = 1$ , resulting in an  $M_w 7.2$  earth-

quake for  $A = 1000 \text{ km}^2$ , with twice the seismic moment and twice the displacement as an  $M_w 7.0$  rupture with identical source area. Correspondingly, Eqs. (14)–(17) lead to higher ground-motion estimates for the larger event, consequently also different seismic hazard, due to small changes in the source-scaling relation.

## Quantifying Uncertainty in Ground-Motion Prediction Equations

Analyzing, quantifying and modeling strong-motion uncertainties is a mandatory part of any reliable seismic-hazard study [168]. The variability in source-scaling relations reflects in part the complexity of the earthquake rupture process, treated as an epistemic uncertainty (i. e. as more data become available and seismologists better understand earthquake dynamics, this variability will eventually decrease). The standard deviation of the residuals, obtained when deriving ground-motion prediction equations by regression analysis measures the aleatory variability (randomness) of ground-motion parameters. This standard deviation,  $\sigma_{lnY}$ , is then partitioned into two error terms for the intra-event and inter-event variability,

$$\sigma_{lnY} = \sqrt{\sigma_{\text{inter}}^2 + \sigma_{\text{intra}}^2} .$$

If the regression uses the geometric mean of the two horizontal components and the component-to-component variability is needed, a third term  $\sigma_{\text{comp}}$  has to be included. The intra-event variability can be further separated into a site-to-site component  $\sigma_{\text{comp}}$ , and the remaining variability  $\sigma_0$  (after accounting for source and site effects):  $\sigma_{\text{intra}} = \sqrt{\sigma_s^2 + \sigma_0^2}$  [52].

The standard deviation  $\sigma_{lnY}$  has previously been found to be a function of magnitude (e. g. [5,157]) with decreasing variability for increasing magnitudes. However, this counter-intuitive result could be affected by a sampling bias due to fewer near-source strong-motion recordings for moderate-to-large earthquakes. More recent work suggests that  $\sigma_{lnY}$  is independent of magnitude [2]. The exact value of  $\sigma_{lnY}$  varies between different studies and ground-motion parameters, but is generally in the range  $0.4 \leq \sigma_{lnY} \leq 0.8$ . Including this standard deviation in the ground-motion estimation is indispensable to assess the variability in shaking intensity – common practice is to use at least one or two standard deviations. However, for critical infrastructures (e. g. nuclear power plants, nuclear waste repositories), it is not clear where to truncate the ground-motion distribution (i. e. how many  $\sigma$ 's to include, [177]) since very long return periods need to be

considered and conclusive physical arguments for upper limits on near-field ground motions have not yet been made. Bommer et al. [33] demonstrate that the choice of the truncation level will significantly affect seismic hazard estimates at very low probabilities (annual frequencies of exceedance of  $10^{-6}$  and lower), while the effect on hazard at probabilities traditionally used in PSHA (annual frequencies of exceedance of  $10^{-4}$ ) is small. The reason is that low-probability seismic hazard may contain contributions from rare but extreme ground-motion values.

The preceding discussion raises the issue of defining upper bounds of near-source ground motion [33]. For instance, the maximum values for *PGA* in the NGA-database is  $PGA = 1.56$  g (for the Tarzana record of the 1994  $M_w$  6.7 Northridge earthquake), but values of over 2 g have been reported for the 2004 Parkfield earthquake [167,168] or the 2003 Miyagi (Japan) earthquake. Note that sites and events showing the largest *PGA*-values generally do not coincide with those showing the highest *PGV*-values. For instance, the largest *PGV*-value reported in the NGA-database is 205 cm/s (observed at station TCU068 for the  $M_w$  7.6 Chi-Chi earthquake, Fig. 4), but the *PGA* at this site was only 0.53 g. *PGV*-values in the NGA-database frequently exceed 75 cm/s, and values of 100 cm/s are not uncommon. Closer inspection of the rapidly growing online databank of near-source recordings, obtained with modern dense observational networks (e.g. the K-Net and KiK-net stations in Japan) may return even higher maximum *PGA*- and *PGV*-values than listed above.

Current research addresses the physical limits to maximum ground motions, imposed by the complexity of the source-rupture process, the wave propagation, the site conditions and the strength of the rock. Investigating the combination of these complex physical processes with numerical methods and large-scale simulations helps to better understand the overall distribution of ground-motion parameters and their complexity.

An innovative observational method to examine maximum ground-motion has been developed by Brune and colleagues [19,46,47,48,49,50], based on precariously balanced rocks. These are free-standing (large) boulders, created by erosional processes, which appear as if small ground-accelerations could overturn (topple) them; this toppling-acceleration can be measured. By also dating the age of these rocks it is possible to determine which ground-motion levels have not been exceeded in the corresponding time interval in this region. These “natural seismoscopes” therefore provide important constraints for maximum shaking levels and probabilistic seismic hazard analysis.

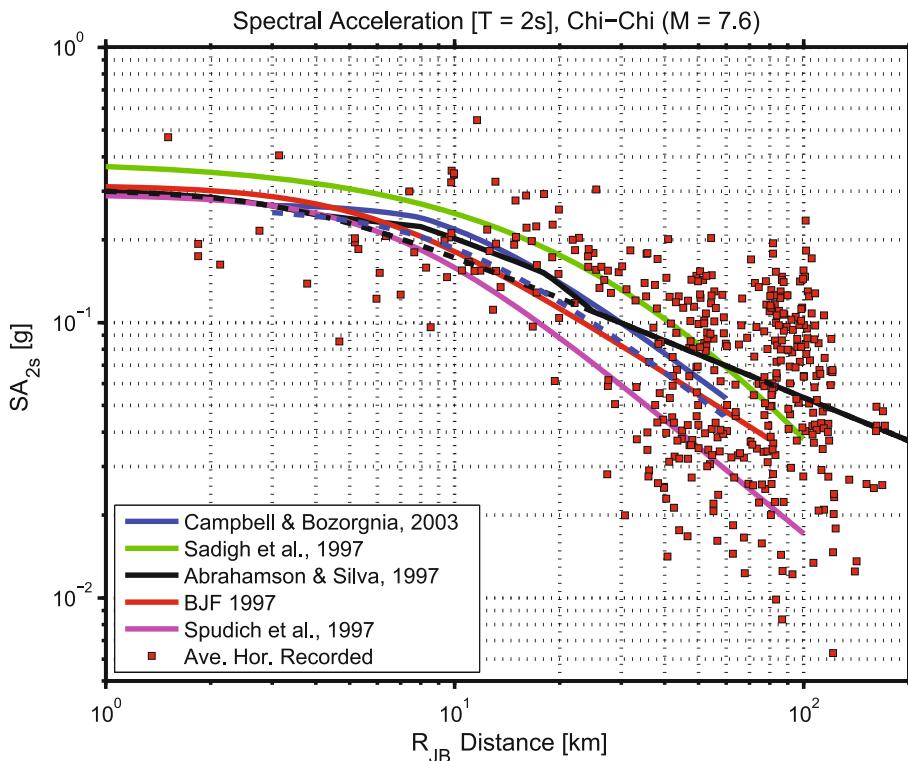
## Comparing Observations with Empirical Predictions

The introduction of this article presents near-field ground-motion observations for two recent well recorded earthquakes, the  $M_w$  6.0 strike-slip earthquake in Parkfield (09/28/2004) and the  $M_w$  7.6 thrust-faulting earthquake in Taiwan (09/20/1999) (Figs. 1–4). I now compare these records against empirical predictions for a number of widely used ground-motion attenuation relations. Figures 15 and 16 display recorded *PGA* for the Parkfield event and spectral acceleration ( $S_A^{T=2\text{ sec}}$ ) for the Chi-Chi earthquake, respectively. In both cases, the empirical predictions are generally consistent with the median of the observations, but the large scatter in the data also leads to observations occur far outside the standard  $1\sigma$ -bounds of the predictions.

For the Parkfield data, observed *PGA*-values are both higher and lower than empirically predicted (for  $r < 100$  km), and no obvious pattern is visible that may explain the large ground-motion variability (Fig. 15). Peak ground acceleration close to and above 1 g were observed at a number of stations. At one site (station “Fault Zone 16”), the *PGA*-value is estimated to be on the order of 2.0–2.5 g [167]. The plotted *PGA*-values are not corrected or grouped according to the soil classification for each site, i.e. parts of the observed variability can be attributed to site effects. However, much of the variability originates from the particular position of each station with respect to the spatial complexity of the earthquake rupture process, and, to a lesser extent, from localized wave-propagation phenomena.

Examining recorded motions and empirical predictions for the Chi-Chi earthquake reveals an overall consistency between the observed median spectral acceleration ( $S_A^{T=2\text{ sec}}$ ) and ground-motion attenuation relations, despite significant scatter in the data (Fig. 16). At large distances ( $R_{JB} \geq 50$  km) many of the recorded motions are significantly larger than any of the empirical prediction, while at distances close to the fault ( $R_{JB} \leq 10$  km) most of the  $S_A$ -values are significantly lower than predicted. This unexpected behavior of lower than-predicted ground-motions for large surface-breaking earthquakes has been recently observed for a number of earthquakes [105], and is attributed to effects of earthquake source dynamics occurring in the uppermost shallow and more compliant crustal layers [126].

These two examples illustrate the large variability of near-source motions recorded at a number of stations for the same earthquake, thus characterizing the intra-event variability described above. Site, path, and source effects



#### Ground Motion: Complexity and Scaling in the Near Field of Earthquake Ruptures, Figure 16

Spectral accelerations ( $SA$ ) at  $T = 2\text{ s}$  for ground-motion recordings of the Sept 20, 2004  $M_w 7.6$  Chi-Chi (Taiwan) earthquake, compared against five different empirical attenuation relationships. Data are plotted using the geometrical mean of the two horizontal components (fault-normal and fault-parallel) ( $SA = \sqrt{SA_{FN} \cdot SA_{FP}}$ ). The Campbell & Bozorgnia [53] and Abrahamson & Silva [5] relations contain a hanging-wall factor, indicated by the *solid blue* and *black lines*, respectively; the corresponding regular relations are plotted with *dashed lines*. The median of the observations is consistent with the empirical predictions, but the data exhibit large variability, with much higher (and lower) motions than empirically predicted. Note that at large distances ( $R_{JB} \geq 50\text{ km}$ ) many sites reveal higher-than-predicted spectral accelerations while at very short distances ( $R_{JB} \leq 10\text{ km}$ ) many sites exhibit significantly lower ground-motions.

are responsible for this large degree of ground-motion complexity, but separating the contributions of each of these effects is difficult. Deciphering the detailed physical processes and corresponding parameters that lead to a specific near-field ground-motion is an active area of research. Ultimately, this will allow us to better model ground-motions and their variability for future earthquakes, a key ingredient for improved seismic hazard analysis.

#### Future Directions

This article reviews the complexity of near-field ground motions, generated by the space-time-dependent heterogeneous earthquake rupture process, transformed by wave-propagation through complex geologic structure and inhomogeneous media, and finally subjected to localized site conditions. As a consequence, the waveform com-

plexity recorded at dense strong-motion arrays provides both a challenge and an opportunity for future research.

A largely overlooked aspect of ground-motion complexity pertains to the rotational motions (also called vorticity) of the displacement field. Translational ground displacements (velocity, acceleration) are recorded as seismograms used for monitoring seismic activity and ground motions; strain measurements capture the deformation of the Earth, but the theoretically predicted [6] ground rotation (a vectorial quantity) has rarely been reliably measured in the past. Observations of rotational motions are challenging, due to their small amplitudes [42] and inadequate instrument sensitivity, but recent work provides evidence that rotational motions are significant [95,97,98,179]. Moreover, they are particularly important for engineering applications [182,183] (see ▶ **Earthquake Source: Asymmetry and Rotation Effects** for

a comprehensive review on rotational motions). As improved sensors are developed to measure rotational motions, they may become a new observable for (engineering) seismology, potentially adding complementary information on earthquake source processes, Earth structure, and ground shaking [98]. For deciphering ground-motion complexity, rotational seismology constitutes a new emerging research field; consequently, the International Working Group on Rotational Seismology (IWGoRS) was formed in 2006 to foster the exchange of ideas, data, and software on rotational seismology.

As strong-motion arrays become more abundant and better equipped, the recorded near-source seismograms provide a wealth of information on the rupture process, the wave-propagation phenomena and the site structure. Harvesting these data to learn more about these complex physical processes will be a research focus for years to come. For example, current earthquake source inversions generally use low-pass filtered seismograms ( $f \leq 1$  Hz) because of incomplete knowledge of Earth structure at shorter wavelengths and the increasing ill-conditioning of the inverse problem as higher frequencies are included. This limitation does not allow proper imaging of high-frequency radiation on the fault, excited at small-scale geometrical complexities or by sudden changes in the dynamic rupture process. Those higher frequencies, however, are particularly damaging to much of the built environment and critical for reliable seismic hazard studies.

Currently, seismologists either estimate earthquake source properties and assume an Earth model, or simplify the source to a known mechanism to understand Earth structure. Future work will increasingly consider earthquake source modeling and imaging of Earth structure as a coupled (joint) inverse problem, potentially including shorter wavelength and/or deriving appropriate stochastic media characterizations. With the advent of high-performance computing facilities and innovative numerical methods (see ▶ [Seismic Wave Propagation in Media with Complex Geometries, Simulation of](#)) [61,112,149], multi-level optimization strategies will be developed to solve the highly non-linear inverse problem of inferring earthquake rupture dynamics from radiated seismic waves that propagate through heterogeneous media.

Empirical ground-motion equations remain important in engineering practice, but accurately quantifying ground-shaking variability for GMPE's continues to be difficult, despite an increasing number of near-source recordings. It is likely that each future large earthquake, recorded by a dense network, may generate unusual ground-motions outside the commonly assumed standard-deviation of current empirical predictions. Pre-

cise and reliable ground-motion estimation is the key challenge for seismologists and earthquake engineers for proper seismic hazard assessment and earthquake loss mitigation in future events. This task requires the ability to model not only median values of ground-motion intensity, but also to accurately capture their variability. Moreover, it will become increasingly important to not only model scalar ground-motion intensities, but to compute large suites of realistic synthetic near-field seismograms that reproduce the observed ground-motion complexity. Thus, innovative source-modeling approaches are required to capture the complexity of the dynamic rupture process (e.g. [77,120]). Multi-scale dynamic simulations with initial conditions based on stochastic parameter distributions [133,151], need to be coupled to large-scale broadband wave-propagation computations (e.g. [138,139]) for many realizations of heterogeneous Earth models (perhaps with stochastic properties at short wave-lengths). Such simulation-based ground-motion prediction is needed to advance the current practice in seismic hazard assessment.

## Acknowledgments

I am indebted to J. Bühler for generating several figures for this article. J. Ripperger provided simulation data for Fig. 5. Thanks to P. Spudich for computing isochrone quantities and synthetics shown in Fig. 11. Strong-motion data were taken from the COSMOS strong-motion database (<http://db.cosmos-eq.org>). Critical comments and helpful suggestions by J. Clinton, G. Cua, and S. Jonsson greatly improved the article. I am also grateful for constructive reviews by R. Harris and Y. Ben-Zion. Parts of this work was supported by the Southern California Earthquake Center; SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. This is SCEC contribution number 1154.

## Bibliography

### Primary Literature

1. Aagaard B, Heaton TH (2004) Near-source ground motions from simulations of sustained intersonic and supersonic fault ruptures. *Bull Seis Soc Am* 94(6):2064–2078
2. Abrahamson NA, Silva WJ (2005) Preliminary results of the A&S 2005 attenuation relation, presented at the USGS workshop on Attenuation Relations to be used in the National USGS Seismic Hazard Maps, Menlo Park CA, October 24
3. Abrahamson NA, Somerville PG (1996) Effects of the hanging-wall and foot-wall on ground motions recorded during the Northridge earthquake. *Bull Seis Soc Am* 86:93–99

4. Abrahamson NA, Shedlock KM (1997) Overview. *Seis Res Lett* 68(1):9–23
5. Abrahamson NA, Silva W (1997) Empirical response spectral attenuation relations for shallow crustal earthquakes. *Seis Res Lett* 68(1):94–127
6. Aki K, Richards PG (2002) Quantitative Seismology. University Science Books, Sausalito
7. Akkar S, Bommer JJ (2007) Empirical prediction equations for peak ground velocity derived from strong-motion records from Europe and the Middle East. *Bull Seis Soc Am* 97(2):511–530
8. Ambraseys NN (1995) The prediction of earthquake peak ground acceleration in Europe. *Earthq Eng Struct Dyn* 24(4):467–490
9. Anderson JG (2002) Strong-motion seismology. In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) International Handbook of Earthquake and Engineering Seismology, Part B. Academic Press, San Diego, 937–965
10. Andrews DJ (1976) Rupture propagation with finite stress in antiplane strain. *J Geophys Res* 81(20):3575–3582
11. Andrews DJ (1980) A stochastic fault model; 1. Static case. *J Geophys Res* 85(B7):3867–3877
12. Andrews DJ (1981) A stochastic fault model; 2. Time-dependent case. *J Geophys Res* 86(11):10821–10834
13. Andrews DJ, Ben-Zion Y (1997) Wrinkle-like slip pulse on a fault between different materials. *J Geophys Res* 102(1): 552–571
14. Aochi H, Fukuyama E (2002) Three-dimensional non-planar simulation of the 1992 Landers earthquake. *J Geophys Res* 107(B2):art 2035. doi:[10.1029/2000JB000061](https://doi.org/10.1029/2000JB000061)
15. Aochi H, Madariaga R (2003) The 1999 Izmit, Turkey, earthquake: Non-planar fault structure, dynamic rupture process, and strong ground motion. *Bull Seis Soc Am* 93(3):1249–1266
16. Archuleta RJ (1984) A faulting model for the 1979 Imperial Valley earthquake. *J Geophys Res* 89(6):4559–4585
17. Atkinson G, Boore DM (1995) New ground motion relations for eastern North America. *Bull Seis Soc Am* 85:17–30
18. Bay F, Fäh D, Malagnini L, Giardini D (2003) Spectral shear-wave ground-motion scaling in Switzerland. *Bull Seis Soc Am* 93(1):414–429
19. Bell JW, Brune JN, Zeng Y (2004) Methodology for obtaining constraints of ground motion from precariously balanced rocks. *Bull Seis Soc Am* 94:285–303
20. Ben-Zion Y (2003) Appendix 2, Key Formulas in Earthquake Seismology. In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) International Handbook of Earthquake and Engineering Seismology, Part B. Academic Press, San Diego, pp 1857–1875
21. Ben-Zion Y, Andrews DJ (1998) Properties and implications of dynamic rupture along a material interface. *Bull Seis Soc Am* 88(4):1085–1094
22. Ben-Zion Y, Sammis CG (2003) Characterization of Fault Zones. *Pure Appl Geophys* 160:677–715
23. Ben-Zion Y, Zhu L (2002) Potency-magnitude scaling relations for southern California earthquakes with  $1.0 < M_L < 7.0$ . *J Geophys Int* 148:F1–F5
24. Benites R, Olsen KB (2005) Modeling strong ground motion in the Wellington Metropolitan Area, New Zealand. *Bull Seis Soc Am* 95(6):2180–2196
25. Beresnev IA (2003) Uncertainties in finite-fault slip inversions: To what extent to believe? (A critical review). *Bull Seis Soc Am* 93(6):2445–2458
26. Beresnev IA, Wen KL (1996) Nonlinear soil response – A reality. *Bull Seis Soc Am* 86:1964–1978
27. Bernard P, Madariaga R (1984) A new asymptotic method for the modeling of near field accelerograms. *Bull Seis Soc Am* 74:539–558
28. Beroza GC, Mikumo T (1996) Short slip duration in dynamic rupture in the presence of heterogeneous fault properties. *J Geophys Res* 101(10):22,449–22,460
29. Beroza GC, Spudich P (1988) Linearized Inversion for Fault Rupture Behavior. Application- to the 1984 Morgan-Hill, California, Earthquake. *J Geophys Res* 93(B6):6275–6296
30. Beyer K, Bommer JJ (2006) Relationships between median values and between aleatory variabilities for different definitions of the horizontal component of motion. *Bull Seis Soc Am* 96(4):1512–1522
31. Bindi D, Parolai S, Grosser H, Milkereit C, Durukal E (2007) Empirical ground-motion prediction equations for northwestern Turkey using the aftershocks of the 1999 Kocaeli earthquake. *Geophys Res Lett* 34:L08305. doi:[10.1029/2007GL029222](https://doi.org/10.1029/2007GL029222)
32. Bolt BA, Abrahamson NA (2002) Estimation of strong seismic ground motions. In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) International Handbook of Earthquake and Engineering Seismology, Part B. Academic Press, San Diego, 983–1001
33. Bommer JJ, Abrahamson NA, Strasser FO, Pecker A, Bard P-Y, Bungum H, Cotton F, Fäh D, Sabetta F, Scherbaum F, Studer J (2004) The challenge of defining upper bounds on earthquake ground motions. *Seis Res Lett* 75(1):82–95
34. Bommer JJ, Martinez-Pereira A (1998) The effective duration of earthquake strong motion. *J Earthq Eng* 3(2):127–172
35. Boore DM (1983) Stochastic simulation of high frequency ground motions based on seismological models of the radiation spectra. *Bull Seismol Soc Am* 73:1865–1894
36. Boore DM (2001) Comparisons of ground motions from the 1999 Chi-Chi earthquake with empirical predictions largely based on data from California. *Bull Seis Soc Am* 91(5):1212–1217
37. Boore DM, Atkinson GM (2008) Updated Reports of Next Generation Attenuation (NGA) Models, PEER Lifelines Program, report and auxiliary material available at <http://peer.berkeley.edu/products/Boore-Atkinson-NGA.html>
38. Boore DM, Joyner WB et al (1997) Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes; a summary of recent work. *Seis Res Lett* 68 (1):128–153
39. Boore DM, Lamprey JW, Abrahamson NA (2006) Orientation-independent measures of ground motion. *Bull Seismol Soc Am* 96(4A):1502–1511. doi:[10.1785/0120050209](https://doi.org/10.1785/0120050209)
40. Borcherdt RD (1994) Estimates of site-dependent response spectra for design (methodology and justification). *Earthquake Spectra* 10(4):617–653
41. Borcherdt RD (2002) Empirical evidence for acceleration-dependent amplification factors. *Bull Seis Soc Am* 92(2):761–782
42. Bouchon M, Aki K (1982) Strain, tilt, and rotation associated with strong ground motion in the vicinity of earthquake faults. *Bull Seismol Soc Am* 72:1717–1738
43. Bouchon M, Streiff D (1997) Propagation of a shear crack on a nonplanar fault: A method of calculation. *Bull Seis Soc Am* 87(1):61–66

44. Bouchon M, Vallee M (2003) Observation of long supershear rupture during the magnitude 8.1 Kunlunshan earthquake. *Science* 301(5634):824–826
45. Brune JN (1970) Tectonic Stress and Spectra of Seismic Shear Waves from Earthquakes. *J Geophys Res* 75(26):4997–5009
46. Brune JN (1996) Precariously balanced rocks and ground-motion maps for southern California. *Bull Seis Soc Am* 86(1): 43–54
47. Brune JN (1999) Precarious rocks along the Mojave section of the San Andreas Fault, California: Constraints on ground motion from great earthquakes. *Seis Res Lett* 70:29–33
48. Brune JN (2003) Precarious rock evidence for low near-source accelerations for trans-tensional strike-slip earthquakes. *Physics of the Earth and Planetary Interiors* 137(1–4):229–239
49. Brune JN, Anooshehpoor A, Purvance MD (2006) Band of precariously balanced blocks between the Elsinore and San Jacinto, California, fault zones: Constraints on ground motion for large earthquakes. *Geology* 34:137–140
50. Brune JN, Whitney JW (1992) Precariously balanced rocks with rock varnish—Paleoindicators of maximum ground acceleration? *Seis Res Lett* 63:21
51. Cakir Z, de Chabalier JB, Armijo R, Meyer B, Barka A, Peltzer G (2003) Coseismic and early post-seismic slip associated with the 1999 Izmit earthquake (Turkey), from SAR interferometry and tectonic field observations. *Geophys J Int* 155 (1):93–110
52. Campbell KW (2002) Strong-motion attenuation relations. In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) International Handbook of Earthquake and Engineering Seismology, Part B. Academic Press, San Diego, pp 1003–1012
53. Campbell KW, Bozorgnia Y (2003) Updated near-source ground-motion (attenuation) relations for the horizontal and vertical components of peak ground acceleration and acceleration response spectra. *Bull Seis Soc Am* 93(3):1413–1413
54. Choi Y, Stewart JP (2005) Nonlinear site amplification as function of 30 m shear-wave velocity. *Earthq Spectr* 21(1):1–30
55. Clinton JF (2004) Modern digital seismology-instrumentation, and small amplitude studies in the engineering world. <http://resolver.caltech.edu/CaltechETD:etd-05202004-225044>
56. Custodio S, Lui P-C, Archuleta RJ (2005) The 2004  $M_w$  6.0 Parkfield, California, earthquake: inversion of near-source ground motion using multiple data sets. *Geophys Res Lett* 32:L23312
57. Davis P, Rubinstein JL, Liu K, Gao S, Knopoff L (2000) Northridge earthquake damage caused by geologic focusing of seismic waves. *Science* 289(5485):1746–1750
58. Day SM (1982a) Three-dimensional finite difference simulation of fault dynamics; rectangular faults with fixed rupture velocity. *Bull Seis Soc Am* 72(3):705–727
59. Day SM (1982b) 3-Dimensional Simulation of Spontaneous Rupture – the Effect of Non-uniform Prestress. *Bull Seis Soc Am* 72(6):1881–1902
60. Douglas J (2003) Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth-Science Reviews* 61:43–104
61. Dumbser M, Käser M (2006) An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes II– The three dimensional case. *Geophys J Int* 167:319–336
62. Dunham E, Archuleta RJ (2004) Evidence for a super-shear transient during the 2002 Denali earthquake. *Bull Seis Soc Am* 94:S256–S268
63. Dunham E, Favreau P, Carlson J (2003) A supershear transition mechanism for cracks. *Science* 299:1557–1559
64. Field EH, Zeng YH, Johnson PA, Beresnev I (1998) Nonlinear sediment response during the 1994 Northridge earthquake: Observations and finite source simulations. *J Geophys Res* 103(B11):26869–26883
65. Frankel A, Clayton RW (1984) A finite-difference simulation of wave propagation in two-dimensional random media. *Bull Seis Soc Am* 74(6):2167–2186
66. Frankel A, Clayton RW (1986) Finite-Difference Simulations of Seismic Scattering Implications– for the Propagation of Short-Period Seismic-Waves in the Crust and Models of Crustal Heterogeneity. *J Geophys Res* 91(B6):6465–6489
67. Fukushima Y, Tanaka T (1990) A new attenuation relation for peak horizontal acceleration of strong earthquake ground motion in Japan. *Bull Seis Soc Am* 80(4):757–783
68. Fukuyama E, Olsen KB (2002) A condition for super-shear rupture propagation in a heterogeneous stress field. *Pure Appl Geophys* 159(9):2047–2056
69. Fäh D, Kind F, Giardini D (2001) A theoretical investigation of average H/V ratios. *Geophys J Int* 145(2):535–549
70. Fäh D, Rüttener E, Noack T, Kruspan P (1997) Microzonation of the city of Basel. *J of Seism* 1:87–102
71. Goff JA, Holliger K (1999) Nature and origin of upper crustal seismic velocity fluctuations and associated scaling properties: Combined stochastic analyses of KTB velocity and lithology logs. *J Geophys Res* 104(B6):13169–13182
72. Graves RW (1996) Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences. *Bull Seis Soc Am* 86(4):1091–1106
73. Graves RW (1998) Three-dimensional finite-difference modeling of the San Andreas fault: source parameterization and ground-motion levels. *Bull Seis Soc Am* 88(4):881–897
74. Graves RW, Clayton RW (1992) Modeling path effects in 3-dimensional basin structures. *Bull Seis Soc Am* 82(1):81–103
75. Graves RW, Pitarka A, Somerville P (1998) Ground-motion amplification in the Santa Monica area: Effects of shallow basin-edge structure. *Bull Seis Soc Am* 88(5):1224–1242
76. Graves RW, Wald DJ (2004) Observed and simulated ground motions in the San Bernardino basin region for the Hector Mine, California, earthquake. *Bull Seis Soc Am* 94(1):131–146
77. Guatteri M, Mai PM, Beroza GC (2004) A pseudo-dynamic approximation to dynamic rupture models for strong ground motion prediction. *Bull Seis Soc Am* 94(6):2051–2063
78. Guatteri M, Mai PM, Beroza GC, Boatwright J (2003) Strong-ground motion prediction from stochastic-dynamic source models. *Bull Seis Soc Am* 93(1):301–313
79. Hanks TC, Bakun WH (2002) A bilinear source-scaling model for M-log A observations of continental earthquakes. *Bull Seis Soc Am* 92(5):1841–1846
80. Harris RA (2004) Numerical simulations of large earthquakes: dynamic rupture propagation on heterogeneous faults. *Pure Appl Geophys* 161:2171–2181
81. Harris RA, Archuleta RJ et al (1991) Fault steps and the dynamic rupture process; 2-D numerical simulations of a spontaneously propagating shear fracture. *Geophys Res Lett* 18(5):893–896
82. Harris RA, Day SM (1993) Dynamics of fault interaction; parallel strike-slip faults. *J Geophys Res* 98(3):4461–4472

83. Harris RA, Day SM (1997) Effects of a low-velocity zone on a dynamic rupture. *Bull Seis Soc Am* 87(5):1267–1280
84. Harris RA, Day SM (1999) Dynamic 3D simulations of earthquakes on en echelon faults. *Geophys Res Lett* 26(14):2089–2092
85. Hartzell S, Guatteri M, Mai PM, Liu P-C, Fiske M (2005) Calculation of broadband time histories of ground motion: Part II, Kinematic and dynamic modeling with theoretical Green's functions. *Bull Seis Soc Am* 95(2):614–645
86. Hartzell S, Liu P-C, Mendoza C, Ji C, Larson K (2007) Stability and uncertainty of finite-fault slip inversions: Application to the 2004 Parkfield, California, earthquake. *Bull Seis Soc Am* 97:1911–1934
87. Heaton TH (1982) The 1971 San-Fernando Earthquake – a Double Event? *Bull Seis Soc Am* 72(6):2037–2062
88. Heaton TH (1990) Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Phys Earth Planet Int* 64:1–20
89. Hikima K, Koketsu K (2004) Source processes of the foreshock, mainshock and largest aftershock in the 2003 Miyagi-ken Hokubu, Japan, earthquake sequence. *Earth Planets Space* 56:87–93
90. Hillers G, Mai PM, Ampuero J-P, Ben-Zion Y (2007) Statistical properties of seismicity of fault zones at different evolutionary stages. *Geophys J Int* 169:515–533. doi:[10.1111/j.1365-246X.2006.03275.x](https://doi.org/10.1111/j.1365-246X.2006.03275.x)
91. Hillers G, Wesnousky S (2008) Scaling relation of strike-slip earthquakes with different rate-state-dependent properties at depth. *Bull Seis Soc Am* 98(3):1085–1101
92. Hisada Y (2001) A theoretical omega-square model considering spatial variation in slip and rupture velocity. Part 2: Case for a two-dimensional source model. *Bull Seis Soc Am* 91(4):651–666
93. Holliger K (1997) Seismic scattering in the upper crystalline crust based on evidence from sonic logs. *Geophys J Int* 128(1):65–72
94. Horikawa H (2001) Earthquake doublet in Kagoshima, Japan: Rupture of asperities in a stress shadow. *Bull Seis Soc Am* 91(1):112–127
95. Huang BS (2003) Ground rotational motions of the 1991 Chi-Chi, Taiwan earthquake as inferred from dense array observations. *Geophys Res Lett* 30(6):1307–1310
96. Ide S (2001) Complex source processes and the interaction of moderate earthquakes during the earthquake swarm in the Hida-Mountains, Japan, 1998. *Tectonophysics* 334(1):35–54
97. Igel H, Cochard A, Wassermann J, Schreiber U, Velikoseltsev A, Dinh NP (2007) Broadband observations of rotational ground motions. *Geophys J Int* 168(1):182–197
98. Igel H, Schreiber U, Flaws A, Schuberth B, Velikoseltsev A, Cochard A (2005) Rotational motions induced by the M8.1 Tokachi-oki earthquake, September 25, 2003. *Geophys Res Lett* 32:L08309. doi:[10.1029/2004GL022336](https://doi.org/10.1029/2004GL022336)
99. Irikura K, Kamae K (1994) Estimation of strong ground motion in broad-frequency band based on a seismic source scaling model and an empirical Green's function technique. *Annali Di Geofisica* 37:1721–1743
100. Iwata T, Sekiguchi H (2002) Source process of the 2000 western Tottori Prefecture earthquake and near-source strong ground motion. *Proceedings of the Japan. Earthquake Eng. Symposium*, vol 11. pp 125–128
101. Johnson KM, Hsu YJ, Segall P, Yu SB (2001) Fault geometry and slip distribution of the 1999 Chi-Chi, Taiwan earthquake imaged from inversion of GPS data. *Geophys Res Lett* 28(11):2285–2288
102. Jonsson S, Zebker H, Segall P, Amelung F (2002) Fault slip distribution of the 1999 M-w 7.1 Hector Mine, California, earthquake, estimated from satellite radar and GPS measurements. *Bull Seis Soc Am* 92(4):1377–1389
103. Joyner WB, Boore DM (1988) Measurement, characterization, and prediction of strong ground motion. Conference on Earthquake Engineering and Soil Dynamics II: Recent Advances in Ground Motion Evaluation, Park ASCE City UT, USA, p 43–102
104. Joyner WB, Spudich P (1994) Including near-field terms in the isochrones integration method for application to finite-fault or Kirchhoff boundary integral problems. *Bull Seis Soc Am* 84:1260–1265
105. Kagawa T, Irikura K, Somerville PG (2004) Differences in ground motion and fault rupture process between surface and buried rupture earthquakes. *Earth Planets Space* 56(1):3–14
106. Kanamori H, Anderson DL (1975) Theoretical basis of some empirical relations in seismology. *Bull Seis Soc Am* 65(5):1073–1095
107. Kawase H (1996) The cause of the damage belt in Kobe: The basin edge effect – constructive interference of the direct S-wave with the basin-induced diffracted Rayleigh waves. *Seis Res Lett* 67(5):25–34
108. Kawase H (2003) Site effects on strong ground motions In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) International Handbook of Earthquake and Engineering Seismology, Part B. Academic Press, San Diego, pp 1013–1030
109. Koketsu K, Kikuchi M (2000) Propagation of seismic ground motion in the Kanto basin, Japan. *Science* 288:1237–1239
110. Kostrov BV (1964) Self-similar problems of propagating shear cracks. *App J Math Mech* 28:1077–1078
111. Kramer SL (1996) Geotechnical Earthquake Engineering. Prentice Hall, Upper Saddle River, New Jersey
112. Käser M, Mai PM, Dumbser M (2007) Geometrically complex finite source rupture models and their accurate treatment using tetrahedral meshes. *Bull Seis Soc Am* 97(5):1570–1586
113. Lavallee D, Archuleta R (2003) Stochastic modeling of slip spatial complexities for the 1979 Imperial Valley, California, earthquake. *Geophys Res Lett* 30:1245. doi:[10.1029/2002GL015839](https://doi.org/10.1029/2002GL015839)
114. Lavallee D, Liu PC, Archuleta R (2006) Stochastic model of heterogeneity in earthquake slip spatial distributions. *Geophys J Int* 165(2):622–640
115. Lee WHK, Shin T-C, Kuo KW, Chen KC, Wu CF (2001) CWB free-field strong-motion data from the 21 September Chi-Chi, Taiwan, earthquake. *Bull Seis Soc Am* 91(5):1370–1376
116. Liu P, Archuleta RJ (2004) A new nonlinear finite fault inversion with three-dimensional Green's functions: Application to the 1989 Loma Prieta, California, earthquake. *J Geophys Res* 109:B02318
117. Liu P-C, Custodio S, Archuleta RJ (2006) Kinematic inversion of the 2004 M 6.0 Parkfield earthquake including an approximation to site effects. *Bull Seis Soc Am* 96:143–158
118. Ma S, Archuleta RJ, Page MT (2007) Effects of large-scale surface topography on ground motions, as demonstrated by a study of the San Gabriel Mountains, Los Angeles, California. *Bull Seis Soc Am* 97(6):2066–2079

119. Madariaga R (1983) High frequency radiation from dynamic earthquake fault models. *Ann Geophys* 1(1):17–23
120. Mai PM (2001) Characterizing earthquake source complexity for improved strong motion prediction. Ph D thesis, Department of Geophysics, Stanford University, California
121. Mai PM (2004) SRCMOD: an online database of finite-source rupture models. <http://www.seismo.ethz.ch/srcmod>, updated July 2007; last accessed Aug 2007
122. Mai PM, Beroza GC (2000) Source-scaling properties from finite-fault rupture models. *Bull Seis Soc Am* 90(3):604–615
123. Mai PM, Beroza GC (2002) A spatial random-field model to characterize complexity in earthquake slip. *J Geophys Res* 107(B11):2308. doi:[10.1029/2001JB000588](https://doi.org/10.1029/2001JB000588)
124. Mai PM, Beroza GC (2003) A hybrid method for calculating near-source broadband seismograms: application to strong motion prediction. *Phys Earth Planet Int* 137:183–199
125. Mai PM, Olsen KB (2005) Broadband ground motion simulations using finite-difference synthetics with local scattering operators. 2005 Annual SCEC Meeting, Palm Springs CA, Southern California Earthquake Center
126. Mai PM, Somerville P, Pitarka A, Dalguer L, Miyake H, Beroza G, Song S-G, Irikura K (2006) Fracture-energy scaling in dynamic rupture models of past earthquakes. *Earthquakes: Radiated Energy and the Physics of Faulting* Geophysical Monograph Series, vol. 170. American Geophysical Union, 10.1029/170GM28:283–294
127. Mai PM, Spudich P, Boatwright J (2005) Hypocenter locations in finite-source rupture models. *Bull Seis Soc Am* 95(3):965–980
128. Monelli D, Mai PM (2008) Bayesian inference of kinematic earthquake rupture parameters through fitting of strong motion data. *Geophys J Int* 173:220–232. doi:[10.1111/j.1365-246X.2008.03733.x](https://doi.org/10.1111/j.1365-246X.2008.03733.x)
129. Nakamura H, Miyatake T (2000) An approximate expression of slip velocity time functions for simulation of near-field strong ground motion. *Zishin (J Seis Soc Jpn)* 53:1–9
130. O'Connell DRH (1999) Replication of apparent nonlinear seismic response with linear wave propagation models. *Science* 283(5410):2045–2050
131. Oglesby DD, Archuleta RJ (2003) The three-dimensional dynamics of a nonplanar thrustfault. *Bull Seis Soc Am* 93(5): 2222–2235
132. Oglesby DD, Day SM (2001) Fault geometry and the dynamics of the 1999 Chi-Chi (Taiwan) earthquake. *Bull Seis Soc Am* 91(5):1099–1111
133. Oglesby DD, Day SM (2002) Stochastic fault stress: Implications for fault dynamics and ground motion. *Bull Seis Soc Am* 92(8):3006–3021
134. Oglesby DD, Dreger DS, Harris RA, Ratchkovski N, Hansen R (2004) Inverse kinematic and forward dynamic models of the 2002 Denali fault earthquake, Alaska. *Bull Seis Soc Am* 94(6):S214–S233
135. Olsen KB (2000) Site amplification in the Los Angeles basin from three-dimensional modeling of ground motion. *Bull Seis Soc Am* 90(6):77–94
136. Olsen KB, Archuleta RJ (1996) Three-dimensional simulation of earthquakes on the Los Angeles fault system. *Bull Seis Soc Am* 86(3):575–596
137. Olsen KB, Archuleta RJ, Matarese JR (1995) 3-Dimensional simulation of a magnitude-7.75 Earthquake on the San Andreas fault. *Science* 270(5242):1628–1632
138. Olsen KB, Day SM, Minster JB, Cui Y, Chourasia A, Faerman M, Moore R, Maechling P, Jordan T (2006) Strong shaking in Los Angeles expected from southern San Andreas earthquake. *Geophys Res Lett* 33:L07305. doi:[10.1029/2005GL025472](https://doi.org/10.1029/2005GL025472)
139. Oprsal I, Fäh D, Mai PM, Giardini D (2005) Deterministic earthquake scenario for the Basel area Simulating – strong motion and site effects for Basel (Switzerland). *J Geophys Res* 110:B04305. doi:[10.1029/2004JB003188](https://doi.org/10.1029/2004JB003188)
140. PEER (2008) Next Generation Attenuation of Ground Motions (NGA) Project. [http://peer.berkeley.edu/products/nga\\_project.htm](http://peer.berkeley.edu/products/nga_project.htm)
141. Peng ZG, Ben-Zion Y (2006) Temporal changes of shallow seismic velocity around the Karadere–Duzce branch of the North Anatolian Fault and strong ground motion. *Pure Appl Geophys* 163:567–600
142. Piatanesi A, Cirella A, Spudich P, Cocco M (2007) A global search inversion for earthquake kinematic rupture history: application to the 2000 western Tottori, Japan earthquake. *J Geophys Res* B07314. doi:[10.1029/2006JB004821](https://doi.org/10.1029/2006JB004821)
143. Pitarka A, Irikura K (1996) Modeling 3D surface topography by finite-difference method; Kobe-JMA station site, Japan, case study. *Geophys Res Lett* 23(20):2729–2732
144. Pitarka A, Irikura K, Iwata T, Kagawa T (1996) Basin structure effects in the Kobe area inferred from the modeling of ground motions from two aftershocks of the January 17, 1995, Hyogo-ken Nanbu earthquake. *J Phys Earth* 44(5):563–576
145. Pitarka A, Irikura K, Iwata T, Sekiguchi H (1998) Three-dimensional simulation of the near-fault ground motion for the 1995 Hyogo-ken Nanbu (Kobe), Japan, Earthquake. *Bull Seis Soc Am* 88(2):428–440
146. Pitarka A, Somerville P, Fukushima Y, Uetake T, Irikura K (2000) Simulation of near-fault strong-ground motion using hybrid Green's function. *Bull Seis Soc Am* 90(3):566–586
147. Pulido N, Kubo T (2004) Near-fault strong motion complexity of the 2000 Tottori earthquake (Japan) from a broadband source asperity model. *Tectonophysics* 390:177–192
148. Reid HF (1910) The Mechanics of the Earthquake: The California Earthquake of April 18, 1906, Report of the State Investigation Commission, Vol. 2. Carnegie Institution of Washington, Washington DC
149. Ripperger J, Mai PM (2004) Fast computation of static stress changes on 2D faults from final slip distributions. *Geophys Res Lett* 31(18):L18610. doi:[10.1029/2004GL020594](https://doi.org/10.1029/2004GL020594)
150. Ripperger J, Mai PM, Ampuero J-P (2008) Near-Field Ground Motion from Dynamic Earthquake Rupture Simulations. *Bull Seis Soc Am* 98(3):1207–1228
151. Ripperger J, Mai PM, Ampuero J-P, Giardini D (2007) Earthquake source characteristics from dynamic rupture with constrained stochastic fault stress. *J Geophys Res* 112:B04311. doi:[10.1029/2006JB004515](https://doi.org/10.1029/2006JB004515)
152. Ritter JRR, Mai PM, Stoll G, Fuchs K (1997) Scattering of teleseismic waves in the lower crust Observations – in Massif Central, France. *Phys Earth Planet Int* 104:127–146
153. Robinson DP, Brough C et al (2006) The M-w 7.8:2001 Kunlunshan earthquake: Extreme rupture speed variability and effect of fault geometry. *J Geophys Res* B08303. doi:[10.1029/2005JB004137](https://doi.org/10.1029/2005JB004137)
154. Romanowicz B (1992) Strike-slip earthquakes on quasi-vertical transcurrent faults; inferences for general scaling relations. *Geophys Res Lett* 19(5):481–484

155. Rosakis AJ, Samudrala O, Coker D (1999) Cracks faster than the shear wave speed. *Science* 284(5418):1337–1340
156. Rubinstein JL, Beroza GC (2004) Evidence for widespread nonlinear strong ground motion in the M-W 6.9 Loma Prieta Earthquake. *Bull Seis Soc Am* 94(5):1595–1608
157. Sadigh K, Chang CY, Egan JA, Makdisi F, Youngs RR (1997) Attenuation relationships for shallow crustal earthquakes based on California strong motion data. *Seis Res Lett* 68(1):180–189
158. Sato H, Fehler M (1998) Seismic wave propagation and scattering in the heterogeneous Earth, Press AIP/Springer, New York
159. Sato H, Fehler M, Wu R-S (2003) Scattering and attenuation of seismic waves in the lithosphere. In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) International Handbook of Earthquake and Engineering Seismology, Part A. Academic Press, San Diego, p 195–208
160. Sato T, Graves RW, Somerville PG (1999) Three-dimensional finite-difference simulations of long-period strong motions in the Tokyo metropolitan area during the 1990 Odawara earthquake (MJ 5.1) and the great 1923 Kanto earthquake (MS 8.2) in Japan. *Bull Seis Soc Am* 89:579–607
161. Sato T, Graves RW, Somerville PG, Kataoka S (1998) Estimates of regional and local strong motions during the great 1923 Kanto, Japan, earthquake (Ms 8.2). Part 2: Forward simulation of seismograms using variable-slip rupture models and estimation of near-fault long-period ground motions. *Bull Seis Soc Am* 88(1):206–227
162. Sato T, Helmberger DV, Somerville PG, Graves RW, Saikia CK (1998) Estimates of regional and local strong motions during the great 1923 Kanto, Japan, earthquake (Ms 8.2). Part 1: Source estimation of a calibration event and modeling of wave propagation paths. *Bull Seis Soc Am* 88(1):183–205
163. Scherbaum F, Schmedes J, Cotton F (2004) On the conversion of source-to-site distance measures for extended earthquake Source models. *Bull Seis Soc Am* 94(3):1053–1069
164. Scholz C (2002) The mechanics of earthquakes and faulting, 2nd edn. Cambridge University Press, Cambridge
165. Sekiguchi H, Irikura K, Iwata T (2002) Source inversion for estimating the continuous slip distribution on a fault-introduction of Green's functions convolved with a correction function to give moving dislocation effects in subfaults. *Geophys J Int* 150(2):377–391
166. Shakal A, Haddadi H, Graizer V, Lin K, Huang M (2006) Some key features of the strong-motion data from the M 6.0 Park-field, California, earthquake of 28 September 2004. *Bull Seis Soc Am* 96(4B):S90–S118
167. Shakal AF, Haddadi HR, Huang MJ (2006) Note on the Very-High-Acceleration Fault Zone 16 Record from the 2004 Park-field Earthquake. *Bull Seis Soc Am* 96(3):S119–S128
168. Sigbjörnsson R, Ambraseys NN (2003) Uncertainty analysis of strong-motion and seismic hazard. *Bull Earthq Eng* 1:321–347
169. Somerville P, Irikura K, Graves R, Sawada S, Wald DJ, Abrahamson N, Iwasaki Y, Kagawa T, Smith N, Kowada A (1999) Characterizing crustal earthquake slip models for the prediction of strong ground motion. *Seis Res Lett* 70(1):59–80
170. Somerville PG (2003) Magnitude scaling of the near fault rupture directivity pulse. *Phys Earth Planet Int* 37:201–212
171. Somerville PG, Smith NF, Graves RW, Abrahamson NA (1997) Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity. *Seis Res Lett* 68(1):199–222
172. Spudich P, Chiou BSJ, Graves R, Collins N, Somerville P (2004) A formulation of directivity for earthquake sources using isochrone theory. United States Geological Survey, Open-File Report 2004-1268
173. Spudich P, Frazer LN (1984) Use of ray theory to calculate high-frequency radiation from earthquake sources having spatially variable rupture velocity and stress drop. *Bull Seis Soc Am* 74(6):2061–2082
174. Spudich P, Hellweg M, Lee WHK (1996) Directional topographic site response at Tarzana observed in aftershocks of the 1994 Northridge, California, earthquake; implications for mainshock motions. *Bull Seis Soc Am* 86(1, Part Suppl B):193–208
175. Spudich P, Joyner WB, Lindh AG, Boore DM, Margaris BM, Fletcher JB (1999) SEA99: a revised ground motion prediction relation for use in extensional tectonic regimes. *Bull Seis Soc Am* 89(1):1156–1170
176. Spudich P, Xu L (2003) Software for calculating earthquake ground motions from finite faults in vertically varying media In: Lee WHK, Kanamori H, Jennings PC, Kisslinger C (eds) International Handbook of Earthquake and Engineering Seismology, Part B. Academic Press, San Diego, pp 1857–1875
177. Strasser FO, Bommer JJ, Abrahamson NA (2008) Truncation of the distribution of ground-motion residuals. *J Seis* 12:79–105
178. Sørensen MB, Opsal I, Bonnefoy SC, Atakan K, Mai PM, Pulido N, Yalciner C (2006) Local site effects in Ataköy, Istanbul, Turkey, due to a future large earthquake in the Marmara Sea. *Geophys J Int* 167(3):1413–1424
179. Takeo M (1998) Ground rotational motions recorded in near-source region of earthquakes. *Geophys Res Lett* 25(6):789–792
180. Tinti E, Fukuyama E, Piatenesi A, Cocco M (2005) A kinematic source-time function compatible with earthquake dynamics. *Bull Seis Soc Am* 95(4):1211–1223
181. Tinti E, Spudich P, Cocco M (2005) Earthquake fracture energy inferred from kinematic rupture models on extended faults. *J Geophys Res* 110:B12303. doi:[10.1029/2005JB003644](https://doi.org/10.1029/2005JB003644)
182. Trifunac MD (1982) A note on rotational components of earthquake motions on ground surface for incident body waves. *Soil Dyn Earthq Eng* 1:11–19
183. Trifunac MD (2006) Effects of torsional and rocking excitations on the response of structures. In: Teisseire R et al (eds) Earthquake Source Asymmetry, Structural Media and Rotation Effects. Springer, Heidelberg, pp 569–582
184. Wald DJ, Heaton TH (1994) Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bull Seis Soc Am* 84(3):668–691
185. Wald DJ, Heaton TH, Hudnut KW (1996) The slip history of the 1994 Northridge, California, earthquake determined from strong-motion, teleseismic, GPS, and leveling data. *Bull Seis Soc Am* 86(1):S49–S70
186. Wang JH, Ou SS (1998) On scaling of earthquake faults. *Bull Seis Soc Am* 88(2):758–766
187. Wells DL, Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bull Seis Soc Am* 84(4):974–1002
188. Working Group on California Earthquake Probabilities (1999) Earthquake probabilities in the San Francisco Bay region: 2000 to 2030—A summary of findings. Open USGS-File Report 99–517

189. Working Group on California Earthquake Probabilities (2003) Earthquake probabilities in the San Francisco Bay region: 2002–2031. Open USGS-File Report 03-214
190. Wu RS, Aki K (1985) The fractal nature of the inhomogeneities in the lithosphere evidenced from seismic-wave scattering. *Pure Appl Geophys* 123(6):805–818
191. Wu Y-M, Shin T-C, Chang C-H (2001) Near real-time mapping of peak ground acceleration and peak ground velocity following a strong earthquake. *Bull Seis Soc Am* 91(5):1218–1228
192. Zeng YH (1993) Theory of scattered P-wave and S-wave energy in a random isotropic scattering medium. *Bull Seis Soc Am* 83(4):1264–1276
193. Zeng YH, Anderson JG, Su F (1995) Subevent rake and random scattering effects in realistic strong ground motion simulation. *Geophys Res Lett* 22(1):17–20
194. Zeng YH, Su F, Aki K (1991) Scattering wave energy propagation in a random Isotropic scattering medium. 1. Theory. *J Geophys Res* 96(B1):607–619

## Books and Reviews

- Abercrombie R, McGarr A, Kanamori H, Di Toro G (eds) (2006) *Earthquakes: Radiated Energy and the Physics of Faulting*. Geophysical Monograph Series, 170. American Geophysical Union, 10.1029/170GM28
- Lay T, Wallace TC (1995) *Modern Global Seismology*. Academic Press, San Diego
- Lee WHK, Kanamori H, Jennings PC, Kisslinger C (2002) *International Handbook of Earthquake an Engineering Seismology*, Part A and B. Academic Press, San Diego

## Web Links

- Database of finite-source rupture models <http://www.seismo.ethz.ch/srcmod>
- COSMOS strong-motion database <http://db.cosmos-eq.org>
- Kyoshin Network (Japan) <http://www.k-net.bosai.go.jp>
- PEER (Pacific Earthquake Engineering Research Center) <http://peer.berkeley.edu>
- NGA Project page at PEER [http://peer.berkeley.edu/products/nga\\_project.html](http://peer.berkeley.edu/products/nga_project.html)
- International Working Group on Rotational Seismology (IWGoRS) <http://www.rotationalseismology.org>

## Group Model Building

ETIËNNE A. J. A. ROUWETTE, JAC A. M. VENNIX  
Institute for Management Research, Radboud University,  
Nijmegen, The Netherlands

### Article Outline

- Glossary
- Definition of the Subject
- Introduction
- Group Model Building: Basic Ideas and Concepts
- Designing Group Model Building Projects

[Conducting Group Model Building Sessions](#)  
[Researching Group Model Building Effectiveness](#)  
[Future Directions](#)  
[Bibliography](#)

## Glossary

- Facilitator** Person who guides the group process in group model building.
- Gatekeeper** Person who forms the linking pin between modeling team and management team.
- Knowledge elicitation** Process of capturing the knowledge contained in the mental models of team members of the management team.
- Modeler** Person who constructs the quantified model during group model building.
- Recorder** Person who takes notes during group model building sessions and constructs workbooks.
- Reference mode** Graph(s) showing the behavior of the problem over time.
- Workbook** Booklet which contains summary of previous group model building sessions and prepares for subsequent sessions.
- Client** Person (or agency) who buys a model.

## Definition of the Subject

Computer (simulation) models have been used to support policy and decision making in the decades after World War II. Over the years modelers learned that the application of these models to policy problems was not as straightforward as had been thought initially. As of the beginning of the 1970s studies started to appear that questioned the use of large-scale computer models to support policy and decision making (cf. [24,31]). Lee's article bears the significant title: "Requiem for large scale models", a statement that leaves little room for ambiguity. Other authors who have studied the impact records of computer models also seem rather sceptical (e.g. [9,22,25,73]). It is interesting to note that Greenberger et al., after interviewing both modelers and policy makers (for whom the models were constructed) found that modelers generally pointed to the fact that they learned a lot from modeling a particular policy issue. Policy makers on the other hand indicated that they did not really understand the models nor had much confidence in them. The results of these studies pointed in the direction of learning from computer models, i. e. conceptual or enlightenment use rather than instrumental use, where policy recommendations could straightforwardly be deduced from the model analysis and outcomes. In other words it is in the process of modeling a policy problem where the learning takes place which

is required to (re)solve a problem. And it is also in this process that one needs to anticipate the implementation of policy changes. By the end of the 1970s system dynamics modelers pointed out that implementation of model outcomes was a neglected area (e. g. [50,74]) and that modelers sometimes naively assumed that implementation was straightforward, thereby neglecting organizational decision making as a political arena.

In other words client participation in the process of model construction and analysis is required for successful modeling and implementation of insights from the model into policy making. Or as Meadows and Robinson put it:

Experienced consultants state that the most important guarantee of modelling success is the interested participation of the client in the modelling process (p. 408 in [34]).

Over the years this has given rise to all kinds of experiments to involve clients in the process of model construction. In the 1990s the term Group Model Building was introduced to refer to more or less structured approaches for client involvement in system dynamics model construction and analysis.

## Introduction

From the early days of the field, the topic of client involvement in the process of model construction has raised attention in the system dynamics literature. Jay Forrester, the founder of the field of system dynamics, has repeatedly indicated that most of the knowledge needed to construct a system dynamics models can be found in the mental database of the participants of the system to be modeled [20,21]. Over the years several system dynamics modelers have experimented with approaches to involve client (groups) in model construction and analysis. This development in the system dynamics community parallels a movement in the operational research and systems fields towards more attention for stakeholders' opinions. A number of authors (e. g. [1]) criticized traditional OR and systems approaches as unsuitable for ill-structured problems that arise from differences between stakeholders' views on the problem. For ill-structured problems a range of new methods was developed [35].

The developments in the system dynamics, operational research and systems communities have given rise to a set of distinct methods and approaches. However, practitioners work on problems that have clear similarities to those encountered in other disciplines and frequently borrow techniques from one another. The boundaries between methods are therefore difficult to draw and there is a de-

gree of overlap between approaches in and between fields. Below we first describe the distinguishing characteristics of system dynamics, as this separates group model building most clearly from other approaches fostering client involvement. We then describe a number of distinct group model building approaches.

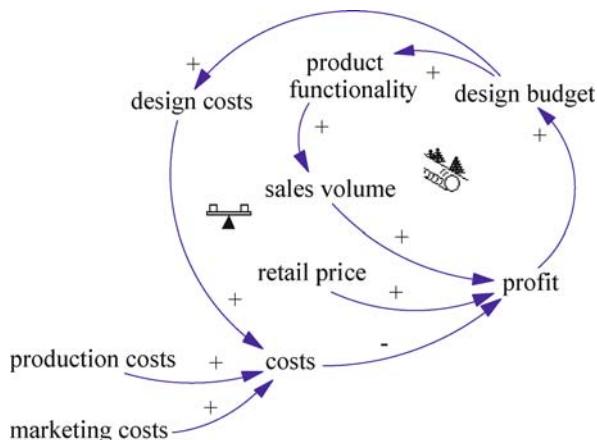
## System Dynamics

System dynamics is most easily characterized by its emphasis on two ideas: (a) the importance of closed loops of information and action for social systems, i. e. social systems as information feedback systems and (b) the need to use formal models to study these loops. System dynamicists assume that the dynamic behavior of a social system is the result of its underlying feedback structure. Actors use the information about the structure as input to their decisions, and by implementing their decision influence system behavior. This creates an interlocked chain of action and information which is also known as a feedback loop. Richardson (see p. 1 in [44]) describes a feedback loop as follows:

The essence of the concept ... is a circle of interactions, a closed loop of action and information. The patterns of behavior of any two variables in such a closed loop are linked, each influencing, and in turn responding to the behavior of the other.

As an illustration of the use of information on the system state in decisions, imagine a simple example on customer behavior. Let us assume that if customers perceive that a product's functionality increases, more products will be bought. This will increase profits and thereby the design budget. An increased design budget can be used to improve the product's design, which will lead more customers to buy the product, and so on. Thus, decisions of actors within the system have an important influence on the system's behavior. If we continue to add other factors and relations to our example and capture these in a model, the diagram in Fig. 1 may result.

As Fig. 1 shows, a causal loop diagram consists of variables, relationships, and feedback loops. Relations can be of two types: positive and negative. A positive relation indicates that both variables change in the same direction. In the model above, an increase in retail price will lead to an increase in profits, indicating a positive relationship. Variables in a negative relationship change in opposite directions. An increase in costs will decrease profits, indicating a negative relationship. The snowball rolling down the slope in the right hand side of Fig. 1 indicates a positive feedback loop. We assumed that an increase in profit re-



**Group Model Building, Figure 1**  
Example of a causal loop diagram

sults in a direct increase in the design budget. A higher budget allows for increased product functionality, which increases sales volume and finally profit. Starting from an increase in profit, the result is a further increase in profit. This is a so-called positive or self-reinforcing loop. However, if we assume that the design department uses its complete budget each year, an increased budget will contribute to design costs and lower profits. This is a negative or balancing loop, indicated by the balance symbol.

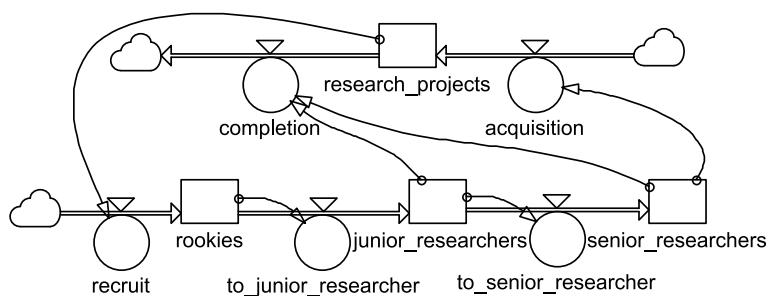
The second important idea in system dynamics is that formal models are necessary to understand the consequences of system structure. Since system dynamics models contain many (often non-linear) relations and feedback loops, it becomes very difficult to predict their behavior without mathematical simulation. Systems are assumed to consist of interacting feedback loops, which may change in dominance over time. Diagrams such as the one depicted above are frequently used in the interaction with clients. Before the dynamic consequences of the structure captured in Fig. 1 can be studied, it is necessary to further

specify both the variables and relations used in the model. Two categories of variables are distinguished: stocks and flows. Stocks are entities existing at a certain time period, for example supplies, personnel, or water in a reservoir. Flows are entities measured over a time period, such as deliveries, recruitment, or inflow of water. Relationships are separated into physical flows and information flows. If we capture differences between stocks and flows and information and physical flows in a diagram, a stock and flows diagram results.

As can be seen in Fig. 2, information links are depicted with a single arrow and physical flows with a double arrow. The physical human resources flow is separated in three stocks: number of rookies, number of junior researchers, and number of senior researchers. Recruitment will lead to an increase in the number of rookies. Two other flows influence the number of people in the stocks: rookies may be promoted to junior researchers and junior researchers may be promoted to senior researchers. The human resources flow is related to the project flow with information links, for instance indicating that acquisition of research projects is determined by the number of senior researchers.

### Group Model Building Approaches

As pointed out before, client involvement has been important to system dynamics from the start of the field. System dynamics emphasizes feedback loops and the use of formal models. In this section we describe how models based on these ideas are built in interaction with actors and stakeholders in the problem at hand. A number of different participative model building formats can be identified, i. e. the reference group approach [43,61]; the stepwise approach [76]; the Strategic Forum [49]; modeling as learning [29]; strategy dynamics [71,72] and Hines' "standard method" [39]. Below we describe each approach briefly.



**Group Model Building, Figure 2**  
Example of a stock and flows diagram

In the Reference Group approach [43,61] participation takes the form of frequent interaction between the modeling team and a group of eight to ten clients. The approach starts with the identification of interest groups, of which representatives are invited to contribute to the modeling effort. The representatives are referred to as referents. In a series of interviews and meetings, the problem to be addressed is defined more specifically. On the basis of this definition and the information gathered in the interviews and meetings, the modeling team develops a preliminary model. In the remainder of the project the modelers are responsible for model improvements while the referents function as critics. This model is elaborated in a series of meetings and is at the same time used as a tool for structuring the discussion. In later sessions, model output is used for developing scenarios. In a scenario discussion the model is run and results are described and analyzed by the modelers. The reference group is then asked to determine to what extent the model's behavior corresponds to their expectations about reality, and if it does not, to suggest changes. These suggestions can trigger changes in the model structure, initiating a new round in the discussions.

The stepwise approach [76] is founded on the idea that full quantification of models is not always possible or desirable. The approach starts with a definition of the problematic behavior. If possible, this definition is given in the form of a behavior over time of the problem of interest. Modeling starts by roughly sketching the feedback loops responsible for this behavior. The key variables related to the cause for concern are identified, followed by the system resources connected to these key variables and their initial states. The resources are used to derive the central stocks in the system. From the resources, the resource flows can then be sketched with the associated rates of conversion. Delays are added to these flows if they are significant. Next, organizational boundaries, flows of information and strategies through which the stocks influence the flows, are added. Again, if there are significant delays, these are added to the information linkages. In the final step, information flows and strategies linking different resource flows are added. The steps are repeated until the relevant feedback loops have all been included. Wostenholme indicates that these steps often provide the insights necessary to infer system behavior from the structure, which reduces the need for quantification. Models can also be analyzed in a qualitative manner.

The steps that make up the Strategic Forum [49] provide a detailed insight of how clients are encouraged to participate in modeling. The strategic forum consists of eight steps, of which the first two are conducted before

the actual meeting (also called the forum) with the client group. The process begins with interviews prepared by a small questionnaire, in which three issues are addressed: ideas on the current situation, a statement of the vision for the future, and agreement on a preliminary map of the problem. On the basis of the interviews, the modeler constructs an integrated map and accompanying computer model. In the second step the project team designs a number of small group exercises that will be used during the forum. The exercises are aimed at discovering important structural and behavioral elements and are similar to the scenario discussions in the reference group approach. The most important difference is that before simulation results are shown, participants have to 'put a stake in the ground', i.e. they have to make a prediction of model behavior on the basis of a change in a policy variable and values for connected parameters. The model is then simulated and results are compared with participants' expectations. Discrepancies between predictions and simulations are identified, and might point to inconsistencies in participants' ideas or lead to model improvements. In the following steps the participants meet in a series of workshops. Each workshop opens with an introduction and a big picture discussion. The heart of the session consists of exercises aimed at internal consistency checks, addressing the consistency between the group's mental model and the computer model. As in the other approaches, model structure will be changed if inconsistencies with the participants' ideas on the problem are revealed. In the final phase of policy design, potential consequences of strategic policies are addressed and the existing capability of realizing the strategic objectives. A wrap-up discussion and identification of follow-up activities concludes the Strategic Forum.

Richmond (see p. 146 in [49]) emphasizes that the main purpose of the Strategic Forum is to check the consistency of strategy. The insights gained by the client therefore frequently lead to changes in strategy or operating policies, but less frequently to changes in objectives or the mission statement. One important element of ensuring an impact on participants' ideas is the (dis)confirmation of expectations on simulation outcomes.

Lane [29] describes a modeling approach developed at Shell International Petroleum, known as 'modeling as learning'. Lane explicitly sets this approach apart from the widely used expert consultancy methodology (e.g. [58]). His approach also puts strong emphasis on involving decision makers in the modeling process. By showing decision makers the benefits of participation early on in the process, an attempt is made to persuade them to spend time in direct interaction with the model. The approach centers on capturing and expressing the client's ideas, initiating a dis-

cussion on the issue with ‘no a-priori certainty regarding quantification, or even cause and effect’ (see p. 70 in [29]). The modelers also strive to include both hard as well as ‘soft’ aspects of the problematic situation. In doing this, it is hoped that the clients’ ideas are included in the model and that ownership is created. This is encouraged by making models and model output transparent to participants, helping the client ‘to learn whichever techniques are used in a project’ (see p. 71 in [29]). Lane states that the focus throughout the approach is on a process of learning, using such elements as experimentation with the model, testing of assumptions and representing and structuring ideas in a logical way.

Hines’ approach [39] starts off by diagnosing the problem. This step comes down to gathering and clustering problem variables. Problem behavior is visualized by sketching the graph over time of the problematic behavior. In the second step the structure underlying the problematic behavior is captured in a causal diagram. This so-called dynamic hypothesis incorporates many of the problem variables identified earlier. The diagram helps to clarify the boundary of the problem that will be addressed and thus limits the project scope. The next step is to identify accumulations in the system, which will form the stocks in the system dynamics model. In the construction of the computer model most work is done by the modelers, with client participation limited to providing data such as numerical values and details of the work processes relevant to the problem at hand. Model structure and behavior is then explained to the client. Discussions with the client then lead to a series of model iterations, increasing confidence of the client in model calibration and validity. Similar to other participative approaches, policy runs are used to test proposed interventions in the problem.

Warren [71,72] describes an approach to participative modeling that strongly focuses on identifying accumulations (stocks) in the system. In order to identify central accumulations, clients are asked to identify the strategic resources in the problem at hand. Increases and decreases in resources then lead to the identification of flows. Warren’s approach differs from the ones described above in the sense that stocks and flows are differentiated from the outset. This means that causal loop diagrams are not used. In addition, graphs over time are recorded next to each variable in the model. By gradually adding elements to the model while visually relating structure and behavior, the clients’ understanding of the problem is gradually increased.

As mentioned before, the boundaries around approaches are not easy to draw and one method may ‘borrow’ techniques of another. Insights and practices from the

operational research and system fields have been merged with those in system dynamics to develop combined methods. For example, modeling as learning is one of the approaches incorporating elements of soft operational research methodologies. Lane and Oliva [30] describe the theoretical basis for integrating system dynamics and soft systems methodology. The cognitive mapping approach (e.g. [17]) also offers tools and techniques that are used in system dynamics studies.

In addition to combining different methods, approaches are sometimes also tailored to use in specific content areas. An example is van den Belt’s [63] mediated modeling, which combines insights from participative system dynamics modeling and consensus building on environmental issues.

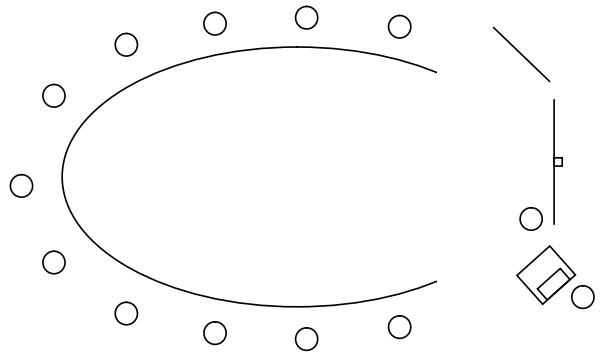
### Group Model Building: Basic Ideas and Concepts

The separate approaches described in the last section continue to be developed and used in practical problems. Although we are not sure that all proponents of these approaches would characterize themselves as using “group model building”, this term has been used more and more in the last decades to refer to system dynamics approaches with client involvement in a general sense. The two approaches that coined the term group model building evolved more or less simultaneously, with considerable cross-fertilization of ideas, at SUNY at Albany and Radboud University Nijmegen in the Netherlands (see [36,69]). In an early application at Radboud University, participants were involved in a Delphi study consisting of mailed questionnaires and workbooks, followed by workshops [66]. In the dissertations by Verburgh [70] and Akkermans [4] a similar approach is used under the name of participative policy modeling and participative business modeling, respectively. In its latest version group model building is a very open approach, which allows for the use of preliminary models or a start from scratch, uses individual interviews, documents and group sessions, qualitative or quantitative modeling and small as well as large models. Vennix [64,65] provides a set of guidelines for choosing among these different approaches, building on and adding to the studies mentioned above. Andersen and Richardson [5] provide a large number of “scripts” that can help in setting up modeling projects. The procedures described are a long way from the earlier descriptions of a set of steps that seem to prescribe standard approaches applicable to most modeling projects. Instead, the guidelines offered have more the appearance of tool boxes, from which the appropriate technique can be selected on the basis of problem characteristics and the clients involved.

Group model building is generally conducted with a group of at least six and up to 15 people. The group is guided by at least two persons: a facilitator and a modeler/recorder. The group is seated in a semi circle in front of a whiteboard and/or projection screen, which serves as a so-called group memory. A projection screen is typically used when a model is constructed with the aid of system dynamics modeling software with a graphic interface (e.g. Vensim, Powersim, Ithink). This group memory documents the model under construction and is used as a parking lot for all kinds of unresolved issues which surface during the deliberations of the group.

In Fig. 3, the small circles indicate the persons present in the session. Apart from the participants, there is a facilitator and a recorder. The facilitator has the most important role in the session as he or she guides the group process. His/her task, as a neutral outsider, is to (a) elicit relevant knowledge from the group members, (b) to (help) translate elicited knowledge into system dynamics modeling terms, and (c) make sure that there is an open communication climate so that in the end consensus and commitment will result. The recorder keeps track of the elements of the model. In Fig. 3 (s)he is seated behind a computer and the model is projected on the screen in front of the group. A separate whiteboard (upper right hand corner) is used to depict the reference mode of behavior and record comments or preliminary model structure. As the model is visible to all participants, it serves as a group memory that at each moment reflects the content of the discussion up to that point. A group model building session is generally conducted in the so-called chauffeured style, where only the facilitator uses electronic support and projection equipment, while participants do not have access to electronic communication media [38]. The central screen or whiteboard will be used to depict the model, as shown in Fig. 3.

The role of liaison between the organization and the modeling team is performed by the gatekeeper, who is generally also a member of the participant group. The gatekeeper is the contact between both parties, and has an important role in the decision which participants to involve in the sessions. Apart from the gatekeeper, the facilitator and the recorder, two other roles may be important in a modeling session [46], i.e. a process and a modeling coach. The process coach functions as an observer and primarily pays attention to the group process. The model coach needs to be experienced in system dynamics modeling but might also be an expert in the content area as well. As Richardson and Andersen [46] point out, all roles are important in group model building but not all of them have to be taken up by a single person. One person might



**Group Model Building, Figure 3**

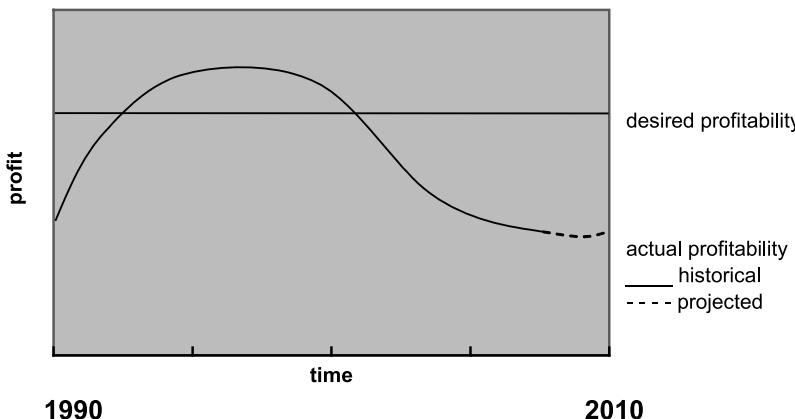
Typical room layout for group model building with participants seated in a semi-circle, white board and facilitator in front, and computer and overhead projector (adapted from [5])

for instance combine the roles of facilitator and process coach. Taken together, these different roles constitute the facilitation or modeling team.

In principle the group follows the normal steps in the construction of a system dynamics model. This means that the first step is the identification of the strategic issue to be discussed, preferably in the form of a so-called reference mode of behavior, i.e. a time series derived from the system to be modeled which indicates a (historical) undesirable development over time. As an example let us take the sales of a software product. An initial problem statement might be falling profit. Typically the problematic behavior will be depicted in a graph over time as in Fig. 4.

In the graph above, a projection of the expected behavior is included for the years after 2008.

The next step is to elicit relevant variables with which the model construction process can be started. Depending on the type of problem this will take the form of either a causal loop diagram or a stocks and flow diagram and is generally referred to as the conceptualization stage. The following step is to write mathematical equations (model formulation) and to quantify the model parameters. As described in the introductory section, most of the model formulation work is done backstage as it is quite time consuming and members of a management team generally are not very much interested in this stage of model construction. In this stage, the group is only consulted for crucial model formulations and parameter estimations. Experienced group model builders will start to construct a simple running model as soon as possible and complicate it from there on if required. In the end the model should of course be able to replicate the reference mode of behavior (as one of the many validity tests) before it can be sensibly be used as a means to simulate the potential effects of strategies and scenarios.



**Group Model Building, Figure 4**

Problem description in graphical form: reference mode of behavior

### Objectives of Group Model Building

As mentioned in the introduction, the founder of system dynamics has repeatedly pointed out that much of the knowledge and information which is needed to construct a model can be found in the mental models of system participants. At first sight it may seem that the most important objective of building a system dynamics model is to find a robust strategy to solve the problem of the organization. In the end that is why one builds these models. From this perspective the most important issue in group model building is how to elicit the relevant knowledge from the group. However, as stated before, decision making in organizations has its own logic, and in many cases there is quite some disagreement about the problem and how it should be tackled. No wonder that implementation of model outcomes is difficult if the model building process is not well integrated with decision making processes in organizations, when it comes to creating agreement and commitment with a decision. From that perspective knowledge elicitation is only one element in the process of model construction. It is not so much the model but to a greater extent the process of model construction which becomes important. Somewhat simplified one could say that in the "standard" approach when an organization is confronted with a strategic problem it hires a modeler (or group of modelers) to construct a model and come up with recommendations to "solve" the problem. However, in most cases these recommendations become part of the discussion in the management team and get misunderstood, or adapted or frequently just disappear as a potential solution from the discussion. Hence Watt's title of his paper: "Why won't anyone believe us?" becomes very much understandable from the point of view of the mod-

eler. So rather than creating a situation where modelers "take away" the problem from the organization and (after considerable time) return with their recommendations, the model building process is now used to structure the problem, guide communication about it and test the robustness of strategies taking into account other criteria and information which is not included in the model, but does play a role for the organization when making the decision. Stated differently, the model building process now becomes intertwined with the process of decision making in an organization. And this in turn means that other objectives than knowledge elicitation become important.

Simultaneously with the attempts to involve clients in the process of constructing system dynamics models the objectives of group model building have been defined at several levels, i. e. the individual, the group and the organizational level (cf. [7,65]). The main goal at the individual level is change of mental models and learning. The idea is that participants should better understand the relationship between structure and dynamics and how their interventions may create counterintuitive results. Unfortunately research has revealed that this is hardly the case. Even after extensive training people have difficulty to understand the relationship between structure and dynamics (for a review see [8,53]). A second goal at the individual level is behavioral change. Frequently the conclusions of a modeling intervention point in the direction of behavioral change, for example implementing a new job rotation scheme, or a change in purchasing policy. The question can then be asked how insights from the modeling intervention are translated to changes in behavior. Rouwette (e.g. [52,53,57]) uses a framework from social psychology to understand the impact of modeling on behavior. The theory of Ajzen [2,3] explains behavior on the basis of

### Group Model Building, Table 1

Objectives of group model building (cf. [53])

Individual	Positive reaction Mental model refinement Commitment Behavioral change
Group	Increased quality of communication Creation of a shared language Consensus and alignment
Organization	System changes System improvement or results
Method	Further use Efficiency

(a) attitude, (b) perceptions of norms and (c) perceptions of control. It seems likely that each of these concepts is influenced in modeling sessions. When for example model simulations reveal unexpected levers for improving system behavior, we can expect that perceived control will increase. Another example: let's imagine that a manager is participating in a modeling session, where another participant reveals positive outcomes of a certain policy option. If these positive outcomes were previously not known to the manager, hearing them might make his/her attitude towards that option more positive (cf. [42]).

At the group level objectives refer to mental model alignment [28] and fostering consensus [51,67,75]. Creating consensus should not be confused with premature consensus, i. e. not discussing conflicting viewpoint. Here it concerns creation of consensus after critical debate and discussion of opinions has taken place. This type of discussion which needs to take place in a cooperative communication climate is helpful to also create commitment with the resulting decision.

At the organizational level goals have been discussed as system process change (are things done differently) and system outcome change (are customers impacted differently) [11]. Although it has to be pointed out that in many cases system changes are the result of changes in attitude and behavior of participants in the system. An overview of group model building objectives is given in Table 1. In this table finds a number of additional objectives such as positive reaction and creation of a shared language, that are more fully reviewed by Huz et al. [28], Rouwette et al. [55] and Rouwette and Vennix [53].

### Designing Group Model Building Projects

When designing group model building projects there are a number of questions that need to be addressed. The first concerns the suitability of system dynamics for the prob-

lem at hand. System dynamicists generally say that a problem needs to be dynamically complex in order to be suitable to model it through system dynamics. This means that one should at least hypothesize that there are positive and negative feedback process underlying the problem. From a more practical point of view one could say that one should be able to represent the problem in the form of a reference mode of behavior. If the latter is not possible one should seriously question the use of system dynamics for the problem.

A second issue which needs to be given some thought is the question whether to use qualitative or quantitative modeling. Within the system dynamics community there is still a debate about the question whether qualitative modeling (or: mapping) can be considered system dynamics (see [13,14,23]). In short those who disagree point out that without quantification and simulation one cannot reliably develop a robust policy simply because the human mind is not capable of predicting the dynamic effects of (interventions in) a dynamically complex structure. Those who do use mapping on the other hand point out that mapping in itself can have the beneficial effect to structure the problem and at least will make managers aware of potential underlying feedback loops and their potential counterintuitive effects when intervening in a dynamically complex system. Basically the issue to quantify or not depends on the goals of the group model building intervention. If the ultimate goal is to find robust policies then quantification is required. However, if the aim is to structure a problem and to create consensus on a strategic issue then qualitative modeling may be all that is needed. This links up with Zagonel's [77] distinction between the use of models as micro worlds or as boundary objects. When used as a boundary object the emphasis is on supporting negotiation and exchange of viewpoints in a group. This is clearly the case when problems are messy, i. e. connected to other problems and when there is much divergence of opinion on what the problem is and sometimes even whether there is a problem at all.

A third issue is the question who to involve in the sessions. There are a number of criteria which are generally employed. First it is important to involve people who have the power to make decisions and changes. A second criterion is to involve people who are knowledgeable about the problem at hand. A third criterion is to involve a wide variety of viewpoints, in order to make sure that all relevant knowledge about the problem is included. Of course these guidelines may create dilemmas. For example, involving more people in the process will make the group communication process more difficult. This may in turn endanger the creation of consensus and commitment.

**Group Model Building, Table 2**

Group model building scripts (cf. [5])

Phase in modeling	Script
Defining a problem	Presenting reference modes
	Eliciting reference modes
	Audience, purpose, and policy options
Conceptualizing model structure	Sectors, a top down approach
	Maintain sector overview while working within a sector
	Stocks and flows, by sector
	Name that variable or sector
Eliciting feedback structure	Direct feedback loop elicitation
	Capacity utilization script
	System archetype templates
	"Black box" means-ends script
Equation writing and parametrization	Data estimation script
	Model refinement script
	"Parking lot" for unclear terms
Policy development	Eliciting mental model-based policy stories
	Create a matrix that links policy levers to key system flows
	"Complete the graph" policy script
	Modeler/reflector feedback about policy implications
	Formal policy evaluation using multiattribute utility models
	Scripts for "ending with a bang"

Another issue is whether to use a preliminary model or to start from scratch (see [45]). Although using a preliminary model may speed up the process the inherent danger is that it will be difficult to build group ownership over the model. Group ownership is clearly required to create consensus and commitment.

Finally, a range of methods and techniques is available to elicit relevant knowledge both from individuals and from groups. When it comes to individuals, well known methods are interviews, questionnaires and so-called workbooks. The latter are a kind of modified questionnaires, which are used in between sessions to report back to the group and ask new question in preparation of the next session. Interviews are being used routinely as a preparation for group model building sessions.

If a decision is made on the issues discussed above, the next important question is how to plan and execute the modeling sessions. This question is a central topic in the group model building literature and its success heavily depends on the correct choice of available techniques and the quality of the facilitator.

### Conducting Group Model Building Sessions

Although careful preparation of group model building sessions is a necessity, the most important part of the whole

project is what happens in the group model building sessions themselves. During the sessions not only the analysis of the problem takes place (and the model is constructed), but also the interaction process between members of the management team unfolds. It is this interaction process which needs to be guided in such a way that consensus and commitment will emerge and implementation of results will follow. As pointed out the process is guided by the group facilitator, generally someone who is not only specialized in facilitation of group processes but also in system dynamics model construction. The facilitator is supported by a recorder or modeler who helps constructing the system dynamics model while the facilitator interacts with the management team.

The facilitator may choose from a wide variety of techniques in setting up and conducting a session. As a foundation for choosing techniques, Andersen and Richardson [5] develop a set of guiding principles and so-called scripts for group model building sessions. Guiding principles capture basic ideas in the interaction with clients, such as break task/group structure several times each day, clarify group products, maintain visual consistency and avoid talking heads. Scripts are more concrete instances of these principles and refer to small elements of the interaction process [5,32]. The Table 2 shows scripts described in Andersen and Richardson's [5] original paper.

In choosing a script it is first important to be aware of the phase that is relevant in the project at that time. A common starting point, as we saw in the description of group model building approaches, is to define the central problem of interest. The reference mode of behavior can function as a guideline for involving clients in this phase. Once the central problem is clear, a logical next step is to move towards model conceptualization. In this step again a number of options are available. Andersen and Richardson [5] describe a script for identifying sectors that are important in the problem. An alternative is to start with more concrete variables in the problem, using a Nominal Group Technique [15].

Whatever scripts and techniques a facilitator employs it is important that (s)he displays the right attitude and uses the correct skills. Several different aspects of attitude are important. First of all the facilitator is not the person who thinks (s)he knows the best solution, but needs to be helpful in guiding the group to find a solution to the strategic problem the organization is faced with. Second, a facilitator should be neutral with respect to the problem that is being discussed. Being too knowledgeable about a particular problem area (e.g. strategic alliances) may thus be dangerous, because it creates the tendency to participate in the discussions. Rather than being an expert, having an inquiry attitude (i.e. asking questions rather than providing answers) is more helpful to the group. Finally, integrity and being authentic is important. Relying on tricks to guide the process will prove counterproductive, because people will look through them.

When it comes to skills, a thorough knowledge and experience in constructing system dynamics models is of course indispensable. Second, a facilitator needs to be knowledgeable about group process and have the skills to structure both the strategic problem as well as the group interaction process. For the latter, special group process techniques (e.g. brainstorming, Nominal Group Technique, Delphi) may be used, and knowledge about and skills in applying these techniques is of course a prerequisite for a successful group model building intervention. Finally, communication skills are important. Reflective listening is a skill which will help to prevent misunderstanding in communication, both between participants and the facilitator and between group members. For a more thorough discussion of these attitudes and skill in the context of group model building we refer to Vennix [64].

### Researching Group Model Building Effectiveness

In the previous sections we described goals of group model building projects and principles and scripts for guiding the

modeling process. In this section we consider the empirical evidence for a relation between modeling interventions and these intended outcomes. Empirical evidence can be gathered using a variety of research strategies, such as (field) experiments, surveys or (in-depth) case studies. According to the review of modeling studies by Rouwette et al. [55], the case study is the most frequently used design to study group model building interventions. We first report on the results found by these authors and then turn to other designs.

In the meta-analysis of Rouwette et al. [55], the majority of group model building studies uses a case study design and assesses outcomes in a qualitative manner. Data are collected using observation, and a minority of studies employs individual group interviews. Case reports may be biased towards successful projects and are frequently not complete. The outcomes of the modeling projects were scored along the dimensions depicted in Table 1 in the section on modeling goals. The findings show positive outcomes in almost all dimensions of outcomes. Learning about the problem seems to be a robust outcome of group model building, for example:

- Of 101 studies that report on learning effects, 96 indicate a positive effect;
- Of 84 studies focusing on implementation of results, 42 report a positive effect.

Another set of studies, using quantitative assessment of results is described by Rouwette and Vennix [53]. Although the research surveyed so far indicates positive effects of modeling on outcomes such as mental model refinement, consensus and implementation of results, important challenges remain. Research so far has paid little attention to the complexity of the intervention as described in the previous section. Pawson and Tilley [40] urge us not to assume that interventions are similar and lead to similar effects, since this would confuse meaningful differences between studies. Rouwette and Vennix [53] describe two ways to learn more about the process through which outcomes of modeling are created: base research more on theory and/or to conduct research in more controlled settings. At present only few studies address elements of group model building in a controlled setting. Shields [59,60] investigates the effect of type of modeling and facilitation on a group task. Most research on the use of system dynamics models concerns so-called management flight simulators. These studies aim to mimic the important characteristics of decision making in complex, dynamic problems, and test the effectiveness of various decision aids. Results are reviewed by Sterman [62], Hsiao and Richardson [26] and Rouwette, Größler and Vennix [56].

Increased attention to theories may shed more light on the way in which modeling effects group decisions. Theories can help in specifying relations which can then be tested. Explanatory research is needed to connect the components and outcomes of group model building interventions (see p. 194 in [7]). In the field of system dynamics modeling, two attempts at formulating theories on modeling components and outcomes are the work of Richardson et al. [47] and Rouwette [52]. The framework formulated by Rouwette [57] builds on the theories of Ajzen [2,3] and Petty and Cacioppo [42] described earlier. Richardson et al. [47] separate mental models into means, ends and means-ends models. The ends model contains goals, while the means model consists of strategies, tactics, and policy levers. The means-ends model contains the connection between the two former types of models and may contain either detailed "design" logic or more simple "operator" logic. On the basis of research on participants in a management flight simulator [6], the authors conclude that operator logic, or high level heuristics, is a necessary condition for improving system performance. Therefore, providing managers with operator knowledge is the key to implementation of system changes.

## Future Directions

The success of group model building and problem structuring methodologies in general depends on a structured interaction between theory, methodology refinement and application in practical project accompanied by systematic empirical evaluation.

Rouwette and Vennix [53] indicate three areas for further development of theories:

- Review related methodologies used in complex organizational problems, to determine which theories are used to explain effects. Examples that come to mind are theories used in the operational research and systems fields [37].
- Forge a closer connection to research on electronic meeting systems. In this field, studies are usually conducted in controlled settings [16,18,41] and theory development seems to be at a more advanced stage. Research on electronic meeting systems is interesting both because of the empirical results and explanatory theories used and because of insights on the intervention process. A recent development in the field is research on ThinkLets [10]. A ThinkLet is defined as a named, packaged facilitation intervention and thus seems very similar to the concept of a group model building script.
- A third source of theories is formed by research in psychology and group decision making. Theories from

these fields inform the definition of central concepts in group model building (see Table 1) and theories on modeling effectiveness. Rouwette and Vennix [54] review literature on group information processing and relate this to elements of group model building interventions.

From theories and evaluation research will come insights to further develop the methodology along the lines of (a) determining what kind of problem structuring methodology is best suited in what kind of situation, (b) refinement of procedures, (c) better understanding the nature of the intervention, and (d) better guidelines for facilitators how to work in different kinds of groups and situations.

## Bibliography

### Primary Literature

1. Ackoff RA (1979) The future of operational research is past. *J Oper Res Soc* 30(2):93–104
2. Ajzen I (1991) The theory of planned behavior. *Organ Behav Human Decis Process* 50:179–211
3. Ajzen I (2001) Nature and operation of attitudes. *Annu Rev Psychol* 52:27–58
4. Akkermans HA (1995) Modelling with managers: Participative business modelling for effective strategic decision-making. Ph D Dissertation, Technical University Eindhoven
5. Andersen DF, Richardson GP (1997) Scripts for group model-building. *Syst Dyn Rev* 13(2):107–129
6. Andersen DF, Maxwell TA, Richardson GP, Stewart TR (1994) Mental models and dynamic decision making in a simulation of welfare reform. Proceedings of the 1994 International System Dynamics Conference: Social and Public Policy. Sterling, Scotland, pp 11–18
7. Andersen DF, Richardson GP, Vennix JAM (1997) Group model-building: Adding more science to the craft. *Syst Dyn Rev* 13(2):187–201
8. Andersen DF, Vennix JAM, Richardson GP, Rouwette EAJA (2007) Group model building: Problem structuring, policy simulation and decision support. *J Oper Res Soc* 58(5):691–694
9. Brewer GD (1973) Politicians, bureaucrats and the consultant: A critique of urban problem solving. Basic Books, New York
10. Briggs RO, de Vreede GJ, Nunamaker JF (2003) Collaboration engineering with ThinkLets to pursue sustained success with group support systems. *J Manag Inf Syst* 19:31–63
11. Cavalieri S, Sterman JD (1997) Towards evaluation of systems thinking interventions: A case study. *Syst Dyn Rev* 13(2):171–186
12. Checkland P (1981) Systems thinking, systems practice. Wiley, Chichester/New York
13. Coyle G (2000) Qualitative and quantitative modelling in system dynamics: Some research questions. *Syst Dyn Rev* 16(3):225–244
14. Coyle G (2001) Maps and models in system dynamics: Rejoinder to Homer and Oliva. *Syst Dyn Rev* 17(4):357–363
15. Delbecq AL, van de Ven AH, Gustafson DH (1975) Group techniques for program planning: A guide to nominal group and delphi processes. Scott, Foresman and Co, Glenview

16. Dennis AR, Wixom BH, van den Berg RJ (2001) Understanding fit and appropriation effects in group support systems via meta-analysis. *Manag Inf Syst Q* 25:167–183
17. Eden C, Ackermann F (2001) SODA – The principles. In: Rosenhead J, Mingers J (eds) Rational analysis for a problematic world revisited. Problem structuring methods for complexity, uncertainty and conflict. Wiley, Chichester, pp 21–42
18. Fjermestad J, Hiltz SR (1999) An assessment of group support systems experimental research: Methodology and results. *J Manag Inf Syst* 15:7–149
19. Fjermestad J, Hiltz SR (2001) A descriptive evaluation of group support systems case and field studies. *J Manag Inf Syst* 17:115–160
20. Forrester JW (1961) Industrial Dynamics. MIT Press, Cambridge
21. Forrester JW (1987) Lessons from system dynamics modelling. *Syst Dyn Rev* 3(2):136–149
22. Greenberger M, Crenson MA, Crissey BL (1976) Models in the policy process: Public decision making in the computer era. Russel Sage Foundation, New York
23. Homer J, Oliva R (2001) Maps and models in system dynamics: A response to Coyle. *Syst Dyn Rev* 17(4):347–355
24. Hoos IR (1972) Systems analysis in public policy: A critique. University of California Press, Berkeley/Los Angeles/London
25. House PW (1982) The art of public policy analysis: The arena of regulations and resources, 2nd printing. Sage, Thousand Oaks
26. Hsiao N, Richardson GP (1999) In search of theories of dynamic decision making: A literature review. In: Cavana RY et al (eds) Systems thinking for the next millennium – Proceedings of the 17th international conference of the system dynamics society. Wellington, New Zealand
27. Huz S (1999) Alignment from group model building for systems thinking: Measurement and evaluation from a public policy setting. Unpublished doctoral dissertation, SUNY, Albany
28. Huz S, Andersen DF, Richardson GP, Boothroyd R (1997) A framework for evaluating systems thinking interventions: an experimental approach to mental health system change. *Syst Dyn Rev* 13(2):149–169
29. Lane DC (1992) Modelling as learning: A consultancy methodology for enhancing learning in management teams. *Eur J Oper Res* 59(1):64–84. Special issue of Eur J Oper Res: Morecroft JDW, Sterman JD (eds) Modelling for learning
30. Lane DC, Oliva R (1998) The greater whole: Towards a synthesis of system dynamics and soft systems methodology. *Eur J Oper Res* 107(1):214–235
31. Lee DB (1973) Requiem for large-scale models. *J Am Inst Plan* 39(1):163–178
32. Luna-Reyes L, Martinez-Moyano I, Pardo T, Cresswell A, Andersen D, Richardson G (2006) Anatomy of a group model-building intervention: Building dynamic theory from case study research. *Syst Dyn Rev* 22(4):291–320
33. Maxwell T, Andersen DF, Richardson GP, Stewart TR (1994) Mental models and dynamic decision making in a simulation of welfare reform. In: Proceedings of the 1994 international system dynamics conference, Stirling, Scotland. Social and Public Policy. System Dynamics Society, Lincoln, pp 11–28
34. Meadows DH, Robinson JM (1985) The electronic oracle: Computer models and social decisions. Wiley, Chichester/New York
35. Mingers J, Rosenhead J (2004) Problem structuring methods in action. *Eur J Oper Res* 152:530–554
36. Morecroft JDW, Sterman JD (1992) Modelling for learning. Special issue *Eur J Oper Res* 59(1):28–41
37. Morton A, Ackermann F, Belton V (2003) Technology-driven and model-driven approaches to group decision support: focus, research philosophy, and key concepts. *Eur J Inf Syst* 12:110–126
38. Nunamaker JF, Dennis AR, Valacich JS, Vogel DR, George JF (1991) Electronic meetings to support group work. *Commun ACM* 34(7):40–61
39. Otto PA, J Struben (2004) Gloucester fishery: Insights from a group modeling intervention. *Syst Dyn Rev* 20(4):287–312
40. Pawson R, Tilley N (1997) Realistic evaluation. Sage, London
41. Pervan G, Lewis LF, Bajwa DS (2004) Adoption and use of electronic meeting systems in large Australian and New Zealand organizations. *Group Decis Negot* 13(5):403–414
42. Petty RE, Cacioppo JT (1986) The elaboration likelihood model of persuasion. *Adv Exp Soc Psychol* 19:123–205
43. Randers J (1977) The potential in simulation of macro-social processes, or how to be a useful builder of simulation models. Gruppen for Ressursstudier, Oslo
44. Richardson GP (1991) Feedback thought in social science and systems theory. University of Pennsylvania Press, Philadelphia
45. Richardson G (2006) Concept models. Größler A, Rouwette EAJA, Langer RS, Rowe JI, Yanni JM (eds) Proceedings of the 24th International Conference of the System Dynamics Society, Nijmegen
46. Richardson GP, Andersen DF (1995) Teamwork in group model-building. *Syst Dyn Rev* 11(2):113–137
47. Richardson GP, Andersen DF, Maxwell TA, Stewart TR (1994) Foundations of mental model research. Proceedings of the 1994 international system dynamics conference, Stirling, Scotland. pp 181–192. System Dynamics Society, Lincoln
48. Richmond B (1987) The strategic forum: From vision to strategy to operating policies and back again. High Performance Systems Inc, Highway, Lyme
49. Richmond B (1997) The strategic forum: Aligning objectives, strategy and process. *Syst Dyn Rev* 13(2):131–148
50. Roberts EB (1978) Strategies for effective implementation of complex corporate models. In: Roberts EB (ed) Managerial applications of system dynamics. Productivity Press, Cambridge, pp 77–85
51. Rohrbaugh JW (1992) Collective challenges and collective accomplishments. In: Bostron RP, Watson RT, Kinney ST (eds) Computer augmented team work: A guided tour. Van Nostrand Reinhold, New York, pp 299–324
52. Rouwette EAJA (2003) Group model building as mutual persuasion. Ph D Dissertation, Radboud University Nijmegen
53. Rouwette EAJA, Vennix JAM (2006) System dynamics and organizational interventions. *Syst Res Behav Sci* 23(4):451–466
54. Rouwette EAJA, Vennix JAM (2007) Team learning on messy problems. In: London M, Sessa VI (eds) Work group learning: Understanding, improving & assessing how groups learn in organizations. Lawrence Erlbaum Associates, Mahwah, pp 243–284
55. Rouwette EAJA, Vennix JAM, van Mullekom T (2002) Group model building effectiveness: A review of assessment studies. *Syst Dyn Rev* 18(1):5–45
56. Rouwette EAJA, Größler A, Vennix JAM (2004) Exploring influencing factors on rationality: A literature review of dynamic decision-making studies in system dynamics. *Syst Res Behav Sci* 21(4):351–370

57. Rouwette EAJA, Vennix JAM, Felling AJA (2008) On evaluating the performance of problem structuring methods: An attempt at formulating a conceptual framework. Paper accepted for Group Decision and Negotiation
58. Schein EH (1987) Process consultation, vol II. Addison-Wesley, Reading
59. Shields M (2001) An experimental investigation comparing the effects of case study, management flight simulator and facilitation of these methods on mental model development in a group setting. In: Hines JH, Diker VG, Langer RS, Rowe JI (eds) Proceedings of the 20th international conference of the system dynamics society, Atlanta
60. Shields M (2002) The role of group dynamics in mental model development. In: Davidsen PI, Mollona E, Diker VG, Langer RS, Rowe JI (eds) Proceedings of the 20th international conference of the system dynamics society, Palermo
61. Stenberg L (1980) A modeling procedure for public policy. In: Randers J (ed) Elements of the system dynamics method. Productivity Press, Cambridge, pp 292–312
62. Sterman JD (1994) Learning in and about complex systems. *Syst Dyn Rev* 10(2–3):291–330
63. van den Belt M (2004) Mediated modeling. A system dynamics approach to environmental consensus building. Island Press, Washington
64. Vennix JAM (1996) Group model-building: Facilitating team learning using system dynamics. Wiley, Chichester
65. Vennix JAM (1999) Group model building: Tackling messy problems. *Syst Dyn Rev* 15(4):379–401
66. Vennix JAM, Gubbels JW, Post D, Poppen HJ (1990) A structured approach to knowledge elicitation in conceptual model-building. *Syst Dyn Rev* 6(2):31–45
67. Vennix JAM, Schepers W, Willems R (1993) Group model-building: What does the client think of it? In: Zepeda E, Machuca J (eds) The role of strategic modelling in international competitiveness. Proceedings of the 1993 international system dynamics conference Mexico, Cancun, pp 534–543
68. Vennix JAM, Akkermans HA, Rouwette EAJA (1996) Group model-building to facilitate organizational change: An exploratory study. *Syst Dyn Rev* 12(1):39–58
69. Vennix JAM, Andersen DF, Richardson GP (eds) (1997) Special issue on group model building. *Syst Dyn Rev* 13(2):187–201
70. Verburgh LD (1994) Participative policy modelling: Applied to the health care industry. Ph D Dissertation, Radboud University Nijmegen
71. Warren K (2000) Competitive strategy dynamics. Wiley, Chichester
72. Warren K (2005) Improving strategic management with the fundamental principles of system dynamics. *Syst Dyn Rev* 21(4):329–350
73. Watt CH (1977) Why won't anyone believe us? *Simulation* 28:1–3
74. Weil HB (1980) The evolution of an approach for achieving implemented results from system dynamic projects. In: Randers J (ed) Elements of the system dynamics method. MIT, Cambridge
75. Winch GW (1993) Consensus building in the planning process: Benefits from a "hard" modelling approach. *Syst Dyn Rev* 9(3):287–300
76. Wolstenholme EF (1992) The definition and application of a stepwise approach to model conceptualisation and analysis. *Eur J Oper Res* 59:123–136
77. Zagonel AA (2004) Reflecting on group model building used to support welfare reform in New York state. Unpublished doctoral dissertation, SUNY, Albany

## Books and Reviews

- Meadows DH, Richardson J, Bruckmann G (1982) *Groping in the dark: The first decade of global modelling*. Wiley, Chichester
- Schwartz RM (1994) *The skilled facilitator: Practical wisdom for developing effective groups*. Jossey-Bass, San Francisco

## Growth Models for Networks

SERGEY N. DOROGOVTSEV<sup>1,2</sup>

<sup>1</sup> Departamento de Física, Universidade de Aveiro, Aveiro, Portugal

<sup>2</sup> A. F. Ioffe Physico-Technical Institute, St. Petersburg, Russia

## Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Equilibrium and Nonequilibrium Random Networks](#)

[Random Recursive Graphs \(Trees\)](#)

[Recursive Trees Versus Equilibrium Connected Trees](#)

[The Barabási–Albert Model](#)

[General Preferential Attachment](#)

[Finite Size Effects](#)

[Hidden Variables](#)

[Condensation of Edges](#)

[Weighted Growing Networks](#)

[Connected Components in Growing Networks](#)

[Significance of Loops](#)

[Accelerated Growth of Networks](#)

[Critique of the Preferential Attachment Mechanism](#)

[Optimization-Based Models](#)

[Deterministic Graphs](#)

[Future Directions](#)

[Acknowledgments](#)

[Bibliography](#)

## Glossary

**Random network** A random network is a statistical ensemble: a given set of graphs and their statistical weights – realization probabilities. Formulating a network model, one defines the full set of members of the ensemble and indicates corresponding statistical weights or rules generating these weights.

**Nonequilibrium random network** Nonequilibrium random networks are nonequilibrium statistical ensembles of graphs. A growing network is a very particular case of a nonequilibrium one, with a growing number of edges and vertices.

**Recursive network** A recursive network is a growing network, where new edges connect only new vertices to already existing ones. In recursive networks, no new edges emerge between already existing vertices.

**Tree** A tree is a graph without loops (cycles). The absence of loops crucially simplifies the description of these networks.

**Random recursive graph (tree)** In graph theory, the random recursive graph is a tree growing in the following way: at each time step, a new vertex is attached to one of the existing vertices chosen with equal probability.

**Degree** Degree – a total number of connections of a vertex – is the simplest local property of a network. A vertex degree distribution is the probability that a vertex in a network has a given degree.

**Heavy-tailed degree distributions** These are degree distributions with divergent higher moments in the infinite network limit. In particular, power-law distributions are heavy tailed.

**Scale-free networks** In many natural and artificial networks, a sufficiently wide region of a degree distribution may be approximately fitted by a power law. These networks are called scale-free.

**Preferential attachment** In the preferential attachment process, vertices for connection/attachment are chosen preferentially – usually, with probability proportional to a given function of their degree. The preferential attachment mechanism effectively generates networks with complex architectures including networks with heavy-tailed degree distribution.

## Definition of the Subject

For several decades mathematicians, biologists, and physicists have studied growing networks and their models. In physics and physical chemistry, models of this kind were first proposed in the theory of polymers [1,2]. In graph theory and its applications, e.g., to biology and to computer science, very simple growing networks – random recursive trees – were an issue of wide interest for many years, see for example early articles [3,4]. Growth models for networks have become, maybe, the hottest topic in statistical mechanics, graph theory, and multidisciplinary research after the work [5], where these models were used to explain universal complex structures of the Internet, the World Wide Web, and other real networks.

## Introduction

Numerous natural and artificial networks have essentially more complex architectures than classical random graphs in graph theory. The classical random graphs have a rapidly decreasing Poisson distribution of the number of connections of a vertex – degree distributions. In contrast, real-world networks have degree distributions decaying in a much slower fashion. Usually, networks with a more complex structure than the classical random graphs are called complex networks. Complex networks include practically all important real-world nets: various cellular networks, the Internet, the World Wide Web, social networks, and many others.

Why do real-world networks have complex architectures? This fundamental problem essentially determines the evolution of a multidisciplinary research field, which is often called the science of complex networks. A few concepts explaining complex network structures were developed. Two of them – self-organization and optimization – were extensively studied during the last years. Modeling the creation of complex network structures is mostly based on these concepts. A significant part of real networks are evolving, usually growing networks, which strongly differ from equilibrium nets. Furthermore, technically, it is easier to arrive at heavy-tailed distributions of connections in evolving networks than in equilibrium ones. These are the main reasons for the intense interest in growth models for networks.

In this very brief survey we only discuss basic demonstrative models highlighting the set of key ideas of this field. A huge number of growth models, including more realistic ones, remain out of the scope of this article. For the sake of demonstration, we discuss only networks which are undirected (networks with undirected connections) and one-partite (there is only one sort of vertices). Many important real networks are directed, the World Wide Web, for example. Chemical reaction networks and many others are bipartite. Fortunately, generalizations to directed and multi-partite networks are rather straightforward.

## Equilibrium and Nonequilibrium Random Networks

In statistical mechanics and graph theory, a random network is not a single graph but a statistical ensemble. A given set of graphs are members of this ensemble. Each of these graphs has its probability of realization – statistical weight. These weights may be stationary or may be evolving, that is depending on time  $t$ . One of the possible ways to define a statistical ensemble is by a master equation for statistical weights. In a wide range of situations, these

equations are linear. Let  $\mathcal{P}(g, t)$  be the statistical weight of a graph  $g$  belonging to the set of graphs of a statistical ensemble  $G$ ,  $g \in G$ , and  $W(g, g')$  be transition rates from graph  $g'$  to graph  $g$ . Then the master equation is of the form

$$\frac{\partial \mathcal{P}(g, t)}{\partial t} = \sum_{g' \in G} [W(g, g')\mathcal{P}(g', t) - W(g', g)\mathcal{P}(g, t)]. \quad (1)$$

In equilibrium random networks, the left-hand side of the master equation is zero, and the right-hand side provides one with a set of detailed balance conditions for stationary statistical weights  $\mathcal{P}(g)$ . In nonequilibrium networks,  $\partial \mathcal{P}(g, t)/\partial t \neq 0$ , and a detailed balance is not reached. Growing networks represent a particular case of nonequilibrium ones, and in many network growth models, equations describing the evolution of statistical properties of a network are simple consequences of Eq. (1). Some characteristics of a growing network may become stationary as  $t \rightarrow \infty$ . This does not mean that the network approaches an equilibrium state in this limit. For example, in numerous network growth models, a degree distribution approaches a stationary limit, which is very convenient for their analysis.

### Random Recursive Graphs (Trees)

In recursive networks new edges connect only new vertices to already existing vertices, new connections between already existing vertices are not possible. Most simple growth models for networks are recursive. The simplest recursive network is *the random recursive graph* from graph theory [6]. In the random recursive graph, the growth starts with a single vertex. At each time step,

- (i) add a new vertex and
- (ii) attach it to one of the already existing vertices chosen with equal probability.

The random recursive graph has no loops – it is a tree by definition. This model is one of two fundamental random graphs in graph theory: the Erdős–Rényi graph [7] is a basic equilibrium network, and the random recursive graph is a basic growing network.

The degree distribution  $P(q)$  of the random recursive graph (the average fraction of vertices of degree  $q$ ) decays exponentially. In the infinite network limit,

$$P(q) = 2^{-q-1}, \quad (2)$$

which is valid for all degrees. This distribution decreases much slower than the Poisson degree distribution of the Erdős–Rényi model, which asymptotically decays as  $1/q!$ .

On the other hand, all moments of this distribution are finite.

The absence of loops allows one to easily describe the global organization of the network. The first characteristic of the global organization is *a mean intervertex distance*  $\bar{\ell}$  – the average length of the shortest path between two vertices in the network. The dependence of  $\bar{\ell}$  on the total number of vertices in the network,  $N$ , characterizes the network compactness. In finite-dimensional networks – “large worlds” –  $\bar{\ell} \sim N^{1/d}$ , where  $d$  may be regarded as the dimensionality of a network (more rigorously, its Hausdorff dimension). In networks with the small-world property, which are infinite-dimensional dimensional objects – “small-worlds” –  $\bar{\ell}$  increases with  $N$  slower than any positive power of  $N$ . In the random recursive tree,

$$\bar{\ell}(N) = \frac{2}{N-1} [(N+1)H_N - 2N] \stackrel{N \gg 1}{\cong} 2 \ln N - (4 - 2\gamma_e) \quad (3)$$

(see [8] and references therein) so that this network is a small world. Here  $H_N = \sum_{i=1}^N 1/i$  is the  $N$ th harmonic number and  $\gamma_e = 0.5772\dots$  is Euler’s number. The average distance of vertices from the root, approaching  $(1 - \gamma_e) \ln N$  as  $N \rightarrow \infty$ , and the diameter (the average maximum distance between two vertices in a graph of this statistical ensemble) behave similarly.

Remarkably, the relative width of the intervertex distance distribution,  $\mathcal{R}(\ell, N)$ , of the random recursive tree approaches zero in the infinite network limit. The width of this distribution grows as  $\sqrt{\bar{\ell}} \sim \sqrt{\ln N} \ll \bar{\ell}$ , and  $\mathcal{R}(\ell, N) \rightarrow \delta(\ell - \bar{\ell}(N))$  as  $N \rightarrow \infty$ . Here  $\delta(x)$  is the  $\delta$ -function. This property is rather typical for small worlds.

### Recursive Trees Versus Equilibrium Connected Trees

A difference between growing and equilibrium networks may be striking. For the sake of demonstration, let us compare two random networks – the random recursive trees and the equilibrium (or homogeneous) connected trees, [9]. The equilibrium statistical ensemble of connected trees is defined in the following way. Each of its members is a labeled connected tree. (The term “a labeled network” means the following. All vertices in these networks have individual labels. Labeled networks which differ only by permutations of the labels are counted as distinct. We discuss only labeled networks as is usual in statistical mechanics and graph theory.) The ensemble includes all possible labeled connected trees of a given size  $N$ . All the members have equal probabilities of realization. It turns out that the mean intervertex distance of this

random network scales with  $N$  as  $\bar{\ell}(N) \sim N^{1/2}$ , i.e., this kind of network may be regarded as a two-dimensional object. This is in a sharp contrast to the random recursive trees, which are small worlds, as we explained above. In the recursive tree ensemble, only those connected labeled trees are present that are allowed by the causality of the construction. So the great majority of the members – labeled trees – of the equilibrium tree ensemble are absent in the random recursive tree. This results in a quite different global organization of this equilibrium network. As for local properties, the degree distribution of equilibrium connected trees is  $P(q) = e^{-1}/(q-1)!$ , which is a much more rapidly decaying function than the exponential degree distribution (2) of the random recursive trees.

### The Barabási–Albert Model

The preferential attachment process is a very straightforward way to generate networks with numerous hubs. In this process, highly connected vertices receive new connections with higher probability (“popularity is attractive”); as a result, they become even more attractive for connection, get new edges, and so on – a network organizes itself. Actually, this very general idea is not new. In a non-network aspect, Yule in 1925, [10], and later Simon, [11], used this idea to explain skewed distributions in nature. Price applied this self-organization mechanism to scientific citations [12]. Nonetheless, it was the very timely work of Barabási and Albert [5] (see also [13]) that triggered an avalanche of studies using the preferential attachment to explain complex networks.

The Barabási–Albert model is a recursive graph, growing according to the following rules. The growth starts from some particular initial configuration of vertices and edges, which is not very important. At each time step,

- (i) add a new vertex and
- (ii) attach it to  $m \geq 1$  preferentially selected already existing vertices. Each of these vertices is chosen with a probability proportional to its degree – proportional preference.

That is, in the Barabási–Albert model, the probability of attachment to vertex  $i$  of degree  $q_i$  is  $\Pi_i = q_i / \sum_{j=1}^N q_j$ . In particular, if the initial configuration is a single vertex, and  $m = 1$ , i.e. each new vertex has only one connection, the Barabási–Albert network is a nonuniform recursive tree [14]. However, it turns out that the resulting degree distribution does not depend on  $m$  (we ignore a normalization factor), and in this respect loops generated for  $m > 1$  are not important. The degree distribution

approaches a stationary form in the large network limit,

$$P(q) = \frac{2m(m+1)}{q(q+1)(q+2)} \xrightarrow{q \rightarrow \infty} q^{-3}. \quad (4)$$

This degree distribution is heavy-tailed, its second and higher moments diverge.

One of the simple ways to obtain the proportional preference is as follows. Choose at random an edge and attach a new vertex to one or to both ends of this edge [15]. One can see that the end vertices will indeed be selected with probability proportional to their degrees. Furthermore, the attachment of a new vertex to all nearest neighbors of a randomly chosen vertex leads to the same effect – proportional preference. In a more specific situation, in uncorrelated networks, one can arrive at the proportional preference also in the following way: choose a vertex at random and then choose at random one of its nearest neighbors for attachment. The last algorithm is sometimes applicable, since the proportional preferential attachment generates recursive networks with very weak correlations (see below).

In respect of degree-degree correlations, the Barabási–Albert model is very special. The simplest degree correlations in a network are correlations between degrees of the nearest neighbor vertices. These correlations are present if the joint distribution of the degrees of the nearest neighbors does not factor in the following way:  $P(q, q') \neq qP(q)q'P(q')/(q)^2$ . If there is equality, the network is uncorrelated. Empirically, it is convenient to use a less informative object for these correlations. This is the dependence of the mean degree of the nearest neighbors of a vertex versus degree of this vertex,  $\langle q \rangle_{nn}(q)$  [16,17]. In uncorrelated networks,  $\langle q \rangle_{nn}$  does not depend on  $q$ . These correlations were observed in practically all real-world networks. We stress that only equilibrium networks may be uncorrelated. Growing (and, more general, nonequilibrium) networks, including the Barabási–Albert model and random recursive graphs, have these correlations. However, in the Barabási–Albert model and other recursive models based on the proportional preferential attachment, the degree-degree correlations are anomalously weak, and so the dependence  $\langle q \rangle_{nn}(q)$  is nearly flat. This is a quite unusual situation for growing networks.

When  $m > 1$ , this network has loops (or cycles in terms of graph theory). A loop is a closed path visiting each vertex only once. Loops of length 3 (triangles) are directly related to *clustering*. The clustering coefficient  $C_i$  of vertex  $i$  is the ratio of the number  $n_i$  of triangles attached to this vertex and  $q_i(q_i - 1)/2$  which is the maximum possible number of these triangles. Here  $q_i$  is the degree of vertex  $i$ . The *mean clustering* is the average clustering co-

efficient of a vertex:

$$\langle C \rangle = \sum_q P(q)C(q). \quad (5)$$

The clustering coefficient of a network is

$$\begin{aligned} C &= \frac{\sum_q P(q)C(q)q(q-1)}{\sum_q P(q)q(q-1)} \\ &= \frac{3 \times (\text{total number of triangles})}{\text{number of connected triples of vertices}}. \end{aligned} \quad (6)$$

In uncorrelated networks,  $C = \langle C \rangle$ . In growing networks, this is not the case. The mean number of triangles in the Barabási–Albert model is  $(m-1)m(m+1)(\ln N)^3/48$ , while the mean number of pairs of adjacent edges is  $m(m+1)(N \ln N)/2$  [18]. So the clustering coefficient of the Barabási–Albert model equals

$$C = \frac{m-1}{8} \frac{(\ln N)^2}{N}. \quad (7)$$

The mean clustering differs from expression (7) only by a factor,  $\langle C \rangle \sim (\ln N)^2/N$  [19]. Thus, a nonzero clustering coefficient in these networks is only a finite size effect:  $C$  and  $\langle C \rangle$  approach 0 as  $N \rightarrow \infty$ . More generally, the number of finite loops of length  $L$  in the Barabási–Albert model is also relatively small,  $\sim (\ln N)^L$  as  $N \rightarrow \infty$  [20]. This indicates that this network has a locally tree-like structure.

The Barabási–Albert model is a small world. However, the size dependence of the mean intervertex distance (and of diameter) noticeably differs for different  $m$ . As  $N \rightarrow \infty$ ,

$$\bar{\ell}(N, m=1) \sim \ln N, \quad \bar{\ell}(N, m>1) \sim \frac{\ln N}{\ln \ln N} \quad (8)$$

[21], which demonstrates the importance of loops in the network.

### General Preferential Attachment

In general, the attachment probability may be some given function  $f(q)$  of the degree of a vertex. For linear preference,  $f(q) = q + A$ , where the constant  $A \equiv am > -m$  is an additional attractiveness, the resulting network is scale-free. The asymptotics of the degree distribution is power-law,  $P(q) \sim q^{-\gamma}$  in the infinite network limit, where exponent  $\gamma$  is

$$\gamma = 3 + \frac{A}{m}. \quad (9)$$

So  $\gamma$  can be in the range from 2 to  $\infty$ . For an arbitrary  $q$  and  $N \rightarrow \infty$ ,

$$\begin{aligned} P(q) &= (2+a) \frac{\Gamma(m(1+a)+a+2)}{\Gamma(m(1+a))} \\ &\cdot \frac{\Gamma(k+ma)}{\Gamma(k+ma+3+a)} = \stackrel{q \rightarrow \infty}{\sim} q^{-\gamma} \end{aligned} \quad (10)$$

[22], where  $\Gamma(x)$  is the gamma-function. At  $a = A/m = 0$ , we arrive at the degree distribution (4) of the Barabási–Albert model.

A similar effect – changing exponent  $\gamma$  – can be obtained by combining random attachment and preferential one, with proportional preference. Say, for each new connection, with a probability  $p$ , choose a vertex at random and, with the complementary probability  $1-p$ , choose a vertex with probability proportional to its degree. In this case,  $\gamma = 1 + 2/(1-p)$ , for details and more realistic examples see book [23].

A more detailed characteristic in these networks is a degree distribution of a vertex born at time  $s$ , which we denote by  $p(q, s, t)$ . The age of the network asymptotically approaches its size,  $t \cong N$ , if the growth starts from a finite cluster. The following scaling form is valid for growing scale-free networks:

$$p(q, s, t) = (s/t)^\beta f(q(s/t)^\beta), \quad (11)$$

where  $f(x)$  a scaling function, and exponent  $\beta$  is related to the  $\gamma$  exponent of the degree distribution [22]:

$$\beta = \frac{1}{\gamma - 1}. \quad (12)$$

Usually, even in growing networks with wide degree distributions, the function  $f(x)$  has a narrow peak form. The mean degree of a vertex born at time  $s$  measured in a scale-free network of large age  $t$  is asymptotically expressed as

$$\langle q \rangle(s, t) \propto \left(\frac{s}{t}\right)^{-\beta}. \quad (13)$$

We have mentioned that these networks have strong correlations between degrees of the nearest neighbors [24]. In particular, in the scale-free recursive networks,

$$\langle q \rangle_{nn}(q) \propto q^{\gamma-3}. \quad (14)$$

This demonstrates that

- (i) if  $\gamma > 3$ , the correlations are *assortative* on terms of [25], that is the nearest neighbors of a strongly connected vertex, as a rule, also have many connections and vice versa;
- (ii) if  $\gamma < 3$ , the correlations are *disassortative*, that is the nearest neighbors of a strongly connected vertex, as a rule, have small degrees and vice versa; and
- (iii) in the unique case of  $\gamma = 3$ , as in the Barabási–Albert model, these correlations are very weak.

More general forms of the preference function were studied, but it turned out that only the linear preference pro-

vides power-law degree distributions. In particular, networks with a power-law preference  $f(q) = q^\gamma$  were investigated, [26]. When  $0 \leq \gamma < 1$ , the resulting degree has a stretched exponential form:

$$P(q) \propto q^{-\gamma} \exp[-g(y)q^{1-\gamma}], \quad (15)$$

where  $g(y)$  depends only on exponent  $\gamma$ . On the other hand, if  $\gamma > 1$ , a small number of the oldest vertices attract all connections. In particular, if  $\gamma > 2$ , there is even a finite probability that the first vertex gets all links, and the network is a star.

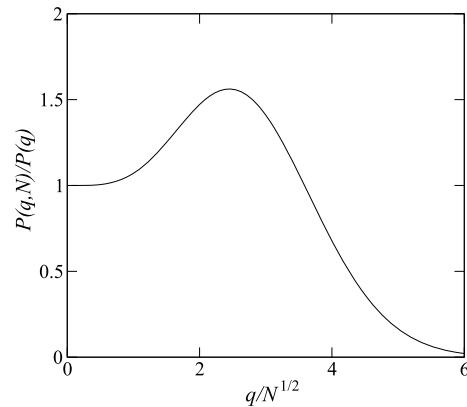
In networks, which we discussed, once added, vertices never disappear. The network evolution, however, may combine two parallel processes – both the addition and removal of vertices. Remarkably, the deletion of vertices with all their connections results in the preferential removal of edges from vertices with many connections. Similarly to Sect. “The Barabási–Albert Model”, the probability of this removal is proportional to degree – proportional preference. In a network model studied in work [27], (i) new vertices are attached to preferentially chosen existing vertices and, simultaneously, (ii) randomly selected vertices with all their connections are removed from the network. The form of the resulting degree distribution depends on the relation between the rates of these processes and also on the mean degree of the network. The exponent of the degree distribution grows as the removal rate increases. Moreover, at sufficiently high removal rates, the degree distribution may become exponentially decaying. Networks, in which connections of a removed vertex are not deleted but are redistributed in the neighborhood of this vertex, were described in [28]. The degree distributions of these networks also vary from a power law to an exponential one.

### Finite Size Effects

Even the largest artificial network, the World Wide Web, is actually small – between  $10^{10}$  and  $10^{11}$  vertices (Web pages). So finite size effects in networks are of primary importance. Here we only touch upon the size dependence of a degree distribution in scale-free recursive networks. It readily follows from Eq. (11) that the degree distribution is of the form:

$$P(q, N) = q^{-\gamma} F(q/N^{1/(\gamma-1)}), \quad (16)$$

where  $F(x)$  is an exponentially rapidly decreasing scaling function whose precise form, in principle, depends on initial conditions. A typical scaling function  $F(x)$  for exponent  $\gamma = 3$  is shown in Fig. 1 from [15]. The scaling functions for arbitrary  $\gamma$  were computed in Refs. [29,30].



Growth Models for Networks, Figure 1

Ratio  $P(q, N)/P(q, N \rightarrow \infty)$  versus the scaling variable  $q/N^{1/2}$  for a scale-free recursive network with degree distribution exponent 3 [15]. The growth starts with a pair of vertices with a connecting edge. With other initial configurations, the form of the hump will be different

Network finiteness leads to a cutoff of the degree distribution,  $q_{\text{cut}}(N)$ . In the scale-free recursive networks,  $q_{\text{cut}}(N) \sim N^{1/(\gamma-1)}$  for arbitrary  $\gamma$ , as Eq. (16) shows. This dependence coincides with a so-called “natural cutoff”, whose meaning is clear from the following estimate:  $N \int_{q_{\text{cut}}(N)}^{\infty} dq P(q) \sim 1$ , where  $P(q) = P(q, N \rightarrow \infty)$ . That is, in a single realization of a network, only one vertex of degree above the natural cutoff value may be found, and so such vertices are practically unobservable. Note, however, that the cutoff is essentially model dependent, and in some models its position is far below the natural cutoff value.

### Hidden Variables

Equation (13) shows that it is the oldest vertices that become hubs in the Barabási–Albert model and other scale-free recursive networks, where a preference function is  $f(q) = Bq + A$  with constant  $A$  and  $B$ . This is certainly not the case in real-world networks. Google, for example, has received a far greater number of hyperlinks than a huge majority of earlier created Web sites. There is a simple way to weaken “the older the richer effect” in growing networks, which was proposed in [31]. Suppose that  $B$  and  $A$  be random variables with values varying from vertex to vertex,  $A(s)$  and  $B(s)$ , where  $s = 0, 1, \dots, t$  is a label of a vertex. These  $A(s)$  and  $B(s)$  are usually called *hidden variables*, the term *fitness* is also used. It turns out that under some conditions, even with random  $A(s)$  and  $B(s)$ , a degree distribution remains heavy tailed, see [31] and, for more detail, [32]. On the other hand, the randomness of the hidden variables leads to strong fluctuations of degrees

of vertices with a given birth time. Even a young vertex  $s$  with large  $A(s)$  or/and  $B(s)$  may become rapidly a strongly connected hub.

More generally, a random preference function for a recursive graph may be a function of three variables: (i) the degree  $q$  of a vertex, which is selected for attachment, (ii) its birth time  $s$ , and (iii) the age  $t$  of the network. That is, in general, one should consider  $f(q, s, t)$ . In principle, it is possible to arrive at heavy-tailed degree distributions even when the preference function is independent of  $q$ . For example, a power-law  $f(s, t) \propto (s + 1)^{-\alpha}$  results in scale-free networks. Other forms of  $f(q, s, t)$  allow one, in particular, to take into account aging of vertices, where  $f(q, s, t) = g(q, t - s)$  [33].

### Condensation of Edges

If the coefficient  $B(s)$  of the preference function strongly fluctuates, a new phenomenon may occur [34]. The large fluctuations of  $B(s)$  may lead to condensation of edges on a few vertices in the growing network. Namely, a finite fraction of connections in a large network turns out to be attracted by a few vertices (or even by a single one).

This effect is observable even when only a single vertex in a recursive network, say, vertex  $u$ , has a fitness essentially greater than others:  $B(s) = 1 + g\delta_{su}$ . Here  $\delta_{su}$  is a Kronecker symbol. If  $g$  is greater than some critical value  $g_c = \gamma_0 - 1$ , where  $\gamma_0$  is the degree distribution exponent in the “pure” network with  $g = 0$ , then a finite fraction of edges will be attracted by vertex  $u$ . This fraction is proportional to the difference  $g - g_c$ . Other vertices have a power-law degree distribution with exponent  $\gamma = 1 + g > \gamma_0$ .

### Weighted Growing Networks

The edges of weighted networks have their own *weights* – positive real numbers. The introduction of the weights allows one to take into account the diversity of edges in many real networks [35]. For example, transport networks are weighted. Let  $w_{ij}$  be the weight of the edge connecting vertices  $i$  and  $j$ . Then the *strength* of vertex  $i$  is defined as the sum of the weights of its edges:  $s_i = \sum_j w_{ij}$ . Empirical data show that in many real-world weighted networks both the strength distribution and the degree distribution are power-law [35,36]. In some of these networks, the strength of a vertex, in average, is proportional to its degree,  $s_i \propto k_i$ , and in other ones,  $s_i \propto k_i^\theta$  with exponent  $\theta > 1$ .

A simple growth model for scale-free weighted networks is defined as follows [35,36]. At each time step,

- (i) add a new vertex and attach it to one or several preferentially chosen existing vertices by edges with some initial weight, say 1; the probability to select vertex  $i$  is taken to be proportional to its degree  $s_i$ ;
- (ii) in addition, increase the weights of the edges of each of these selected vertices: for vertex  $i$ , the edge weights are increased as  $w_{ij} \rightarrow w_{ij}(1 + \delta/s_i)$ , where  $\delta$  is a given positive parameter.

Thus, the attachment to a vertex increases its degree by 1 and its strength by  $1 + \delta$ . In the resulting network, the strengths of vertices are proportional to their degrees,  $s_i \propto k_i$ , and the strength and degree distributions are power laws with the same exponent

$$\gamma = 2 + \frac{1}{1 + 2\delta}. \quad (17)$$

One can arrive at the same results in another way [37]: choose an edge with a probability proportional to its weight, increase this weight by  $\delta$ , and simultaneously attach a new vertex by a unit weight link to one of the ends of this selected edge.

Bianconi showed how to construct a growing network with a power-law relation between vertex strength and degree [38]. Simply combine two kinds of the preferential attachment. Namely, new vertices should be attached at random either to vertices selected with a probability proportional to a vertex strength or to vertices selected with a probability proportional to a vertex degree.

### Connected Components in Growing Networks

So far we have discussed networks consisting of a single connected component, which is a rather special situation. All vertices in these networks are mutually approachable. In general, a network consists of a set of connected components of various sizes. Typically, in a random network, either (i) all connected components are relatively small (much smaller than  $N$ ) or (ii) one of the components is “giant” and the rest are relatively small. The emergence of the *giant connected component* (containing, by definition, a finite fraction of all vertices in an infinite net) crucially changes the architecture of a network. Furthermore, the statistics of connected components essentially characterizes the network organization. These statistics for growing random networks may remarkably differ from those for equilibrium networks.

A specific transition of the birth of the giant connected component in a growing network was discovered in work [39], where a very simple network was studied. The growth starts from some initial configuration, which

is not important over long periods. The network grows due to two parallel processes:

- (i) With the unit rate, bare vertices are added to the network.
- (ii) With a rate  $b$ , new edges interconnect randomly chosen vertices.

The rate  $b$  plays the role of a control parameter. When  $b$  is greater than some critical value  $b_c$ , a giant connected component is present. The degree distribution of this random graph has an exponential decay. In contrast to equilibrium networks, the birth of the giant component in this model is an infinite order phase transition. At the critical point, the size of the giant connected component has a specific singularity

$$S \propto \exp(-\text{const}/\sqrt{b - b_c}), \quad (18)$$

coinciding with that in the famous Berezinskii–Kosterlitz–Thouless phase transition. It was found that the result (18) is also valid for growing scale-free networks [40], where the value of the constant in the exponential is determined by exponent  $\gamma$ . This singularity was later observed in many other growing networks [41,42,43,44].

The size distribution of finite components in the networks with this transition has an exponential decay in the presence of the giant connected component. Without a giant component, the component size distribution is a power-law in scale-free and non-scale-free growing networks. For comparison, if an uncorrelated network has a rapidly decreasing degree distribution, then the component size distribution slowly decreases only at the birth point of a giant component.

### Significance of Loops

We have emphasized that trees are simple for analysis. However, it is loops that make networks really interesting. In numerous studied random networks, in the infinite network limit, any finite environment of a vertex has no loops, i. e., it is tree-like, and the so-called *tree ansatz* is valid for many problems. It turns out, however, that it is infinite loops that are important for the global organization of networks. In sparse random recursive networks, there are usually few finite loops and many loops of length greater than a network diameter. These large loops may make a loopy network much more compact than a tree.

Let us compare a mean intervertex distance  $\bar{\ell}(N)$  in the two kinds of scale-free recursive networks: trees ( $m = 1$ ) and sparse networks with large loops ( $m > 1$ ), see Sects. “[The Barabási–Albert Model](#)” and “[General Preferential Attachment](#)”.

- (i) *Recursive trees.* For any  $\gamma > 2$ , asymptotically  $\bar{\ell}(N) \sim \ln N$  [45,46].
- (ii) *Sparse recursive networks with loops.*
  - (a) When  $\gamma > 3$ ,  $\bar{\ell}(N) \sim \ln N$ .
  - (b) When  $\gamma = 3$ ,  $\bar{\ell}(N) \sim \ln N / \ln \ln N$  [21].
  - (c) When  $2 < \gamma < 3$ , we expect  $\bar{\ell}(N) \sim \ln \ln N$  or, maybe, another slow dependence of this sort (in this range, we know only estimates for uncorrelated networks [47]).

Thus, the difference in  $\bar{\ell}(N)$  in these networks is strong if  $\gamma \leq 3$ .

### Accelerated Growth of Networks

In networks, which we discussed above, the number of edges  $L$  grows with time proportionally to (or linearly in) the total number of vertices  $N$ . Asymptotically,  $L/N \cong \text{const}$  in large networks – the growth is linear. A nonlinear growth, where the dependence  $L(N)$  is nonlinear at large  $N$ , is poorly studied. A growing network with a power-law  $L \propto N^{1+a}$ , where  $a < 1$ , was considered in [48]. Since the number  $a$  in real networks is usually positive, this type of growth is called *accelerated*. The diverse architectures of these networks are strongly model dependent. Their degree distributions may be nonstationary even in the limit  $N \rightarrow \infty$ . Some of these networks are scale-free, with exponent  $\gamma$  in a wide range from 1 to  $\infty$ . Cutoffs in these degree distributions strongly differ from those for linearly growing networks.

The accelerated growth essentially changes the form of the size dependence of the mean intervertex distance. Since these networks become more dense with  $N$ , the dependence  $\bar{\ell}(N)$  may be stationary or even decreasing. Numerous real-world examples of this effect are described in [49].

### Critique of the Preferential Attachment Mechanism

The preferential attachment mechanism and, more generally, self-organization models for networks have been criticized in [50]. The authors of that work stressed that these models are “only evocative” and “not explanatory”. An evocative model only “can reproduce the phenomenon of interest but does not necessarily capture and incorporate the true underlying cause”. In contrast, an explanatory model “also captures the causal mechanisms (why and how, in addition to what)”.

This critique was based on two arguments. First, it was demonstrated in [50] that the Barabási–Albert model fails to reproduce many empirical data for real networks. We believe that this part of the critique is not very strong.

Indeed, the Barabási–Albert model is only the simplest model in a wide class of models. There exist much more realistic models with preferential attachment. The second argument is essentially more serious. That is: the preferential attachment concept itself does not provide us with a reason for the preferential attachment. In this respect, the preferential attachment model falls short of explaining the genuine nature of complex network architectures. The authors of [50] suggested that optimization-based models explain complex networks in a more natural and convincing way. Unfortunately, the optimization concept of complex networks is much less developed than the preferential attachment. The difficulty is that typically, numerical optimization of networks is a heavy computational task. On the other hand, few optimization models for complex networks were studied analytically.

### Optimization-Based Models

In the optimization-based models, new connections optimize (e.g., minimize) some functional of network characteristics. Optimization may be global or local. In the global optimization, each new connection is made in such a way that it minimizes a function or functional of global network characteristics. For example, this *cost function* may be a linear function of the mean intervertex distance and other global characteristics of a network [51]. In this case, optimization leads to more compact architectures. The global optimization of a network structure is a time consuming computational problem. Only small networks were generated using these algorithms. Their sizes did not allow one to arrive at solid conclusions about their structures.

In the local optimization, for a new connection, we select “the best vertex” in a network. The variables of a cost function in the local optimization are characteristics of a vertex, e.g., a shortest path distance from some other vertex or vertices. For each selected vertex, the cost function is minimum among all vertices of the network.

A local optimization model proposed in [52] allows a strict analytical treatment, since it is a tree. The model is formulated in the following way. The growth starts from a single vertex (root with label 0) at some point of a restricted two-dimensional (which is not that important) area. At each time step,

- (i) place a new vertex  $t$  at a random point of this area and
- (ii) attach it to that existing vertex  $i$  where the cost function  $\ell_{0,i} + \alpha(t)d_{i,t}$  has the minimum value.

Here  $\ell_{0,i}$  is the shortest-path network distance between vertex  $i$  and the root – “the depth” of vertex  $i$ ;  $d_{i,t}$  is the

Euclidean distance between vertices  $t$  and  $i$  in this plane; the coefficient  $\alpha(t)$  is some given function of the network size  $t = N + 1$ . One can see that this model is based on “optimized trade-offs” between two conflicting objectives: a network objective and a geographic one – connect to the root (the center of a network) versus connect to the geographically closest vertex. It turns out that when  $\alpha(t)$  is a power law, the resulting degree distribution has a power-law region. Unfortunately, depending on a form of the function  $\alpha(t)$ , either this power-law region is quite narrow or the great majority of vertices are leaves [53]. So this model does not produce a really scale-free network in the limit  $N \rightarrow \infty$ .

Another local optimization model was described in [54]. This is also a tree. The growth starts from the root 0 at some point of a one-dimensional (which is important) interval. At each time step,

- (i) Place a new vertex  $t$  at a random point of this interval, and
- (ii) attach it to that existing vertex  $i$  where the cost function  $\ell_{0,i} + \alpha n_{i,t}$  has the minimum value.

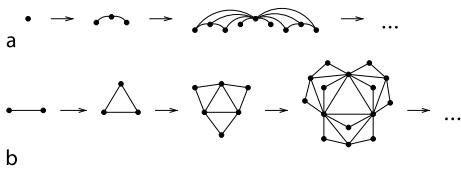
In this cost function,  $\ell_{0,i}$  is the shortest-path network distance between vertex  $i$  and the root;  $n_{i,t}$  is the number of already existing vertices between vertices  $t$  and  $i$  in this interval; and the coefficient  $\alpha$  is a constant. One can show that this very special optimization process actually generates preferential attachment. The resulting degree distribution has a power-law form with a cutoff depending only on the value of  $\alpha$ :  $q_{\text{cut}} \sim 1/\alpha$ .

### Deterministic Graphs

Deterministic growing graphs provide a wide range of complex architectures. These networks are quite simple: for each size, a deterministic network “ensemble” consists of a single member – a single graph. Nonetheless, one can build deterministic graphs with properties surprisingly similar to those of random networks. One can construct deterministic graphs with a small-world feature, deterministic scale-free graphs, where a discrete spectrum of degrees has a power-law envelope, and so on. Figure 2 shows two simple constructions proposed in Refs. [55,56] which are scale-free deterministic small worlds. For other deterministic graphs with a small-world property, see, e.g., Refs. [57,58].

### Future Directions

In this brief survey, we discussed only a few basic growth models. For example, we did not mention a mechanism



**Growth Models for Networks, Figure 2**

Two examples of scale-free deterministic graphs. Graphs **a** and **b** were introduced in [55] and [56], respectively

based on the merging of vertices [59]. This merging results in strongly connected hubs and power laws in networks. Moreover, we did not touch upon some important network characteristics. Among these missed characteristics, there are for example, the betweenness centrality and its distribution [60,61]. We considered only exponential and power-law degree distributions and omitted more general multifractal distributions. We did not discuss numerous applications of these models and empirical data. The reader should refer to books and reviews in the list of references for more detail.

Here we mention a few works where the growth models for networks were used for analysis and interpretation of empirical data. A link copying model of the World Wide Web was described in [62]. In this model links are directed. Each new vertex connects to one of the nearest neighbors of a randomly chosen vertex, i.e. “copies” the corresponding directed link. This process effectively generates preferential attachments and so leads to heavy-tailed degree distributions. The authors of work [16] applied the preferential attachment concept to the Internet both on the Autonomous System level and for the network of routers. Duplication of proteins also leads to preferential attachment and explains specific architectures of protein interaction networks [42], see also [63]. Several networks of scientific collaborations were empirically studied in the work [64]. With these data, the parameters of the preferential attachment process were found. One should note that collaboration networks may be treated as bipartite graphs where the vertices of the first type are collaborators and the vertices of the second type are acts of collaboration (scientific papers for example). The preferential attachment mechanism has been generalized to networks of this kind in [65].

Finally, we present a short list of research directions which, we believe, give a good perspective.

- (i) More realistic models for specific real-world networks. These models should not only reproduce basic network characteristics but also uncover the underlying mechanism.

- (ii) Optimization-based models. The goal is to find efficient optimization algorithms generating large networks with heavy-tailed degree distributions.
- (iii) Accounting for loops in growing networks. The aim is to find a way to compute the contribution of loops to the properties and characteristics of networks.
- (iv) Cooperative models on growing networks, flows, and processes taking place in networks. The problem is how the architectures of growing networks influence phenomena in interacting systems placed on these networks.
- (v) Co-evolving systems – networks and interacting agents on them. Interacting agents placed on the top of a network may also influence the structure of the network itself. How do these complex systems evolve? How are they organized?

## Acknowledgments

The author thanks A. V. Goltsev, J. F. F. Mendes, and A. N. Samukhin for numerous discussions. This work was partially supported by projects POCI: FAT/46241, MAT/46176, FIS/61665, and BIA-BCM/62662, and DYSONET.

## Bibliography

### Primary Literature

1. Flory PJ (1941) Molecular size distribution in three-dimensional polymers: I, II, III. *J Am Chem Soc* 63:83–100
2. Stockmayer WH (1943/1944) Theory of molecular size distribution and gel formation in branched chain polymers. *J Chem Phys* 11:45–55; 12:125–134
3. Na HS, Rapoport A (1970) Distribution of nodes of a tree by degree. *Math Biosci* 6:313–329
4. Moon JW (1974) The distance between nodes in recursive trees. *Lond Math Soc Lect Notes Ser* 13:125–132
5. Barabási A-L, Albert R (1999) Emergence of scaling in complex networks. *Science* 286:509–512
6. Mahmoud H (1991) Limiting distributions for path lengths in recursive trees. *Probab Eng Info Sci* 5:53–59
7. Erdős P, Rényi A (1959) On random graphs. *Publ Math Debrecen* 6:290–297
8. Dobrow RP (1996) On the distribution of distances in recursive trees. *J Appl Prob* 33:749–757
9. Bialas P, Burda Z, Jurkiewicz J, Krzywicki A (2003) Tree networks with causal structure. *Phys Rev E* 67:066106
10. Yule GU (1925) A mathematical theory of evolution based on the conclusions of Dr. JC Willis. *Phil Trans Royal Soc Lond B* 213:21–87
11. Simon HA (1955) On a class of skew distribution functions. *Biometrika* 42:425–440
12. Price DJ de S (1976) A general theory of bibliometric and other cumulative advantage processes. *J Amer Soc Inform Sci* 27:292–306

13. Albert R, Barabási A-L, Jeong H (1999) Mean-field theory for scale-free random networks. *Phys A* 272:173–187
14. Szymański J (1987) On a nonuniform random recursive trees. *Ann Discret Math* 33:297–306
15. Dorogovtsev SN, Mendes JFF, Samukhin AN (2001) Size-dependent degree distribution of a scale-free growing network. *Phys Rev E* 63:062101
16. Pastor-Satorras R, Vázquez A, Vespignani A (2001) Dynamical and correlation properties of the Internet. *Phys Rev Lett* 87:258701
17. Vázquez A, Pastor-Satorras R, Vespignani A (2002) Large-scale topological and dynamical properties of the Internet. *Phys Rev E* 65:066130
18. Bollobás B, Riordan O (2003) Mathematical results on scale-free graphs. In: Bornholdt S, Schuster HG (eds) *Handbook of Graphs and Networks*. Wiley, Weinheim, pp 1–34
19. Klemm K, Eguíluz VM (2002) Growing networks with small-world behavior. *Phys Rev E* 65:057102
20. Bianconi G, Capocci A (2003) Number of loops of size  $h$  in growing scale-free networks. *Phys Rev Lett* 90:078701
21. Bollobás B, Riordan OM (2004) The diameter of a scale-free random graph. *Combinatorica* 24:5–34
22. Dorogovtsev SN, Mendes JFF, Samukhin AN (2000) Exact solution of the Barabási-Albert model. *Phys Rev Lett* 85:4633–4636
23. Dorogovtsev SN, Mendes JFF (2003) *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press, Oxford
24. Krapivsky PL, Redner S (2001) Organization of growing random networks. *Phys Rev E* 63:066123
25. Newman MEJ (2002) Assortative mixing in networks. *Phys Rev Lett* 89:208701
26. Krapivsky PL, Redner S, Leyvraz F (2000) Connectivity of growing random networks. *Phys Rev Lett* 85:4629–4632
27. Bauke H, Sherrington D (2007) Topological phase transition in complex networks. [arXiv:0710.0831](https://arxiv.org/abs/0710.0831)
28. Ben-Naim E, Krapivsky PL (2007) Addition-deletion networks. *J Phys A* 40:8607–8619
29. Krapivsky PL, Redner S (2002) Finiteness and fluctuations in growing networks. *J Phys A* 35:9517–9534
30. Waclaw B, Sokolov IM (2007) Finite size effects in Barabási-Albert growing networks. *Phys Rev E* 75:056114
31. Bianconi G, Barabási A-L (2001) Competition and multiscaling in evolving networks. *Europhys Lett* 54:436–442
32. Ergun G, Rodgers GJ (2002) Growing random networks with fitness. *Phys A* 303:261–272
33. Dorogovtsev S, Mendes JFF (2000) Evolution of networks with aging of sites. *Phys Rev E* 62:1842–1845
34. Bianconi G, Barabási A-L (2001) Bose-Einstein condensation in complex networks. *Phys Rev Lett* 86:5632–5635
35. Barrat A, Barthélémy M, Pastor-Satorras R, Vespignani A (2004) The architecture of complex weighted networks. *Proc Natl Acad Sci* 101:3747–3752
36. Barrat A, Barthélémy M, Vespignani A (2004) Weighted evolving networks: Coupling topology and weights dynamics. *Phys Rev Lett* 92:228701
37. Dorogovtsev S, Mendes JFF (2005) Evolving weighted scale-free networks. *AIP Conf Proc* 776:29–36
38. Bianconi G (2005) Emergence of weight-topology correlations in complex scale-free networks. *Europhys Lett* 71:1029–1035
39. Callaway DS, Hopcroft JE, Kleinberg JM, Newman MEJ, Strogatz SH (2001) Are randomly grown graphs really random? *Phys Rev E* 64:041902
40. Dorogovtsev SN, Mendes JFF, Samukhin AN (2001) Anomalous percolation properties of growing networks. *Phys Rev E* 64:066110
41. Lancaster D (2002) Cluster growth in two growing network models. *J Phys A* 35:1179–1194
42. Kim J, Krapivsky PL, Kahng B, Redner S (2002) Infinite-order percolation and giant fluctuations in a protein interaction network. *Phys Rev E* 66:055101
43. Coulomb S, Bauer M (2003) Asymmetric evolving random networks. *Eur Phys J B* 35:377–389
44. Krapivsky PL, Derrida B (2004) Universal properties of growing networks. *Phys A* 340:714–724
45. Szabó G, Alava M, Kertész J (2002) Shortest paths and load scaling in scale-free trees. *Phys Rev E* 66:026101
46. Bollobás B, Riordan OM (2004) Shortest paths and load scaling in scale-free trees. *Phys Rev E* 69:036114
47. Cohen R, Havlin S (2003) Scale-free networks are ultra-small. *Phys Rev Lett* 90:058701
48. Dorogovtsev SN, Mendes JFF (2001) Effect of the accelerating growth of communications networks on their structure. *Phys Rev E* 63:025101(R)
49. Leskovec J, Kleinberg J, Faloutsos C (2007) Laws of graph evolution: Densification and shrinking diameters. *ACM TKDD* (1)2 physics/0603229
50. Willinger W, Govindan R, Jamin S, Paxson V, Shenker S (2002) Scaling phenomena in the Internet: Critically examining criticality. *Proc Natl Acad Sci* 99:2573–2580
51. Ferrer I, Cancho R, Sole RV (2003) Optimization in complex networks. In: Pastor-Satorras R, Rubí M, Diaz-Guilera A (eds) *Statistical Mechanics of Complex Networks*. Springer, Berlin, pp 114–125; cond-mat/0111222
52. Fabrikant A, Koutsoupias E, Papadimitriou CH (2002) Heuristically optimized trade-offs: A new paradigm for power laws in the Internet. *Lecture Notes in Computer Science*, vol 2380. Springer, Berlin, pp 110–122
53. Berger N, Bollobás B, Borgs C, Chayes J, Riordan O (2003) Degree distribution of the FKP network model. *Lecture Notes in Computer Science*, vol 2719. Springer, Berlin, pp 725–738
54. D’Souza RM, Borgs C, Chayes JT, Berger N, Kleinberg RD (2007) Emergence of tempered preferential attachment from optimization. *Proc Natl Acad Sci* 104:6112–6117
55. Barabási A-L, Ravasz E, Vicsek T (2001) Deterministic scale-free networks. *Phys A* 299:559–564
56. Dorogovtsev SN, Goltsev AV, Mendes JFF (2002a) Pseudofractal scale-free web. *Phys Rev E* 65:066122
57. Jung S, Kim S, Kahng B (2002) Geometric fractal growth model for scale-free networks. *Phys Rev E* 65:056101
58. Andrade JS Jr, Herrmann HJ, Andrade RFS, da Silva LR (2005) Apollonian networks. *Phys Rev Lett* 94:018702
59. Kim BJ, Trusina A, Minnhagen P, Sneppen K (2005) Self organized scale-free networks from merging and regeneration. *Eur Phys J B* 43:369–372
60. Goh KI, Kahng B, Kim D (2001) Universal behavior of load distribution in scale-free networks. *Phys Rev Lett* 87:278701–278704
61. Goh KI, Oh E, Jeong H, Kahng B, Kim D (2002) Classification of scale-free networks. *Proc Natl Acad Sci* 99:12583–12588
62. Kleinberg JM, Kumar R, and Raghavan P, Rajagopalan S, Tomkins AS (1999) The Web as a graph: Measurements, mod-

- els and methods. Lecture Notes in Computer Science, vol 1627. Springer, Berlin, pp 1–17
63. Colizza V, Flammini A, Maritan A, Vespignani A (2005) Characterization and modeling of protein-protein interaction networks. *Phys A* 352:1–27
  64. Barabási A-L, Jeong H, Neda Z, Ravasz E, Schubert A, Vicsek T (2002) Evolution of the social network of scientific collaborations. *Phys A* 311:590–614
  65. Ramasco JJ, Dorogovtsev SN, Pastor-Satorras R (2004) Self-organization of collaboration networks. *Phys Rev E* 70:036106

## Books and Reviews

- Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74:47–97
- Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D-U (2006) Complex networks: Structure and dynamics. *Phys Rep* 424:175–308
- Caldarelli G (2007) Scale-Free Networks: Complex Webs in Nature and Technology. Oxford Finance Series. Oxford University Press, Oxford
- Cohen R, Havlin S, Ben-Avraham D (2003) Structural properties of scale free networks. In: Bornholdt S, Schuster HG (eds) *Handbook of Graphs and Networks*. Wiley, Weinheim, pp 85–110
- Dorogovtsev SN, Mendes JFF (2002) Evolution of networks. *Adv Phys* 51:1079–1187
- Dorogovtsev SN, Goltsev AV, Mendes JFF (2008) Critical phenomena in complex networks. *Rev Mod Phys* 80(3) [arXiv:0705.0010](https://arxiv.org/abs/0705.0010)
- Durrett R (2006) Random Graph Dynamics. Cambridge University Press, Cambridge
- Kim J, Krapivsky PL, Kahng B, Redner S (2002) Infinite-order percolation and giant fluctuations in a protein interaction network. *Phys Rev E* 66:055101
- Newman MEJ (2003) The structure and function of complex networks. *SIAM Review* 45:167–256
- Pastor-Satorras R, Vespignani A (2004) Evolution and Structure of the Internet: A Statistical Physics Approach. Cambridge University Press, Cambridge
- Pastor-Satorras R, Vazquez A, Vespignani A (2001) Dynamical and correlation properties of the Internet. *Phys Rev Lett* 87:258701–258704
- Smythe RT, Mahmoud HM (1995) A survey of recursive trees. *Theor Prob Math Statist* 51:1–27

---

## Growth Phenomena in Cellular Automata

JANKO GRAVNER

Mathematics Department, University of California,  
Davis, USA

### Article Outline

- [Glossary](#)
- [Definition of the Subject](#)
- [Introduction](#)
- [Final Set](#)

- [Asymptotic Shapes](#)
- [Nucleation](#)
- [Future Directions](#)
- [Bibliography](#)

## Glossary

**Asymptotic density** The proportion of sites in a lattice occupied by a specified subset is called asymptotic density, or, in short, density.

**Asymptotic shape** The shape of a growing set, viewed from a sufficient distance so that the boundary fluctuations, holes, and other lower order details disappear, is called the asymptotic shape.

**Cellular automaton** A cellular automaton is a sequence of configurations on a lattice which proceeds by iterative applications of a homogeneous local update rule. A configuration attaches a state to every member (also termed a site or a cell) of the lattice. Only configurations with two states, coded 0 and 1, are considered here. Any such configuration is identified with its set of 1's.

**Final set** A site whose state changes only finitely many times is said to fixate, or attain a final state. If this happens for every site, then the sites whose final states are 1 comprise the final set.

**Initial set** A starting set for a cellular automaton evolution is called initial set, and may be deterministic or random.

**Metastability** Metastability refers to a long, but finite, time period in an evolution of a cellular automaton rule, during which the behavior of the iterates has identifiable characteristics.

**Monotone cellular automaton** A cellular automaton is monotone if addition of 1's to the initial configuration always results in more 1's in any subsequent configuration.

**Nucleation** Nucleation refers to (usually small) pockets of activity, often termed nuclei, with long range consequences.

**Solidification** A cellular automaton solidifies if any site which achieves state 1 remains forever in this state.

### Definition of the Subject

In essence, analysis of growth models is an attempt to study properties of physical systems far from equilibrium (e.g., [52] and its more than 1300 references). Cellular automata (CA) growth models, by virtue of their simplicity and amenability to computer experimentation [59,66], have become particularly popular in the last

30 years in many fields, such as physics [15, 59, 60], biology [18], chemistry [15, 50], social sciences [12], and artificial life [51]. In contrast to voluminous empirical literature on CA in general and their growth properties in particular, precise mathematical results are rather scarce. A general CA theory is out of the question, since a Turing machine can be embedded in a CA, so that examples as “simple” as elementary one-dimensional CA [17] and Conway’s Game of Life [7] are capable of universal computation. Even the most basic parameterized families of CA systems exhibit a bewildering variety of phenomena: self-organization, metastability, turbulence, self-similarity, and so forth [1, 22, 24, 43]. From a mathematical point of view, CA can be rightly viewed as discrete counterparts to partial differential equations, and so they are able to emulate many aspects of the physical world, while at the same time they are easy to experiment using widely available platforms (from many available simulation programs we mention just [66]).

Despite their resistance to traditional methods of deductive analysis, CA have been of interest to mathematicians from their inception and we will focus on the rigorous mathematical results about their growth properties. The scope will be limited to CA with deterministic update rule – random rules are more widely used in applications [12, 18, 50], but fit more properly within probability theory (however, see e. g. [37] for a connection between deterministic and random aspects of CA growth). Adding randomness to the rule in fact often makes them more tractable, as ergodic properties of random systems are much better understood than those of deterministic ones ([10] provides good examples).

Even though mathematical arguments are the ultimate objective, computer simulations are an indispensable tool for providing the all important initial clues for the subsequent analysis. However, as we explain in a few examples in the sequel, caution needs to be exercised while making predictions based on simulations. Despite the increased memory and speed of commercial computers, for some CA rules highly relevant events occur on spatial and temporal scales far beyond the present-day hardware. Simply put, mathematics and computers are both important, and one ignores each ingredient at one’s own peril.

By nature, a subject in the middle of active research contains many exciting unresolved conjectures, vague ideas that need to be made precise, and intriguing examples in search of proper mathematical techniques. It is clear from what has already been accomplished that, often in sharp contrast with the simplicity of the initially posed problem, such techniques may be surprising, sophisticated, and drawn from such diverse areas as combi-

natorics, geometry, probability, number theory, and PDE. This is a field to which mathematicians of all stripes should feel invited.

## Introduction

Let us begin with the general set-up. We will consider exclusively binary CA. Accordingly, a *configuration* will be a member of  $\{0, 1\}^{\mathbb{Z}^d}$ , that is, an assignment of 0 or 1 to every site in the  $d$ -dimensional lattice  $\mathbb{Z}^d$ . This divides the lattice into two sets, those that are assigned state 1, called the *occupied* sites, and those in state 0, which are the *empty* sites. A configuration is thus represented by its occupied set  $A$ . This set will change in discrete time, its evolution given by  $A_0, A_1, A_2, \dots \subset \mathbb{Z}^d$ .

The configuration changes subject to a CA rule, which is, in general, specified by the following two ingredients. The first is a finite *neighborhood*  $\mathcal{N} \subset \mathbb{Z}^d$  of the origin, its translate  $x + \mathcal{N}$  then being the neighborhood of point  $x$ . By convention, we assume that  $\mathcal{N}$  contains the origin. Typically,  $\mathcal{N} = B_\nu(0, \rho) \cap \mathbb{Z}^d$ , where  $B_\nu(0, \rho) = \{x \in \mathbb{R}^d : \|x\|_\nu \leq \rho\}$  is the ball in the  $\ell^\nu$ -norm  $\|\cdot\|_\nu$  and  $\rho$  is the *range*. When  $\nu = 1$  the resulting  $\mathcal{N}$  is called the Diamond neighborhood, while if  $\nu = \infty$  it is referred to as the Box neighborhood. (In particular, when  $d = 2$ , range 1 Diamond and Box neighborhoods are also known as von Neumann and Moore neighborhoods, respectively.) The second ingredient is a map  $\pi : 2^{\mathcal{N}} \rightarrow \{0, 1\}$ , which flags the *sufficient configurations* for occupancy. More precisely, for a set  $A \subset \mathbb{Z}^d$ , we let  $\mathcal{T}(A) \subset \mathbb{Z}^d$  consist of every  $x \in \mathbb{Z}^d$  for which  $\pi((A_t - x) \cap \mathcal{N}) = 1$ . Then, for a given *initial subset*  $A_0 \subset \mathbb{Z}^d$  of occupied points, we define  $A_1, A_2, \dots$  recursively by  $A_{t+1} = \mathcal{T}(A_t)$ .

To explain this notation on arguably the most famous CA of all time, the Game of Life [7, 26] has  $d = 2$ , Moore neighborhood  $\mathcal{N}$  consisting of the origin and nearest eight sites, so that the neighbor of  $x$  is

$$x + \mathcal{N} = \begin{array}{c} \bullet & \bullet & \bullet \\ \bullet & x & \bullet \\ \bullet & \bullet & \bullet \end{array},$$

and  $\pi(S) = 1$  precisely when either  $0 \in S$  and  $|S| \in \{3, 4\}$ , or  $0 \notin S$  and  $|S| = 3$ . Here,  $|S|$  is the size (cardinality) of  $S \subset \mathcal{N}$  and note that the center  $x$  of the neighborhood itself is counted in the occupation number.

Usually, our starting set  $A_0$  will consist of a possibly large, but finite set of 1’s surrounded by 0’s. However, other initial states are worthy of consideration, for example, half-spaces, wedges, and sets with finite complements, called *holes*. Finally, for understanding self-organizational abilities and statistical tendencies of the CA rule, the most

natural starting set is the random “soup”  $\Pi(p)$  to which every site is adjoined independently with probability  $p$ .

As already mentioned, we need to consider special classes if we hope to formulate a general theorem. Mathematically, the most significant restriction is to the class of *monotone* (or *attractive*) CA rules, for which  $S_1 \subset S_2$  implies  $\pi(S_1) \leq \pi(S_2)$ . To avoid the trivial case we will also assume that monotone CA have  $\pi(\mathcal{N}) = 1$ .

Another important notion is that of solidification: we say that the CA *solidifies* if  $\pi(S) = 1$  whenever  $0 \in S$ . In words, this means that once a site becomes occupied, it cannot be removed. To every CA on  $\mathbb{Z}^d$  given by the rule  $(\mathcal{N}, \pi)$  one can associate “space-time” solidification CA on  $\mathbb{Z}^d \times \mathbb{Z}$ , with unique solidification rule given by the neighborhood  $\mathcal{N}' = (\mathcal{N} \times \{-1\}) \cup \{0^{d+1}\}$ , and  $\pi'$  such that  $\pi'(S \times \{-1\}) = \pi(S)$  for  $S \subset \mathcal{N}$ . This construction is useful particularly for one-dimensional CA whose space-time version interprets their evolution as a two-dimensional object [64], but we prefer to focus on the growth phenomena in the rule’s “native” space.

A more restrictive, but still quite rich, class of rules is the *Threshold Growth* (TG) CA, which is a general *totalistic* monotone solidification CA rule. For such rules,  $\pi(S)$  depends only on the cardinality  $|S|$  of  $S$  whenever  $0 \notin S$ ; therefore, for such  $S$  there exists a *threshold*  $\theta \geq 0$  such that  $\pi(S) = 0$  whenever  $|S| < \theta$  and  $\pi(S) = 1$  whenever  $|S| \geq \theta$ .

We will universally assume that a 1 cannot spontaneously appear in a sea of 0’s, that is, that 1’s only grow by contact:  $\pi(\emptyset) = 0$ . We also find it convenient to assume that  $\pi$  is *symmetric*:  $-\mathcal{N} = \mathcal{N}$  and  $\pi(-S) = \pi(S)$ . This is not a necessary assumption in many contexts, but its absence makes many statements unnecessarily awkward.

Next is a very brief historical lesson. The first paper in CA modeling is surely [62], a precursor to the research into nucleation and self-organization in CA. The follow-up to this pioneering work had to wait until the 70’s, when the influential work [42] appeared. The earliest work on CA growth is that of S. Willson [63,64,65], which still stands today as one of the notable achievements of mathematical theory. The importance of growth properties of CA, from theoretical and modeling perspectives, was more widely recognized in the mid-80’s [53,54,59]. At about the same time, statistical physicists recognized the value of mathematical arguments in studying nucleation and metastability and hence the need to build tractable models [60,61]. Bootstrap percolation ([2,4,67], and references therein), one of the most studied CA, which we discuss in some detail in Sect. “**Nucleation**”, originates from that period. Since the beginning of the 90’s there has been a great expansion in the popularity of CA modeling [18,50], while

mathematical theory, which we review in the next three sections, proceeds at a much more measured pace.

The rest of the article is organized as follows. In Sect. “**Final Set**” we consider properties of the set which the CA rule generates “at the end of time.” In particular, we discuss when the CA eventually occupies the entire available space and, when it fails to do so, what proportion of space it does fill. Section “**Asymptotic Shapes**” then focuses on the occupation mechanism, in particular on shapes attained from finite seeds. The main theme of Sect. “**Nucleation**” are sparse randomly populated initializations. We conclude with Sect. “**Future Directions**”, a summary of issues in need of further research.

## Final Set

Perhaps the most basic question that one may ask is: what proportion of space does a CA rule ultimately fill? Clearly we need to specify more precisely what is meant by this, but it should be immediately suspected that the answer in general depends on the initial state, even if we only restrict to finite ones. Indeed, consider the TG CA with Moore neighborhood and  $\theta = 3$ . It is easy to construct an initial set which stops growing, say, one containing fewer than 3 sites. It is not much harder to convince oneself that there exist finite sets (even some with only 3 sites) which eventually make every site occupied. It is a combinatorial exercise to show that these two are the only possibilities in this example. Is this dichotomy valid in any generality? This is one of the questions we address in this section.

Assume a fixed CA rule and the associated transformation  $\mathcal{T}$ , and fix an initial state  $A_0$ . If every  $x \in \mathbb{Z}^d$  *fixates*, that is, changes state only finitely many times, then the *final set*  $A_\infty = \mathcal{T}^\infty(A_0)$  exists. Notice that this is automatically true for every solidification rule, in which no site can change state more than once.

We say that  $A_0$  *fills space* if  $\mathcal{T}^\infty(A_0) = \mathbb{Z}^d$ . One cannot imagine a greater ability of a CA rule to “conquer” the environment than if a finite set is able to fill space. Thus it is natural to ask whether there exist general conditions that assure this property, and indeed they do for monotone CA.

Induced by  $\mathcal{T}$  is a growth transformation  $\tilde{\mathcal{T}}$  on closed subsets of  $\mathbb{R}^d$ , given by

$$\tilde{\mathcal{T}}(B) = \{x \in \mathbb{R}^d : 0 \in \mathcal{T}((B - x) \cap \mathbb{Z}^d)\}.$$

In words, one translates the lattice so that  $x \in \mathbb{R}^d$  is at the origin, and applies  $\mathcal{T}$  to the intersection of Euclidean set  $B$  with the translated lattice. It is easy to verify that the two transformations are *conjugate*,

$$\mathcal{T}(B \cap \mathbb{Z}^d) = \tilde{\mathcal{T}}(B) \cap \mathbb{Z}^d.$$

It will become immediately apparent why  $\tilde{T}$  is convenient. Let  $S^{d-1}$  be the set of unit vectors  $\mathbb{R}^d$  in and let

$$H_u^- = \{x \in \mathbb{R}^d : \langle x, u \rangle \leq 0\}$$

be the closed half-space with unit outward normal  $u \in S^{d-1}$ . Then, *provided that the CA rule is monotone*, there exists a  $w(u) \in \mathbb{R}$  so that

$$\tilde{T}(H_u^-) = H_u^- + w(u) \cdot u$$

and consequently

$$T^t(H_u^- \cap \mathbb{Z}^d) = (H_u^- + tw(u) \cdot u) \cap \mathbb{Z}^d.$$

Monotone CA with  $w(u) > 0$  for every  $u$  are called *supercritical*. A supercritical CA hence enlarges every half-space.

**Theorem 1** *Assume a monotone CA rule. A finite set  $A_0$  which fills space exists if and only if  $w(u) > 0$  for every direction  $u \in S^{d-1}$ .*

See [30,63] for a proof. Before we proceed, a few remarks are in order. First, we should note that one direction of the above theorem has a one-line proof: if  $w(u) \leq 0$  for some  $u$ , then monotonicity prevents the CA from ever occupying a point outside a suitable translate of  $H_u^-$ . The other direction is proved by constructing a sufficiently “smooth” initial set. Moreover, supercriticality can be checked on a finite number of directions, in particular one can prove that a two-dimensional TG CA is supercritical if and only if  $\theta \leq \frac{1}{2}(|\mathcal{N}| - \max\{|\mathcal{N} \cap \ell| : \ell \text{ a line through } 0\})$  [30]. Thus, among the TG CA with Moore neighborhood, exactly those with  $\theta \leq 3$  are supercritical, while this is true for range 2 Box neighborhood when  $\theta \leq 10$ .

A finite set  $A_0$  for which  $\cup_t A_t$  is infinite is said to *generate persistent growth*. Further, a CA for which any set that generates persistent growth has  $A_\infty = \mathbb{Z}^d$  is called *omnivorous* [30]. For an omnivorous rule a finite seed has either a bounded effect or it fills space.

Is every supercritical TG CA omnivorous? The answer is no, and a counterexample in  $d = 2$  is obtained by taking the neighborhood to be the cross of radius 2:  $\mathcal{N} = \{(0,0), (0,\pm 1), (0,\pm 2), (\pm 1,0), (\pm 2,0)\}$ , and  $\theta = 2$ . It is easy to check that for  $A_0 = \{(0,0), (1,0)\}$  the final set  $A_\infty$  consists of the  $x$ -axis, while initialization with a  $2 \times 2$  box results in  $A_\infty = \mathbb{Z}^2$ . On the other hand, the following theorem holds.

**Theorem 2** *The two-dimensional TG CA is omnivorous provided either of the two conditions are satisfied:*

(1)  $\mathcal{N}$  is box neighborhood of arbitrary range.

- (2)  $\mathcal{N} = \bar{\mathcal{N}} \cap \mathbb{Z}^2$ , where  $\bar{\mathcal{N}}$  is a convex set with the same symmetries as  $\mathbb{Z}^2$ , and  $\theta \leq \sigma^2/2$ , where  $\sigma$  is the range of the largest box neighborhood contained in  $\mathcal{N}$ .

The theorem is proved in [8] and [9] by rather delicate combinatorial arguments involving analysis of invariant, or nearly invariant, quantities. The lack of robust methods makes conditions in the theorem far from necessary. In particular, proving a general analogue of Theorem 2 without solidification (while keeping monotonicity) is an intriguing open problem.

For non-monotone solidification rules, any general theory appears impossible, but one can analyze specific examples, and we list some recent results below. All are two-dimensional, therefore we assume  $d = 2$  for the rest of this section.

In many interesting cases, it is immediately clear from computer simulations that  $A_\infty \neq \mathbb{Z}^d$ , but at least  $A_\infty$  is spread out fairly evenly. This motivates the following definition. Pick a set  $A \subset \mathbb{Z}^2$ . Let  $\mu_\epsilon$  be  $\epsilon^2$  times the counting measure on  $\epsilon \cdot A$ . We say that  $A$  has *asymptotic density*  $\rho$  if  $\mu_\epsilon$  converges to  $\rho \cdot \lambda$  as  $\epsilon \rightarrow 0$ . Here  $\lambda$  is Lebesgue measure on  $\mathbb{R}^2$  and the convergence holds in the usual sense:

$$\int f d\mu_\epsilon \rightarrow \rho \cdot \int f d\lambda \quad (1)$$

for any  $f \in C_c(\mathbb{R}^2)$ . Equivalently, for any square  $R \subset \mathbb{R}^2$ , the quantity  $\epsilon^2 \cdot |R \cap (\epsilon \cdot A)|$  converges to the area of  $R$  as  $\epsilon \rightarrow 0$ .

For totalistic solidification CA, the rule is determined by the neighborhood and a *solidification list* of neighborhood counts which result in occupation at the next time. Three neighborhoods have been studied so far: *Diamond* rules with von Neumann neighborhood, *Box* rules with Moore neighborhood and *Hex* rules with the neighborhood  $\mathcal{N}$  consisting of  $(0,0)$  and the six sites  $(\pm 1,0)$ ,  $(0,\pm 1)$ , and  $\pm(1,1)$ . (We note that this last neighborhood is a convenient way to represent the triangular lattice [59].) These rules are often referred to as *Packard snowflakes* [11,36,53]. As an example, in *Hex 135* rule, a 0 turns into a 1 exactly when it “sees” an odd number of already occupied neighbors in its hexagonal neighborhood.

We will assume that 1 is on the solidification list, for otherwise the analysis quickly becomes too difficult (see however [33] and [45] for some results on *Box 2* and *Box 3*) rules. Further, for *Hex* and *Diamond* cases, we will assume 2 is not on this list (or else the dynamics is too similar to a TG CA). We now summarize the main results of [36] and [11].

**Theorem 3** *To each of the four Diamond and 16 Hex Packard snowflakes there corresponds a number  $\rho \in (0, 1]$ ,*

the asymptotic density of  $A_\infty$ , which is independent of the finite seed  $A_0$ . The densities in Diamond cases are

$$\rho_1 = 2/3, \rho_{13} = 2/3, \rho_{14} = 1, \rho_{134} = 29/36.$$

The Hex densities are exactly computable in 8 cases:

$$\begin{aligned} \rho_{13} &= \rho_{135} = 5/6, \rho_{134} = \rho_{1345} = 21/22, \rho_{136} = \rho_{1356} \\ &= \rho_{1346} = \rho_{13456} = 1. \end{aligned}$$

In six other Hex cases, one can estimate, within  $\pm 0.0008$ ,

$$\begin{aligned} \rho_1 &\approx 0.6353, \rho_{14}, \rho_{145} \approx 0.9689, \rho_{15} \approx 0.8026, \\ \rho_{16} &\approx 0.7396, \rho_{156} \approx 0.9378. \end{aligned}$$

The final two Hex rules have densities very close to 1:

$$\rho_{146} \in (0.995, 1), \rho_{1456} \in (0.9999994, 1).$$

The indices in the densities of course refer to the respective rule.

Note that, in each of the two cases,  $\rho_{14} > \rho_{134}$ , testimony to the fundamentally nonmonotone nature of these rules. It is also shown in [36] that observing Hex 1456 from  $A_0 = \{0\}$  on even the world's most extensive graphics array, with millions of sites on a side, one would still be led to the conclusion that  $A_\infty = \mathbb{Z}^2$ . In fact, the site in  $A_\infty^c$  closest to the origin is still at distance of the order  $10^9$ . Nevertheless,  $A_\infty^c$  has a positive density and contains arbitrarily large islands of 0's. This is one illustration of limitations in making empirical conclusions based on simulations.

The fundamental tool to prove the above theorem is the fact that these dynamics have an additive component which creates an impenetrable *web* of occupied sites [36]. This web consists of sites at the edge of the light cone, or, to be more precise, the sites which are occupied at the same time at which TG CA with the same neighborhood and  $\theta = 1$  would occupy them.

The web makes at least an approximate recursion possible, and the basic renewal theory applies. The delicacy of such results is conveyed effectively by comparison to Box solidification. There are 128 such rules with 1 on the solidification list. Although snowflake-like recursive carpets emerge in a great many cases, and exact computations are sometimes feasible, there is no hope of a complete analysis as in the Hex and Diamond settings, and many fascinating problems remain. For instance, the density of Box 1, provided it exists at all, can depend on the initial seed. Namely, it is shown in [33] that Box 1 solidification yields density 4/9 starting from a singleton. Later, D. Hickerson (private communication) engineered finite initial seeds with asymptotic densities 29/64 and 61/128.

**Growth Phenomena in Cellular Automata, Table 1**  
Densities of  $A_\infty$  from  $A_0 = \{0\}$  for some Box rules

Rule	Density
12	2/3
13	28/45
15	43/72
16	385/576
17	35/72
18	4/9

The latter is achieved by an ingenious arrangement of 180 carefully placed occupied cells around the boundary of an  $83 \times 83$  grid. The highest density with which Box 1 solidification can fill the plane is not known, and neither is whether any seed fills with density less than 4/9. Most initial seeds generate what seems to be a chaotic growth with density about 0.51.

Many other Box rules have known asymptotic densities started from a singleton. Table 1 is a sample (D. Griffeath, private communication).

All exact density computations presented in this section are based on explicit recursions, made possible by an additive web. These recursions are in some cases far from simple, for example, D. Griffeath has shown that in the Box 12 case, the following formula holds for  $a_n = |A_\infty \cap B_\infty(0, 2^n - 1)|$ ,  $n \geq 12$ :

$$\begin{aligned} a_n = \frac{8}{3} \cdot 4^n + r_1 \gamma_3^n - \frac{8}{3} \cdot 3^n - \frac{16}{15} \cdot 2^n + \frac{2}{51} \cdot (-2)^n + 4n - 3 \\ + \frac{8}{3} \cdot (-1)^n + r_2 \gamma_1^n + r_3 \cdot \gamma_2^n, \end{aligned}$$

where

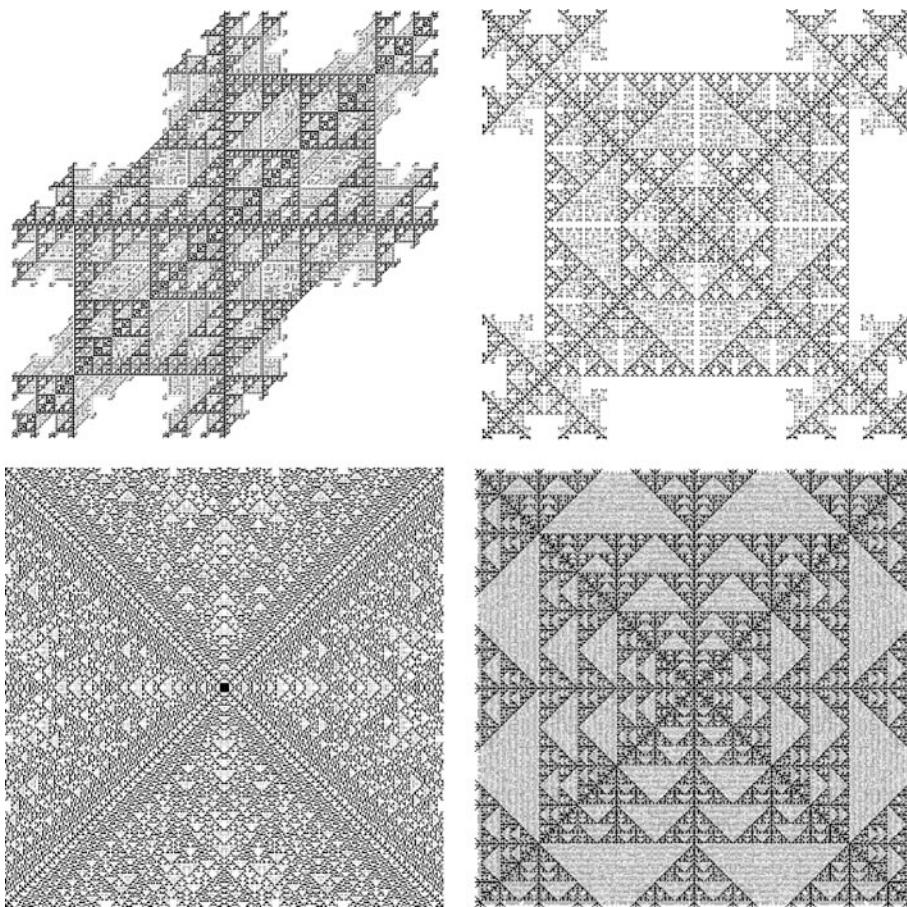
$$\gamma_1 \approx -0.675, \gamma_2 \approx 0.461, \gamma_3 \approx 3.214$$

are the three real roots of the equation  $\gamma^3 - 3\gamma^2 - \gamma + 1 = 0$ , while

$$r_1 \approx -6.614, r_2 \approx -2.126, r_3 \approx 2.434$$

$$\text{solve } 3145r^3 + 19832r^2 - 22688r - 107648 = 0.$$

Apparently very similar rules to those in the above table seem unsolvable, such as Box 14, and the “odd” rule Box 1357 which does have an additive component, but the resulting web from  $A_0 = \{0\}$  “leaks” and growth is apparently chaotic. The same problem plagues almost all Box rules started from general initial set. The sole exception seems to be the 12 rule, the best candidate for a general general theorem among the 128 rules, due to its quasiadditive web [49]. We should also mention that embedded additive dynamics have been used to study other models [23].



**Growth Phenomena in Cellular Automata, Figure 1**

Some Packard snowflakes. Clockwise from top left: Hex 1; Box 1; Box 1357; and again Box 1. First three are started from  $\{0\}$ , and the last from an  $8 \times 8$  box. The web is black, otherwise the updates are periodically shaded. Note that the chaotic growth can result from a chaotic web (bottom left) or from a leaky web (bottom right)

In all considered cases, the web consist of several copies of the final set generated by the space-time solidification associated to a one-dimensional CA. When this CA is linear, the web's fractal dimension can be computed using the method from [65]. For example, the properly scaled webs in the top two frames on Fig. 2 approach a set with Hausdorff dimension  $\log 3 / \log 2$ , while for the bottom right web this dimension is  $\log(1 + \sqrt{5}) / \log 2$ .

Given that all exactly given densities so far are rational, a natural question is whether there is an example of  $A_\infty$  with irrational density. Such example was given by Griffeath and Hickerson in [44], where an initial state for the *Game of Life* is provided for which the set  $t^{-1}A_t$  converges to an asymptotic density  $(3 - \sqrt{5})/90$  on an appropriate finite set  $L$ . This formulation masks the fact that every site  $x$  eventually periodically changes its state, so  $A_\infty$  does not exist. However, a closer look at the construction shows

that the final periods are uniformly bounded. Therefore, if  $p$  is the lowest common multiple of all final periods, the  $p$ 'th iterate of the *Game of Life* rule will generate  $A_\infty$  from the same  $A_0$  and with the same density.

This is the only known example of a computable irrational density, and there is a good reason, which we now explain, why such examples are difficult to come by.

By analogy with statistical physics, we would call a set  $A \subset \mathbb{Z}^2$  *exactly solvable*, if there exists a formula which decides whether a given  $x$  is an element of  $A$ . More formally, we require that there exists a finite automaton which, upon encountering  $x$  as input, decides whether  $x \in A$ . Representation of  $x$  as input is given as  $(\pm i_1^1, \pm i_1^2, i_2^1, i_2^2, \dots)$ , where  $i_1^1, i_1^2$  are the most significant binary digits of the first and second coordinate of  $x$ ;  $i_2^1, i_2^2$  the next most significant, etc. (Some initial  $i_k^1$ 's or  $i_k^2$ 's may be 0, and the representation is finite but of arbitrary length.) This means that  $A$  is *auto-*

matic [5], or equivalently a *uniform tag system* [16]. With a slight abuse of terminology we call a solidification CA *exactly solvable* (from  $A_0$ ) if  $A_\infty$  is exactly solvable.

To our knowledge, the simplest nontrivial example of an exactly solvable CA is *Diamond 1* solidification, for which it can be shown by induction that  $x \notin A_\infty$  iff  $\max\{k: i_k^1 = 1\} = \max\{k: i_k^2 = 1\}$ . It is easy to construct a (two-state) finite automaton that checks this condition, and the density  $\rho$  of  $A_\infty$  evidently must satisfy the equation  $\rho = 1/2 + \rho/4$ , so that  $\rho = 2/3$  as stated in Theorem 3. In fact, all of the CA in Theorem 3 with exactly given densities are exactly solvable, and then, by [16], Theorem 6, these densities must be rational. Therefore, the Griffeath–Hickerson example given above is *not* exactly solvable, and the mechanism that forms  $A_\infty$  must be more complex in this precise sense. We note that none of the other examples from Theorem 3 are exactly solvable either, but for a different reason [36].

This section’s final example, like many other fascinating CA rules, is due to D. Hickerson (private communication). His *Diamoeba* is a rule with the Moore neighborhood and  $\pi(S) = 1$  whenever one of the following two conditions is satisfied:

$$\begin{aligned} 0 \notin S, \text{ and } |S| \in \{3, 5, 6, 7, 8\}, \text{ or} \\ 0 \in S, \text{ and } |S| \in \{6, 7, 8, 9\}. \end{aligned}$$

This would be an easily analyzed monotone rule if the 3 were replaced by a 9, with  $A_\infty = \emptyset$  for every finite  $A_0$ . At first, it seems that the Diamoeba shares this fate. In fact, D. Hickerson has demonstrated that, starting from  $A_0 = B_\infty(0, r) \cap \mathbb{Z}^2$ ,  $A_t = \emptyset$  at the smallest  $t$  given by

$$12r - 8 - 4r_1 + r_{11} + (r \bmod 2),$$

where  $r_1$  and  $r_{11}$  are, respectively, the number of 1’s and the number of 11’s in the binary representation of  $r$ . This interesting formula only gives a small taste of things to come (see [33] for a detailed discussion). One of the most intriguing examples is when  $A_0$  is a  $2 \times 59$  rectangle with

a corner cell removed. This grows to a fairly large set in about a million updates, then apparently stops for several million more, after which another growth spurt is possible. The question whether  $A_\infty = \mathbb{Z}^2$  for this  $A_0$  is tantalizingly left open. However, there does exist an  $A_0$  for which  $A_\infty = \mathbb{Z}^2$ . This initialization was discovered by D. Bell, and is an adaptation of a spaceship found by a search algorithm designed by D. Eppstein [21]. This startling object attests to the remarkable design expertise that *Game of Life* researchers have developed over the years.

### Asymptotic Shapes

After addressing a rule’s *ability* to grow in the previous section, we now turn to the *geometry* of growth: is it possible to predict the shape that the set of 1’s attains as it spreads? It turns out that the complete answer is known in the monotone case.

Naturally, we need a notion of convergence of sets, and the most natural definition is due to Hausdorff (see [29,31] for an introduction to such issues). We say that a sequence of compact sets  $K_n \subset \mathbb{R}^d$  converges to a compact set  $K \subset \mathbb{R}^d$  (in short,  $K_n \rightarrow K$ ) if, for every  $\epsilon > 0$ ,  $K_n \subset K + B_2(0, \epsilon)$  and  $K \subset K_n + B_2(0, \epsilon)$ , for  $n$  large enough. Then we say that a CA has a *linear asymptotic shape*  $L$  from a finite initial seed  $A_0$  if

$$\frac{1}{t}A_t \rightarrow L$$

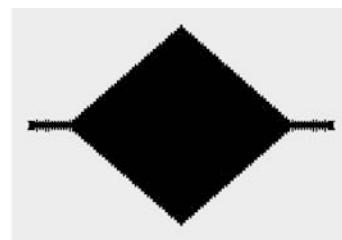
as  $t \rightarrow \infty$ .

Turning to monotone CA, we recall the definition of half-space velocities  $w$ , and set

$$K_{1/w} = \cup\{[0, 1/w(u)] \cdot u : u \in S^{d-1}\}$$

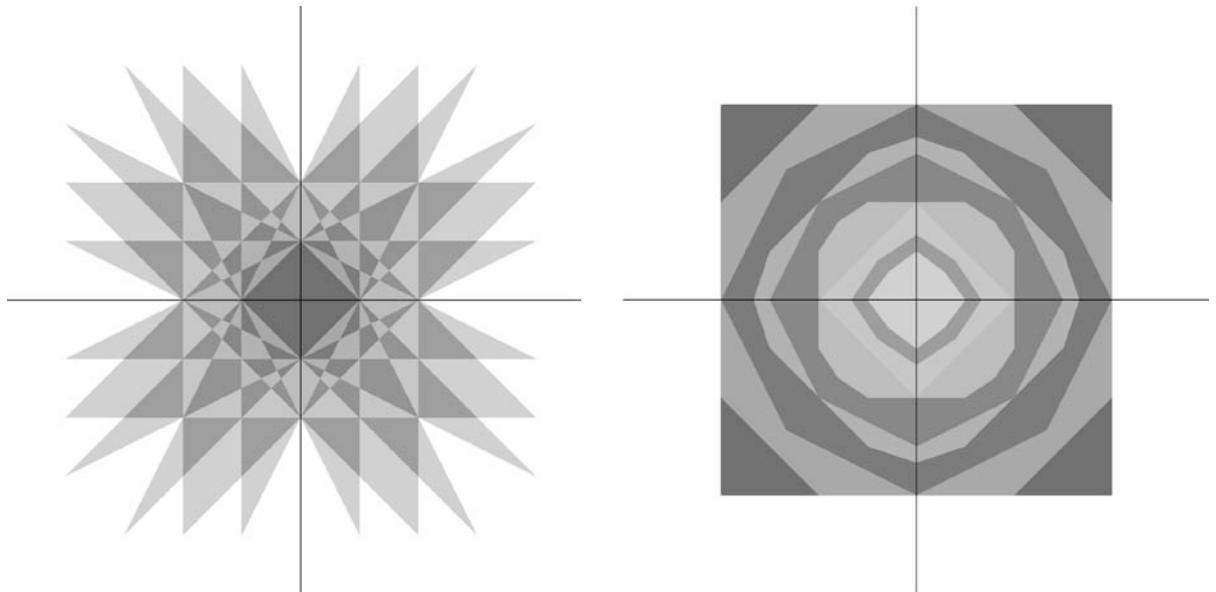
and let  $L$  be the polar transform of  $K_{1/w}$ , that is,

$$L = K_{1/w}^* = \{x \in \mathbb{R}^d : \langle x, u \rangle \leq w(u), \text{ for every } u \in S^{d-1}\}.$$



**Growth Phenomena in Cellular Automata, Figure 2**

The Bell–Eppstein initial set (*left*) that results in  $A_\infty = \mathbb{Z}^2$  for the *Diamoeba* rule. The set  $A_t$ , whose linear asymptotic shape is a rhombus with vertices  $(\pm 1/7, 0)$  and  $(0, \pm 1/8)$ , is shown at  $t = 500$



**Growth Phenomena in Cellular Automata, Figure 3**

The sets  $K_{1/w}$  (left) and the asymptotic shapes for all 10 supercritical range 2 TG CA. Note that there are only 9 shapes, as those with  $\theta = 7$  and  $\theta = 8$  coincide

In general, the polar of a set  $K \subset \mathbb{R}^d$  is given by  $K^* = \{y \in \mathbb{R}^d : \langle x, y \rangle \leq 1 \text{ for every } x \in K\}$ . The set  $L$  is known as a *Wulff shape*, and is a very important notion in crystallography and statistical physics [55]. The next theorem was proved in the classic paper [64]. The core methods in its proof, as well as proofs of similar results [29], are those of convex and discrete geometry.

**Theorem 4** Assume a monotone CA rule with all  $w(u) \geq 0$ . Then there exists a large enough  $r$  so that for every finite initial set  $A_0$ , which contains  $B_2(0, r) \cap \mathbb{Z}^d$ , the linear asymptotic shape from  $A_0$  equals the Wulff shape  $L$ . Even more, the difference between  $A_t$  and  $tL$  is bounded: there exists a constant  $C$ , which depends on the rule and on  $A_0$ , so that  $A_t \subset tL + B_2(0, C)$  and  $tL \subset A_t + B_2(0, C)$  for every  $t \geq 0$ .

Note that supercriticality is not assumed here. If  $w(u) = 0$  for some  $u$ , then  $K_{1/w(u)}$  is an infinite object and  $L$  has dimension less than  $d$ . (The one trivial case is when  $w \equiv 0$  and  $L = \{0^d\}$ .) Finally, note that if there exists a  $u$  so that  $w(u) < 0$  (and hence  $w(-u) < 0$ , by symmetry), then  $A_t$  is sandwiched between two hyperplanes which approach each other and so eventually  $A_t = \emptyset$ .

It is also important to point out that  $K_{1/w}$  is always a polytope,  $L$  is always a convex polytope and both are, for small neighborhoods, readily computable by hand or by computer [30,31,32]. For example, the Moore neigh-

borhood TG CA with  $\theta = 3$  has  $K_{1/w}$  with 16 vertices, of which three successive ones are  $(0, 1)$ ,  $(1, 2)$ ,  $(1, 1)$ , and the remaining 13 are then continued by symmetry. It then follows that the limiting shape  $L$  is the convex hull of  $(\pm 1/2, 0)$ ,  $(0, \pm 1/2)$ ,  $(\pm 1/3, \pm 1/3)$ .

Matters become much murkier when the monotonicity assumption is dropped. We discuss a few interesting two-dimensional solidification examples next. They all hinge on *recursive specification* of iterates  $A_t$  for every  $t$  (see [33] for a definition). This is far from a general approach (and appears to fail even for simple monotone cases), but is the primary technique available.

We begin with the *Box 25* solidification, starting from  $A_0 = B_2(0, r + 1/2) \cap \mathbb{Z}^2$ . As was observed in [33], and can be quickly checked by computer, the linear asymptotic shape exists for  $r = 2$ ,  $r = 9$  and  $r = 13$ , but is in each case different, in fact it is convex in the first case, and nonconvex in the other two cases. This demonstrates that such shapes may depend on the initial seed.

A very interesting example was discovered by D. Hickerson (private communication). Consider *Box 37* solidification, with  $A_0 = B_2(0, 7/2) \cap \mathbb{Z}^2$ . Then  $t^{-1/2}A_t$  converges to  $B_\infty(0, 2\sqrt{2}/3)$  as  $t \rightarrow \infty$ . This demonstrates the possibility of nontrivial *sublinear* asymptotic shapes.

We turn next to the *Hex* rules [36]. These exhibit subsequential limiting shapes, which are not always polygons, as we explain next.

**Theorem 5** Take any of the 16 Hex rules as in Theorem 3, and fix a finite  $A_0$ . There exists a one-parameter family of sets  $S_a$ ,  $a \in [0, 1]$ , so that the following holds: for  $t_n = a \cdot 2^n$ ,

$$2^{-n} A_{t_n} \rightarrow S_a ,$$

as  $n \rightarrow \infty$ .

Furthermore, when 3 and 4 are not both on the solidification list, the family  $S_a$  is called simple and is independent of the initial set. In the opposite, diverse case, initial sets are divided into two classes, distinguished by two different families  $S_a$ .

For rational  $a$ , it can be shown that the Hausdorff dimension of  $\partial S_a$  always exists, and is in principle computable. For example, for the simple  $S_a$  this dimension equals  $5/4$  for  $a = 14/15$ , evidently producing a non-polygonal subsequential shape.

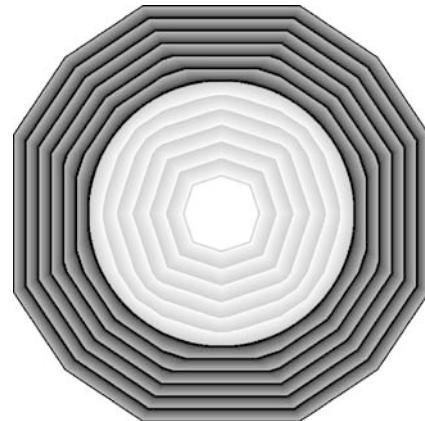
This discussion brings forth the following question, which is probably the most interesting open problem on CA growth. For a *prescribed* set  $L$ , can we find a CA with linear asymptotic shape  $L$ , attained from a “generic” collection of initial sets? In particular, can  $L$  be a circle, thereby giving rise to asymptotic isotropy?

We note that the isotropic construction is possible for probabilistic CA [39], so it seems likely that the answer is yes for a properly constructed chaotic growth. However, techniques for such an approach are completely lacking at present. We should also remark that computational universality should allow for a construction of a CA and a *carefully engineered* initial state with circular (or any other) shape – although this has never been explicitly done. This would, however, violate the requirement of generic initialization.

We conclude this section by a short review of *reverse* shapes [34]. The question here is, if the initial set  $A_0$  is a large hole, and evolves until shortly before the entire lattice is occupied, what is the resulting shape? The initial state has a large and persistent effect on the dynamics and thus the reverse shape geometry will depend on it. The detailed analysis depends on technical convexity arguments, but the cleanest instance is given by the following result.

**Theorem 6** Assume a monotone CA, with  $w \geq 0$  but not identically 0 on  $S^{d-1}$ . Assume also that its rule  $T$  preserves all symmetries of the lattice  $\mathbb{Z}^d$ . Pick a closed convex set  $H \subset \mathbb{R}^d$ , which has all symmetries of  $\mathbb{Z}^d$ , and let  $A_0 = (mH)^c \cap \mathbb{Z}^d$  for some large  $m$ . Moreover, let

$$T = \inf\{t: 0 \in A_t\} .$$



Growth Phenomena in Cellular Automata, Figure 4

Superimposed convergence to the linear asymptotic shape and to the reverse shape, from, respectively, the interior and the exterior, of a large lattice circle. The rule is TG CA with range 2 and  $\theta = 6$ . Iterates are periodically shaded

There is a nonempty bounded convex subset  $\mathcal{R}(H) \subset \mathbb{R}^d$  such that

$$\lim_{M \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{1}{M} \cdot A_{T-M} = \mathcal{R}(H)^c ,$$

in the Hausdorff sense. Moreover, if

$$h_0 = \max\{h > 0: h \cdot H^* \subset K_{1/w}\} ,$$

then

$$\mathcal{R}(H) = (h_0 \cdot H^* \cap \partial K_{1/w})^* .$$

In words, one scales the polar  $H^*$  so that it touches the boundary of  $K^{1/w}$ ; at this point, the intersection determines the reverse shape. (The shape does not change if  $H$  is multiplied by a constant, so  $h_0$  determines its natural scale.) The paper [34] has many more details and examples.

### Nucleation

In this section we assume that the initial state  $A_0$  is the product measure  $\Pi(p)$ , with density  $p > 0$  that is typically very small. Initially, then, there will be no significant activity on most of the space. Certainly this is no surprise as most of the space is empty, but isolated 1’s or small islands of them are often not able to accomplish much either. Most of the lattice is thus in a *metastable* state. However, at certain rare locations there may, by chance, occur local configurations, which are able to spread their influence over large distances until they statistically dominate

the lattice. These locations are called *nuclei*, and their frequency and mechanism of growth are the main self-organizational aspects of the CA rule. The majority of results are confined to two-dimensions, so we will assume  $d = 2$  for the rest of this section and relegate higher dimensions to remarks.

We start with a simple example, for which we give a few details to introduce the basic ideas and demonstrate that a CA can go through more than one metastable state. For this example we do not specify the map  $\pi$ , but instead give a more informal description. In a configuration  $A$ , we call an *insurance* five sites in a cross formation in the state 1, or, more formally, a translate of von Neumann neighborhood which is inside  $A$ . The map  $T$  changes any 0 with a 1 in its von Neumann neighborhood to 1. Moreover, it automatically changes any 1 to 0, except that any 1 whose von Neumann neighborhood intersects with an insurance remains 1. Then, for every  $\epsilon > 0$ , as  $p \rightarrow 0$ ,

$$\begin{aligned} P(0 \in A_t^c \text{ for all } t \leq p^{-1/2+\epsilon}) &\rightarrow 1, \\ P(0 \in (A_t \text{ xor } A_{t+1}) \text{ for all } p^{-1/2-\epsilon} \leq t \leq p^{-5/2+\epsilon}) &\rightarrow 1, \\ P(0 \in A_t \text{ for all } p^{-5/2-\epsilon} \leq t) &\rightarrow 1. \end{aligned}$$

(Here, xor is the exclusive union.) Roughly, most sites are 0 up to time  $p^{-1/2}$ , then periodic with period 2 up to time  $p^{-5/2}$ , and 1 afterwards. (In fact, stronger statements, along the lines of Theorem 7 below, are possible.)

The proof has two phases: the first deterministic and the second probabilistic. For the deterministic one, let  $d_1(x)$  be the  $\ell^1$  distance from  $x$  to  $A_0$ , and assume that  $A_0$  contains no insurance. Then one can prove by induction that, first, none of the  $A_t$  contain an insurance, and second, that for every  $x$  and  $t \geq d_1(x)$ ,  $x \in A_t$  precisely when  $(t - d_1(x)) \bmod 2 = 0$ . On the other hand, an insurance in  $A_0$  centered at the origin will result in  $x \in A_t$  for every  $t \geq d_1(x) - 1$ . The probabilistic part consists of noting that, with overwhelming probability when  $p$  is small,  $B_1(0, p^{-1/2+\epsilon})$  (resp.  $B_1(0, p^{-5/2+\epsilon})$ ) contains no 1 (resp. insurance) in  $A_0 = \Pi(p)$ , while  $B_1(0, p^{-1/2-\epsilon})$  (resp.  $B_1(0, p^{-5/2-\epsilon})$ ) does.

The bulk of the mathematical theory of nucleation and metastability addresses monotone CA, although some work has been done on the *Game of Life* [28] and its generalizations [1,22], excitable media dynamics [19,24,25,42], and artificial life models [51].

Our first general class are supercritical monotone solidification CA. (In fact, the solidification assumption is not necessary, but reduces technical details so much that it is assumed in most published works.) Such rules have two *nucleation parameters*. Let  $\gamma$  be the smallest  $i$  for which there exists an  $A_0$  with  $|A_0| = i$  that generates persis-

**Growth Phenomena in Cellular Automata, Table 2**  
Nucleation parameter  $\nu$  for small Box neighborhood TG CA

	$\theta = 2$	$\theta = 3$	$\theta = 4$	$\theta = 5$	$\theta = 6$	$\theta = 7$
$\rho = 1$	12	42				
$\rho = 2$	40	578	4683	24938	94050	259308

tent growth. Moreover, let  $\nu$  be the number of sets  $A_0$  of size  $\gamma$  that generate persistent growth and have the leftmost among their lowest sites at the origin. (The last requirement ensures that  $\nu$  counts the number of distinct smallest “shapes” that grow.) We call the rule *voracious* if, started from any of the  $\nu$  initial sets  $A_0$  described above,  $A_\infty = \mathbb{Z}^2$ . Voracity is a weak condition, which assures a minimal regularity of growth and can, for any fixed rule, be checked on finitely many cases (which is not true for the more restrictive omnivorous property).

For illustration, we briefly discuss these for range  $\rho$  Box neighborhood TG CA. For relatively small  $\theta$ ,  $\gamma = \theta$ ; for example, when  $\rho = 1$ ,  $\gamma = \theta$  for all three supercritical rules, while when  $\rho = 2$ ,  $\gamma$  exceeds  $\theta$  only for  $\theta = 10$ , when it equals 11. For large  $\rho$ , and  $\theta \sim \alpha\rho^2$ ,  $\gamma$  is asymptotically the smallest possible (that is,  $\gamma \sim \alpha\rho^2$ ) when  $\alpha < \alpha_c$  for some  $\alpha_c \in (1.61, 1.66)$  [32]. One can also compute some  $\nu$ , before they become too large.

Returning to  $A_0 = \Pi(p)$ , the most natural statistics to study is

$$T = \inf\{t: 0 \in A_t\},$$

the first time the CA occupies the origin.

**Theorem 7** Assume a monotone, supercritical, and voracious CA, with nucleation parameters  $\gamma$  and  $\nu$ . Then, as  $p \rightarrow 0$ ,

$$\sqrt{\nu p^\gamma} \cdot T$$

converges in distribution to a nontrivial random variable  $\tau$ , which is a functional of a Poisson point location  $\mathcal{P}$  with unit intensity.

That  $T \approx p^{-\gamma/2}$  can be easily guessed (and proved), but the more precise asymptotics described above require a considerable argument [30], as interaction between growing droplets is nontrivial. In particular, the higher dimensional version has not been proved, and the description of the limiting “movie” from  $\mathcal{P}$  probably cannot avoid viscosity methods from PDE [58].

The most exciting nucleation results have been proved about *critical* models, for which  $w(u)$  vanishes for some direction  $u$  but is positive for others. Although a general framework is presented in [35], we will instead focus on

the most studied examples. Of these the most popular has been the *bootstrap percolation (BP)*, which is TG CA with von Neumann neighborhood and  $\theta = 2$  [2,3,4,67]. Its *modified* version (*MBP*) has the same neighborhood, still solidifies, but when  $0 \notin S$ ,  $\pi(S) = 1$  precisely when  $\{\pm e_1\} \cap S \neq \emptyset$  and  $\{\pm e_2\} \cap S \neq \emptyset$ . (Here  $e_1$  and  $e_2$  are the basis vectors.)

Now  $w(\pm e_1) = w(\pm e_2) = 0$ , so no finite set can generate persistent growth, and it is not immediately clear that  $P(T < \infty) = 1$ . This is true [67], as very large sets are able to use sparse but helpful smattering of 1's around them and so are unlikely to be stopped. To determine the size of  $T$ , one needs more information about the necessary size of these nuclei, and the likelihood of their formation. This was started in [4] and culminated by the following theorem by A. Holroyd [46], which is arguably the crowning achievement of CA nucleation theory to date.

**Theorem 8** *For BP let  $\lambda = \pi^2/18$ , and for MBP let  $\lambda = \pi^2/6$ . Then, for every  $\epsilon > 0$ ,*

$$P(p \log T \in [\lambda - \epsilon, \lambda + \epsilon]) \rightarrow 1$$

as  $p \rightarrow 0$ .

To summarize,  $T \approx \exp(\lambda/p)$ , which is for small  $p$  a long time indeed and amply justifies the description of the almost empty lattice as metastable.

The most common formulation of the theorem above involves finite  $L \times L$  squares with periodic boundary instead of infinite lattices. Then

$$I(L, p) = P(\text{the entire square is eventually occupied})$$

and, as  $p \rightarrow 0$ ,

$$\begin{aligned} I(L, p) &\rightarrow 1 && \text{if } p \log L \geq \lambda + \epsilon, \\ I(L, p) &\rightarrow 0 && \text{if } p \log L \leq \lambda - \epsilon. \end{aligned}$$

Here  $L$  is of course assumed to increase with  $p$ . Before the value of  $\lambda$  was known, this second formulation was used to estimate it by simulation. For example, [3] used  $L$  close to 30,000 and obtained  $\lambda \approx 0.245$  for BP, about a factor of two smaller than the true value 0.548.... Other simulations of BP, MBP, and related models all exhibit a similar discrepancy. The reason apparently is that nuclei are, for realistic values of  $p$ , quite a bit more frequent than the asymptotics would suggest. Indeed, the following result from [38] confirms this.

**Theorem 9** *For BP and MBP,*

$$I(L, p) \rightarrow 1 \quad \text{if } p \log L \geq \lambda - c(\log L)^{-1/2},$$

for an appropriate constant  $c$ .

This alone indicates that to halve the error in approximating  $\lambda$  on an  $L \times L$  system it is necessary to replace  $L$  by  $L^4$ . In addition, [38] shows that for the more tractable MBP one can do explicit calculations to conclude that to get an estimate of  $\lambda$  within 2%, one would need  $L$  at least  $10^{500}$ , a non-achievable size.

For BP, the quantity  $p \log L$  is the “order parameter,” the quantity that, when varied, causes a phase transition (which, in addition, is sharp by Theorem 8). We will list now some other models with known order parameters – we also indicate the status of phase transition, when known:

- CA with von Neumann neighborhood and  $\pi(S) = 1$  when  $|S \setminus \{0\}| \geq 2$ :  $p^2 \log L$  [57];
- TG CA with range  $\rho$  Box neighborhood,  $\theta \in [2\rho^2 + \rho + 1, 2\rho^2 + 2\rho]$ :  $p^{\theta-2\rho^2-\rho} \log L$  [30];
- TG CA with  $\mathcal{N} = \{(0, 0), (0, \pm 1), (\pm 1, 0), (\pm 2, 0)\}$ ,  $\theta = 2$ :  $p^{3/2} L$ , not sharp [30];
- TG CA with  $\mathcal{N} = \{(0, 0), (0, \pm 1), (\pm 1, 0), (\pm 2, 0)\}$ ,  $\theta = 3$ :  $(-\log p)^{-2} p \log L$  [30,68];
- TG CA with range  $\rho$  cross neighborhood  $\mathcal{N} = \{(x, y) : |x| \leq \rho, |y| \leq \rho, xy = 0\}$  and  $\theta = \rho + 1$ :  $p \log L$ , sharp at  $\lambda = \pi^2/(3(\rho + 1)(\rho + 2))$  [48];
- TG CA on  $\mathbb{Z}^d$  with  $\mathcal{N} = B_1(0, 1) \cap \mathbb{Z}^d$  and  $\theta \in [3, d]$ ,  $p^{1/(d-\theta+1)} \log_{\theta-1} L$  (where  $\log_k$  is the  $k$ 'th iterate of  $\log$ ) [14], sharp at  $\lambda = \pi^2/6$  for the modified version when  $\theta = d$  [47].

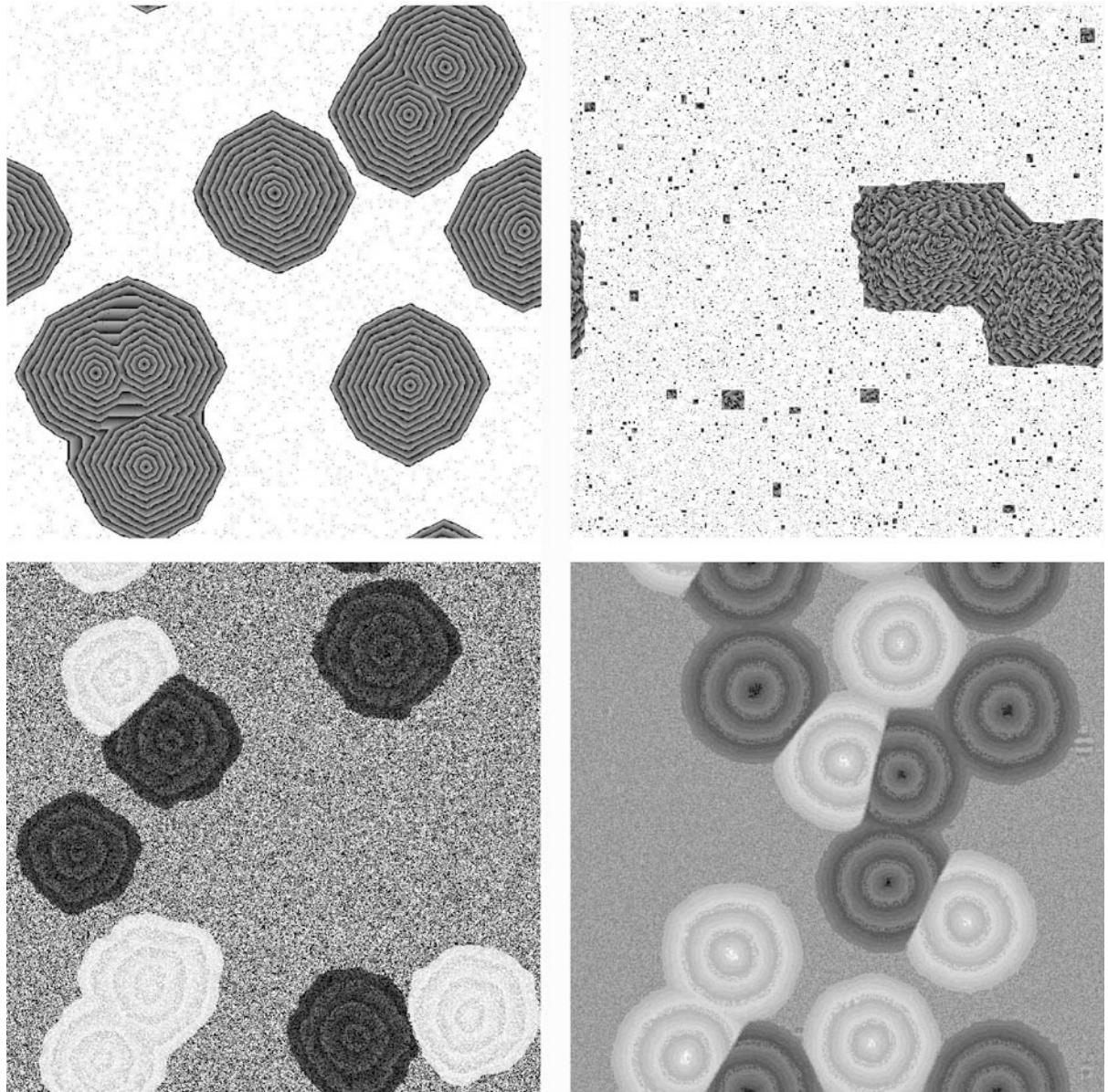
Note that when  $\theta = d = 3$  the last example gives the metastable scale  $\exp(\exp(\lambda/p))$  [13,56], making an even modest experimental approximation of  $\lambda$  impossible.

There are other interesting issues about critical growth models, which do not have to do with nucleation. One is decay rate for the first passage time  $T$  [6], which is connected to the properties of the very last holes to be filled in. Another is its ability to overtake random obstacles [40].

Apart from sending  $p \rightarrow 0$ , one could vary other parameters to get the metastability phenomena, and one natural example is the range. We explain this scenario on a non-monotone CA known as the *Threshold Voter Automaton (TVA)* [20,31]. For simplicity, assume  $\mathcal{N}$  is range  $\rho$  Box neighborhood, and fix a threshold  $\theta$ . This rule makes a site change its “opinion” by contact with at least  $\theta$  of the opposite opinions:

$$\begin{aligned} \pi(S) = 1 &\quad \text{iff } (0 \in S \text{ and } |S^c| < \theta) \\ &\quad \text{or } (0 \notin S \text{ and } |S| \geq \theta). \end{aligned}$$

As the two opinions are symmetric, the most natural initial state of TVA is  $\Pi(1/2)$ . We also assume that  $\rho$  is large and the scaling  $\theta = a|\mathcal{N}|$ , for some  $a \in (0, 1)$ .



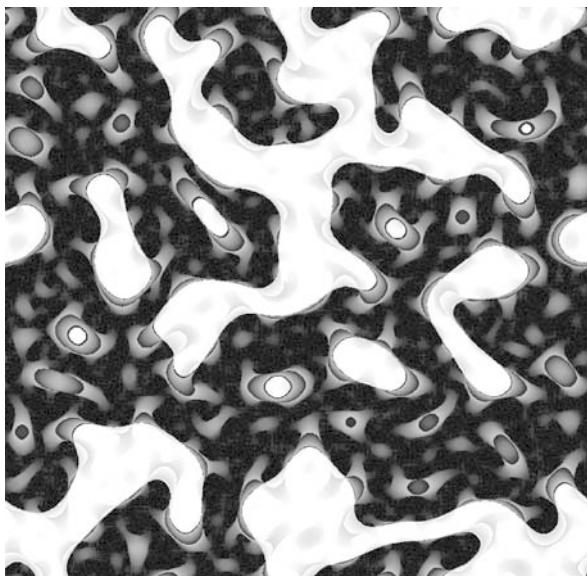
**Growth Phenomena in Cellular Automata, Figure 5**

Four nucleation examples, each on an  $800 \times 800$  array with periodic boundary. Clockwise from top left: TGM CA with Moore neighborhood,  $\theta = 3$ , and  $p = 0.006$ ; bootstrap percolation with  $p = 0.041$ ; TVA with  $\rho = 10$  and  $\theta = 194$ ; TVA with  $\rho = 10$  and  $\theta = 260$ . The iterates are periodically colored to indicate growth and, in the TVA frames, the lighter shades indicate 0's

It is proved in [20] that when  $a > 3/4$ , any fixed  $x \in \mathbb{Z}^2$  changes its opinion only finitely many times with probability approaching 1 as  $\rho \rightarrow \infty$  — and the rigorous results end there. The most interesting rare nucleation questions arise when  $a \in (1/4, 3/4) \setminus \{1/2\}$ . According to simulations, under this assumption the nuclei are rare and eventually tessellate the lattice into regions of consensus with either stable or periodic boundaries [31]. However, the

definition of a nucleus is unclear and consequently their density cannot be estimated. Two torus simulations are given in Fig. 5; it is important to point out that, for such *finite* systems, Lyapunov methods of [27] imply that every site eventually fixates or becomes periodic with period 2.

The *majority* TVA, when  $a = 1/2$ , is perhaps the most appealing of all [43]. The nucleation is now not rare; instead, this CA quickly self-organizes into visually attrac-



**Growth Phenomena in Cellular Automata, Figure 6**

Majority vote: TGM with  $\rho = 10$ ,  $\theta = 221$  on a  $1000 \times 1000$  array with periodic boundary. Again, iterates are periodically colored with the lighter shades reserved for 0's

tive curvature driven dynamics. (Note that flat interfaces between the two opinions are now stable, so one opinion can advance only when it forms a concavity.) For any fixed  $\rho$  this must eventually stop, as finite islands with uniformly small enough curvature of either opinion are stable. However, when  $\rho$  is increased this effect is with large probability not felt on any fixed finite portion of space. (A similar effect is achieved by the Vichniac “twist” [61].) Many fascinating questions remain about this case, especially on the initial nucleating phase, whose analysis depends on delicate properties of random fields and remains an open problem.

### Future Directions

We will identify seven themes, which connect to open problems discussed in previous sections. Progress on each is bound to be a challenge, but also a significant advance in understanding CA growth.

### Regularity of Growth

It is often important, and of independent interest, to be able to conclude that a cellular automaton rule generates growth without arbitrarily large tentacles, holes, or other undesirable features. An omnivorous CA, for example, has this property. The natural goal would be to develop techniques to establish such regularity for much more gen-

eral monotone and non-monotone CA, and for arbitrary dimension. Many rules give the impression that regular growth is a generic trait, i. e., holds for a majority of initial sets.

### Oscillatory Growth

Does there exist a class of CA with growing sets that oscillate on different scales? Hickerson's *Diamoeba* might be able to accomplish this from some initial sets, but perhaps there are other, more tractable, examples with identifiable mechanisms.

### Analysis of Chaotic Growth

One look at the growth of *Box 1* solidification from a  $8 \times 8$  initial box (bottom left frame in Fig. 1) would convince most observers that it has a square asymptotic shape. However, there are no tools to prove, or disprove, this statement. A fully rigorous theory of chaotic CA, tailored to address such asymptotic issues, is almost nonexistent and constitutes perhaps the most important challenge for mathematicians in this area.

### Nucleation Theory for Non-monotone Models

Once nucleation centers are established, growth most often proceeds in a random environment, which consists of debris left over from the nucleation phase. This may help the analysis, as it adds a random perturbation to what may otherwise be intractable dynamics, but on the other hand random environment processes are notoriously tricky to analyze [41].

### Robust Exact Constants and Sharp Transitions

The nucleation phase transition has been proved sharp for a few critical models, by rather delicate arguments. A more robust approach would extend them, and would perhaps provide further insights into the error terms, for which only one sided estimates are now known. The apparent crossover [3] phenomenon would also be interesting to understand rigorously.

### Three-Dimensional Nucleation and Growth

With advances in computer power, extensive three-dimensional CA simulations have become viable on commercial hardware. Therefore, it may be possible to investigate nucleation, droplet interaction, clustering mechanisms, and other staples of two-dimensional CA research,

at least experimentally. Proper visualization tools of complex three-dimensional phenomena may well require some novel ideas in computer graphics.

### Generic Properties of CA with Large Range

A TG CA with range  $\rho$ , say, has on the order of  $\rho^2$  possible thresholds  $\theta$ . When can it be established that some property holds for the majority of relevant choices? One such property (sensitivity of shapes to random perturbations in the rule) was analyzed from this perspective in [37], but it would be interesting to provide further examples for TG or other classes of CA.

## Bibliography

### Primary Literature

1. Adamatzky A, Martínez GJ, Mora JCST (2006) Phenomenology of reaction-diffusion binary-state cellular automata. *Int J Bifur Chaos Appl Sci Eng* 16:2985–3005
2. Adler J (1991) Bootstrap percolation. *Physica A* 171:453–4170
3. Adler J, Staufer D, Aharony A (1989) Comparison of bootstrap percolation models. *J Phys A: Math Gen* 22:L279–L301
4. Aizenman M, Lebowitz J (1988) Metastability effects in bootstrap percolation. *J Phys A: Math Gen* 21:3801–3813
5. Allouche J-P, Shallit J (2003) Automatic sequences: theory, applications, generalizations. Cambridge University Press, Cambridge
6. Andjel E, Mountford TS, Schonmann RH (1995) Equivalence of decay rates for bootstrap percolation like cellular automata. *Ann Inst H Poincaré* 31:13–25
7. Berlekamp ER, Conway JH, Guy RK (2004) Winning ways for your mathematical plays, vol 4, 2nd edn. Peters, Natick
8. Bohman T (1999) Discrete threshold growth dynamics are omnivorous for box neighborhoods. *Trans Am Math Soc* 351:947–983
9. Bohman T, Gravner J (1999) Random threshold growth dynamics. *Random Struct Algorithms* 15:93–111
10. Bramson M, Neuhauser C (1994) Survival of one-dimensional cellular automata under random perturbations. *Ann Probab* 22:244–263
11. Brummitt CD, Delventhal H, Retzlaff M (2008) Packard snowflakes on the von Neumann neighborhood. *J Cell Autom* 3:57–80
12. Bäck T, Dörnemann H, Hammel U, Frankhauser P (1996) Modeling urban growth by cellular automata. In: Lecture Notes in Computer Science. Proceedings of the 4th International Conference on Parallel Problem Solving from Nature, vol. 1141. Springer, Berlin, pp 636–645
13. Cerf R, Cirillo ENM (1999) Finite size scaling in three-dimensional bootstrap percolation. *Ann Probab* 27:1837–1850
14. Cerf R, Manzo F (2002) The threshold regime of finite volume bootstrap percolation. *Stoch Process Appl* 101:69–82
15. Chopard B, Droz M (1998) Cellular automata modeling of physical systems. Cambridge University Press, Cambridge
16. Cobham A (1972) Uniform tag sequences. *Math Syst Theory* 6:164–192
17. Cook M (2005) Universality in elementary cellular automata. *Complex Syst* 15:1–40
18. Deutsch A, Dormann S (2005) Cellular automata modeling of biological pattern formation. Birkhäuser, Boston
19. Durrett R, Steif JE (1991) Some rigorous results for the Greenberg–Hastings model. *J Theor Probab* 4:669–690
20. Durrett R, Steif JE (1993) Fixation results for threshold voter systems. *Ann Probab* 21:232–247
21. Eppstein D (2002) Searching for spaceships. In: More Games of no Chance (Berkeley, CA, 2000). Cambridge University Press, Cambridge, pp 351–360
22. Evans KM (2001) Larger than Life: digital creatures in a family of two-dimensional cellular automata. In: Cori R, Mazoyer J, Morvan M, Mosseri R (eds) Discrete Mathematics and Theoretical Computer Science, vol AA. pp 177–192
23. Evans KM (2003) Replicators and larger than life examples. In: Griffeath D, Moore C (eds) New constructions in cellular automata. Oxford University Press, New York, pp 119–159
24. Fisch R, Gravner J, Griffeath D (1991) Threshold-range scaling for the excitable cellular automata. *Stat Comput* 1:23–39
25. Fisch R, Gravner J, Griffeath D (1993) Metastability in the Greenberg–Hastings model. *Ann Appl Probab* 3:935–967
26. Gardner M (1976) Mathematical games. *Sci Am* 133:124–128
27. Goles E, Martínez S (1990) Neural and Automata Networks. Kluwer, Dordrecht
28. Gotts NM (2003) Self-organized construction in sparse random arrays of Conway's game of life. In: Griffeath D and Moore C (eds) New Constructions in Cellular Automata. Oxford University Press, New York, pp 1–53
29. Gravner J, Griffeath D (1993) Threshold growth dynamics. *Trans Am Math Soc* 340:837–870
30. Gravner J, Griffeath D (1996) First passage times for the threshold growth dynamics on  $\mathbb{Z}^2$ . *Ann Probab* 24:1752–1778
31. Gravner J, Griffeath D (1997) Multitype threshold voter model and convergence to Poisson-Voronoi tessellation. *Ann Appl Probab* 7:615–647
32. Gravner J, Griffeath D (1997) Nucleation parameters in discrete threshold growth dynamics. *Exp Math* 6:207–220
33. Gravner J, Griffeath D (1998) Cellular automaton growth on  $\mathbb{Z}^2$ : theorems, examples and problems. *Adv Appl Math* 21:241–304
34. Gravner J, Griffeath D (1999) Reverse shapes in first-passage percolation and related growth models. In: Bramson M, Durrett R (eds) Perplexing Problems in Probability. Festschrift in Honor of Harry Kesten. Birkhäuser, Boston, pp 121–142
35. Gravner J, Griffeath D (1999) Scaling laws for a class of critical cellular automaton growth rules. In: Révész P, Tóth B (eds) Random Walks. János Bolyai Mathematical Society, Budapest, pp 167–186
36. Gravner J, Griffeath D (2006) Modeling snow crystal growth. I. Rigorous results for Packard's digit snowflakes. *Exp Math* 15:421–444
37. Gravner J, Griffeath D (2006) Random growth models with polygonal shapes. *Ann Probab* 34:181–218
38. Gravner J, Holroyd AE (2008) Slow convergence in bootstrap percolation. *Ann Appl Probab* 18:909–928
39. Gravner J, Mastronarde N Shapes in deterministic and random growth models (in preparation)
40. Gravner J, McDonald E (1997) Bootstrap percolation in a polluted environment. *J Stat Phys* 87:915–927
41. Gravner J, Tracy C, Widom H (2002) A growth model in a random environment. *Ann Probab* 30:1340–1368

42. Greenberg J, Hastings S (1978) Spatial patterns for discrete models of diffusion in excitable media. *SIAM J Appl Math* 4:515–523
43. Griffeath D (1994) Self-organization of random cellular automata: four snapshots. In: Grimmett G (ed) *Probability and Phase Transition*. Kluwer, Dordrecht, pp 49–67
44. Griffeath D, Hickerson D (2003) A two-dimensional cellular automaton with irrational density. In: Griffeath D, Moore C (eds) *New constructions in cellular automata*. Oxford University Press, Oxford, pp 119–159
45. Griffeath D, Moore C (1996) Life without death is P-complete. *Complex Syst* 10:437–447
46. Holroyd AE (2003) Sharp metastability threshold for two-dimensional bootstrap percolation. *Probab Theory Related Fields* 125:195–224
47. Holroyd AE (2006) The metastability threshold for modified bootstrap percolation in  $d$  dimensions. *Electron J Probab* 11:418–433
48. Holroyd AE, Liggett TM, Romik D (2004) Integrals, partitions, and cellular automata. *Trans Am Math Soc* 356:3349–3368
49. Jen E (1991) Exact solvability and quasiperiodicity of one-dimensional cellular automata. *Nonlinearity* 4:251–276
50. Kier LB, Seybold PG, Cheng C-K (2005) *Cellular Automata Modeling of Chemical Systems*. Springer, Dordrecht
51. Lindgren K, Nordahl MG (1994) Evolutionary dynamics of spatial games. *Physica D* 75:292–309
52. Meakin P (1998) Fractals, scaling and growth far from equilibrium. Cambridge University Press, Cambridge
53. Packard NH (1984) Lattice models for solidification and aggregation. Institute for Advanced Study preprint. Reprinted in: Wolfram S (ed) (1986) *Theory and application of cellular automata*. World Scientific, Singapore, pp 305–310
54. Packard NH, Wolfram S (1985) Two-dimensional cellular automata. *J Stat Phys* 38:901–946
55. Pimpinelli A, Villain J (1999) *Physics of Crystal Growth*. Cambridge University Press, Cambridge
56. Schonmann RH (1992) On the behavior of some cellular automata related to bootstrap percolation. *Ann Probab* 20:174–193
57. Schonmann RH (1990) Finite size scaling behavior of a biased majority rule cellular automaton. *Physica A* 167:619–627
58. Song M (2005) Geometric evolutions driven by threshold dynamics. *Interfaces Free Bound* 7:303–318
59. Toffoli T, Margolus N (1997) *Cellular Automata Machines*. MIT Press, Cambridge
60. Vichniac GY (1984) Simulating physics with cellular automata. *Physica D* 10:96–116
61. Vichniac GY (1986) Cellular automata models of disorder and organization. In: Bienenstock E, Fogelman-Soulie F, Weisbuch G (eds) *Disordered Systems and Biological Organization*. Springer, Berlin, pp 1–20
62. Wiener N, Rosenblueth A (1946) The math foundation of the problem of conduction of impulses in a network of connected excitable elements, specifically in cardiac muscle. *Arch Inst Cardiol Mex* 16:205–265
63. Willson SJ (1978) On convergence of configurations. *Discret Math* 23:279–300
64. Willson SJ (1984) Cellular automata can generate fractals. *Discret Appl Math* 8:91–99
65. Willson SJ (1987) Computing fractal dimensions for additive cellular automata. *Physica D* 24:190–206
66. Wójcik M (2001) Mirek's Cellebration: a 1D and 2D Cellular Automata Explorer, Version 4.20. <http://www.mirwoj.opus.chelm.pl/ca/>
67. van Enter ACD (1987) Proof of Straley's argument for bootstrap percolation. *J Stat Phys* 48:943–945
68. van Enter ACD, Hulshof T (2007) Finite-size effects for anisotropic bootstrap percolation: logarithmic corrections. *J Stat Phys* 128:1383–1389
- ### Books and Reviews
- Adamatzky A (1995) *Identification of cellular automata*. Taylor & Francis, London
- Allouche J-P, Courbage M, Kung J, Skordev G (2001) Cellular automata. In: *Encyclopedia of physical science and technology*, vol 2, 3rd edn. Academic Press, San Diego, pp 555–567
- Allouche J-P, Courbage M, Skordev G (2001) Notes on cellular automata. *Cubo, Matemática Educational* 3:213–244
- Durrett R (1988) *Lecture Notes on particle Systems and Percolation*. Wadsworth & Brooks/Cole, Pacific Grove
- Durrett R (1999) Stochastic spatial models. *SIAM Rev* 41:677–718
- Gravner J (2003) Growth phenomena in cellular automata. In: Griffeath D, Moore C (eds) *New Constructions in Cellular Automata*. Oxford University Press, New York, pp 161–181
- Holroyd AE (2007) Astonishing cellular automata. *Bull Centre Rech Math* 10:10–13
- Ilachinsky A (2001) *Cellular Automata: A Discrete Universe*. World Scientific, Singapore
- Liggett TM (1985) *Interacting Particle Systems*. Springer, New York
- Liggett TM (1999) *Stochastic Interacting Systems: contact, voter and exclusion processes*. Springer, New York
- Rothman DH, Zaleski S (1997) *Lattice-Gas Cellular Automata*. Cambridge University Press, Cambridge
- Toom A (1995) Cellular automata with errors: problems for students of probability. In: Snell JL (ed) *Topics in contemporary probability and its applications*. CRC Press, Boca Raton, pp 117–157