Agent-Based Simulation in the Study of Social Dilemmas

Article *in* Artificial Intelligence Review · March 2003

DOI: 10.1023/A:1022120928602 · Source: DBLP

Nick Gotts
Independent Researcher

J. Gary Polhill
James Hutton Institute

Agent-Based Simulation In The Study Of Social Dilemmas

N.M. Gotts, J.G. Polhill and A.N.R. Law
Macaulay Land Use Research Institute
Craigiebuckler
Aberdeen AB15 8QH
Scotland, UK
Telephone: +44-1224-498200
Fax: +44-1224-311556
Email: n.gotts@mluri.sari.ac.uk

Abstract.
This review discusses *agent-based social simulation* (ABSS[1]) in relation to the study of *social dilemmas* such as the *Prisoner's Dilemma* and *Tragedy of the Commons.* Its aims are to explore the place of ABSS in relation to other research methods such as mathematical analysis, to familiarise artificial intelligence researchers (particularly those working on multi-agent systems) with a body of relevant multidisciplinary work, and to suggest directions for future ABSS research on social dilemmas.

ABSS research can contribute greatly to the understanding of social phenomena, but needs to be based on a clear appreciation of the current 'state of play' in the areas where it is used. With regard to 'thin' (simple, general) simulation models, this primarily means attending to what has been or could be discovered by mathematical analysis, to work using other forms of simulation, and to the relevant theoretical disputes; with regard to 'thick' (specific, detailed) models (about which the paper has less to say), linking to the relevant 'thin' models and to the empirical evidence. The bulk of ABSS work on social dilemmas has been concentrated in quite a narrow — though certainly significant — area (reciprocal altruism in the Prisoner's Dilemma), and has sometimes been seriously flawed by over-ambitious claims, and insufficient attention to analytical approaches — although this same work has been very fertile in terms of inspiring further work, both analytical and simulation-based.

Keywords: agents, altruism, cooperation, economics, evolution, games, simulation

Table of Contents

1. Introduction

This review discusses the contribution made by *Agent-Based Social Simulation* (henceforth ABSS), to the study of *social dilemmas* such as the *Prisoner's Dilemma* and *Tragedy of the Commons*, and the directions which such work could most fruitfully take in future.

In ABSS, social phenomena are studied through the medium of computational models which represent agents and their interactions with each other individually. In this review, ABSS covers all such computational models, however simple the agents and interactions in the model may be. The agents are most frequently thought of as representing individual organisms (humans, other animals, or even plants), but may also stand for collectives such as firms or states, or artificial entities (robots or softbots). As an approach to the simulation of social phenomena, ABSS can be contrasted with *numerical simulation*, in which the salient features of the social system being modelled are represented by numbers, generally related to each other by differential or difference equations. (The features represented might be, for example, the prices of goods in an economic simulation, the numbers of people in different age groups in a demographic model, or the proportions of individuals pursuing particular strategies in a situation involving conflict or competition. The last of these possibilities is the most relevant in this paper.)

In a social dilemma, two or more participants must each choose between following their own immediate interests, or the common interest of all participants. For each participant, choosing the 'selfish' option is immediately advantageous, *whatever* the other participants do, but if enough participants take this option, *all* end up worse off than if enough had made the 'altruistic' or 'cooperative' choice. The review provides an overview of the highly interdisciplinary body of research on social dilemmas, while emphasising ABSS work. It is aimed at all researchers interested in social cognition and behaviour, particularly those working in ABSS itself, and in the wider areas of multi-agent systems and artificial life.

Some influential models of human behaviour suggest that it should be largely or even completely selfish, notably the *Homo economicus* model within economics, and what might be called the *struggle for existence* or *Social Darwinist* model within evolutionary biology. These appear at odds with empirical evidence of cooperation and altruism in both experimental and natural settings. This has led to sustained work on social dilemmas in many disciplines over recent decades.

Social dilemmas are highly relevant to environmental issues, which often arise from the cumulative actions of many independent agents creating serious pollution, or depleting a finite resource. Most of the more applications-oriented studies reviewed here focus on this area. However, in any context where collective action is necessary, whether the aim is economic, political, or simply sociable, and whether or not a formal organisation exists to coordinate it, some individuals may cheat, shirk, or otherwise undermine the collective good for their own ends. Social dilemmas also arise in contexts where agents are in economic competition, or in political or even military conflict, so long as there are some rules or *norms* in force, which it may be advantageous for an agent to break, but which provide a common benefit. Any sufficiently sophisticated form of social modelling will therefore need to consider social dilemmas at some point.

In one view, ABSS is to the larger research domain of multi-agent systems (henceforth MAS) as cognitive modelling is to artificial intelligence (AI). Most MAS research treats real-world social processes and structures simply as potential inspiration for computational approaches to be judged on their efficiency, robustness and so forth, while ABSS aims to advance understanding of those real-world processes and structures. However, ABSS can also be seen as a subdomain of artificial life (henceforth Alife). MAS has focused on interactions between small numbers of relatively complex agents (with internal symbolic representations of their environment, and abilities such as planning). Alife has paid more attention to the emergent properties of large collections of agents, responding to their immediate environment — including each other — according to simple rules or regularities.

It is striking that several recent books on MAS (Tokoro, 1996; Müller et al., 1999; Ferber, 1999; Weiss, 1999), have little to say about ABSS. Indeed Wooldridge and Jennings (1998)[p.389] argue that dense, complex webs of interaction between large numbers of agents should be avoided in agent-based software engineering, as their dynamics are complex and often chaotic. In contrast, the dynamics of such webs are of great interest to social scientists using ABSS. However Moss (2001) responds that if MAS is to be useful in relation to 'messy distributed software systems' such as the Internet and federated databases, the dynamics of similarly 'messy' real-world social systems become highly relevant. We also note that MAS for use in online trading, or in collaboration between organisations which remain competitors or potential opponents, are necessarily *embedded* in real social systems, where the principals that the software agents represent may try to use the MAS to cheat each other. Investigation of cooperation despite such temptation in real social systems will be highly relevant to such MAS applications.

Doran and Gilbert (1994), discussing the wider topic of scientific models with dynamic aspects, contrast *exploratory* modelling, aimed at gaining new insights into the system modelled, and *focused* modelling, typically aimed at predicting the system's behaviour in specific conditions. Similarly Kliemt (1996) distinguishes *thick* and *thin* simulations. 'Thick' simulations are detailed, draw on copious empirical data, and tell the investigator a lot about a specific question (such as the optimum distribution of sales staff in a specific retail outlet), but only about that. Such simulations are useful in domains which have traditionally employed case studies. 'Thin' simulations are a tool for 'controlled speculation', useful in disciplining theory formation; they often use simplifying and distorting assumptions. Most of the ABSS studies described here are 'thin', reflecting the balance of the social dilemmas literature, and use rather simple agents, which would not meet many MAS practitioners' criteria for agenthood.

The relationship between various forms of simulation and mathematical analysis is a major theme of this paper. ABSS at the thin end of the spectrum can be seen either as complementary to, or in competition with, analytical approaches. Specifically, thin ABSS work on social dilemmas is closely related to work in evolutionary game theory, as explored at some length below. Doran and Gilbert (1994) cite as a particular problem for computer simulation the 'trap of verisimilitude': the temptation to add plausible detail to the model simply because it can be added. They note that the corresponding trap for those using mathematical analysis is the 'trap of tractability': subordinating everything to producing a model with analysable behaviour. Indeed, simulation may be useful when the choice lies between models which are analysable but make implausible assumptions, and models which are more plausible, but because of greater complexity (frequently, because of nonlinear interactions between parts of the model), are analytically intractable.

Chattoe (1996), taking economics as an application domain, argues that in several respects simulation is *more* rigorous than mathematical modelling. He contends that reliance on sets of analytically soluble equations to represent social processes makes it effectively impossible even to consider alternatives to economic orthodoxy, despite the implausibility of the assumption that real economic agents make use of such sets of equations, and the lack of any adequate alternative justification for them. Also, it is typically difficult to identify those *parts* of an equation-based model responsible for poor performance. Parts of the model do not generally correspond in any clear way to parts of the system modelled, and if such a part can be identified, replacing it may render the system of equations insoluble. Simulations, Chattoe claims, are much more readily developed in a modular fashion. Moreover, they generally produce 'emergent data', which can be used as a check on their operation (Chattoe, 1996) [p.100, emphasis in original]:

> For example, in a model of firm interaction, all that need be simulated is the pricing decisions of individual firms, and their survival or bankruptcy... The age profile of firms comprises an independent *emergent* piece of data resulting from the independent actions of many firms, which is not designed into the simulation and can be compared with the age profile of real firms.

Ideally, analysis and simulation can guide each other. Partial mathematical analyses (of special cases, for example), may be sought and found when a general problem is intractable, and then act as a valuable check on simulation results, which can in turn suggest new goals for analytical work. However, as both Johnson (1999) and Halpin (1999) warn, there is no point simulating when an analytical solution is already known or within reach. Using simulation simply because the researcher lacks the mathematical knowledge or skills to undertake analytical work risks precisely this.

Section 2 briefly describes the work which brought social dilemmas to the attention of research communities, and reviews some of the main findings and current areas of interest in recent empirically oriented work. Section 3 outlines a range of mechanisms which have been proposed as explanations for human cooperation and apparent altruism. Sections 4 to 6 review work on the Prisoner's Dilemma, while sections 7 and 8 consider multi-agent social dilemmas. Section 9 outlines implications and applications of social dilemmas research, and section 10 assesses the existing and potential contributions of ABSS to such research. Some readers may prefer to skim the discussion of individual papers in sections 4-8 initially.

2. Social Dilemmas

2.1. Game Theory, Evolution and Altruism

While ideas underlying research into social dilemmas can be traced back several centuries (Hobbes, 1914), new light was thrown upon the area with the development of game theory (von Neumann and Morgenstern, 1944; Nash, 1951), and particularly the theory of *non zero-sum games* in the 1940s and 1950s.

Game theory models strategic interactions between agents (generally called *players*) with interests that are at least partially incompatible. In a *zero-sum game*, one player's gain is always balanced by equal losses on the part of one of more of the others. In non zero-sum games, such gains and losses need not balance. Game theory focuses on analysing the properties of *strategies*, that is, rules specifying which of a (finite or infinite) set of actions an agent should choose in any situation. A game may give an agent one choice between options, or a succession of choices, when information on the results of preceding choices can be taken into account. A strategy may be *pure* (each choice being uniquely determined, given the feedback from previous choices, if any), or *mixed* (involving a probabilistic choice between pure strategies).

Much game theory work (*classical* game theory as it will be called), makes strong assumptions about agents' rationality. Specifically, it assumes *instrumental rationality* — that agents act as if they have consistent preferences and unlimited computational capacity — and *common knowledge of rationality* (CKR), which means that each agent *assumes*:

- That all agents are instrumentally rational.

- That all agents are aware of other agents' rationality-related assumptions. (This produces an infinite recursion of shared assumptions.)

The assumption that agents act as if they have a consistent set of preferences (referred to as the *maximisation of expected utility* theory) requires that any single agents' *payoffs* (rewards or punishments) can be expressed and compared in numerical terms so (for example), a payoff of 3 would have the same value as a 50% chance of a payoff of -2 and a 50% chance of a payoff of 8. (It is not required that the payoffs of different agents be comparable in this way.) An instrumentally rational agent will act so as to maximise their expected payoff. There is experimental evidence that in at least some circumstances, people act in ways incompatible with this model. Specifically, they frequently violate the so-called *independence* axiom, which dictates that an agent's preferences between a given pair of options should be stable under specified types of contextual change (see Hargreaves Heap and Varoufakis (1995) [pp.13-14] for details). More broadly, maximisation of expected utility effectively assumes that *all* an agent's concerns are taken into account in all decisions. Most cognitive scientists, viewing human action as *goal-directed*, would find this assumption highly implausible. To quote an eminent cognitive scientist *and* economist (Simon, 1997) [p.324]:

> Neoclassical economic theory has nothing to say about how items are placed on the agenda for decision... In any system that can deal with only a limited number of problems at one time, there must be a mechanism to determine which problems will receive attention.

The key question is not whether instrumental rationality is a completely accurate model of how human beings make decisions, but when — or even whether — it is a useful approximation.

For a recent introduction to classical game theory see Gibbons (1997); for a critical overview, see Hargreaves Heap and Varoufakis (1995). Some recent work in game theory has broadened to include experimental evidence (Camerer, 1997), or evolutionary considerations (Maynard Smith, 1982; Hofbauer and Sigmund, 1998). In evolutionary game theory, the players are no longer taken to be rational. Instead, it is assumed that the strategies which are most successful at a particular time will have the best chance of being followed in future. Evolutionary game theory thus stresses adaptation, like Alife, but differs from it in stressing mathematical analysis rather than computational simulation. However, evolutionary game theorists can and do use simulation.

Game theory, like neoclassical economics, tends to assume that individual motivation is mainly or even exclusively selfish, and hence that altruism has no significant role in determining human behaviour. As Binmore (1994) [p.22] stresses, altruism is not actually incompatible with the *Homo economicus* model of neoclassical economics:

> Nothing says that *homo economicus* cannot be passionate about the welfare of others.

Nevertheless, Binmore, a game theorist whose underlying assumptions are declaredly those of neoclassical economics (Binmore, 1994) [p.15], argues that, outside close-knit groups such as families, tribes, street gangs and platoons, altruism with more than a trivial cost to the altruist is rare. He expects selective forces (biological or social) to suppress altruism beyond these limits. Several theoretical grounds for doubting this are explored in section 3. The assumption of fundamental selfishness has an important function in classical game theory and neoclassical economics, which often rely on the assumption that agents know each others' preferences, at least probabilistically. If there are significant numbers of marauding altruists going around helping others without thought of reward, this becomes much less plausible. These approaches also treat agents' motivations as *exogenous* — outside the scope of what is to be explained, and fixed for any particular system or phenomenon to be investigated. If different social systems can produce marked differences in the prevalence of altruistic behaviour, as seems to be the case (Roth, 1995; Henrich, 2000; Henrich et al., 2001), the explanatory autonomy of economics is compromised.

Genuine altruism also seems incompatible with Darwinism as originally conceived, according to which natural selection acts on individual organisms. As Darwin (1968) [p.133] put it:

natural selection can act only through and for the good of each being.

It should thus tend to suppress altruistic tendencies, if there is any heritable variation for them in a population. Darwin himself, according to Cronin (1991) [pp.325-8] found human altruism, specifically, difficult to explain. Section 3, however, discusses a number of explanations for altruism which Darwinists have put forward.

2.2. The Prisoner's Dilemma

Much of the ABSS literature on altruism and cooperation concerns variations on a particular two-player non zero-sum game, the *Prisoner's Dilemma* (PD). This phenomenon is by no means unique to ABSS: work on the PD and its variants is prominent in the literatures of economics, politics, anthropology, sociology, social psychology, philosophy and evolutionary biology.

According to Poundstone (1992), the PD stems from experiments reported in Flood (1952), although Miller (1996) states that it was first defined by Tucker (1950). The anecdote often used to explain it is not given here, but the game-theoretical description of the PD is as follows. Two players are each, separately, offered a choice between two options generally called *Cooperate* (C) and *Defect* (D). They cannot communicate with each other. If both choose C, each will receive a payoff of R *units* (the **R**eward for mutual cooperation). If both choose D, each will receive P units (the **P**unishment for mutual defection). If one chooses D and the other C, the defector (choosing D) will receive T units (the **T**emptation), the cooperator, S units (the **S**ucker's payoff). (Table 1 in the appendix lists the standard abbreviations for the PD moves and payoffs.) The PD requires that $T > R > P > S$, and it is often also required that $2R > T + S$ — see below.

Typical payoffs are:

- If both players choose C, $R = 3$ units each.
- If both choose D, $P = 1$ unit each.
- If one chooses C and the other D, the defector (D) gets $T = 5$ units, the cooperator, $S = 0$.

Both players are assumed to understand the entire setup, including the other player's options. If both try to maximise their own payoff regardless of the other player (as is also assumed), both will choose D; each will then do better than by choosing C, whichever choice the other makes. However, each will thus get less than if *both* had chosen

C. (One may object that the assumption of indifference to the other player's payoff does not reflect how real people feel. Support for this point can be drawn from experimental economics (section 2.4), but it does not affect the strategic situation in the PD as posed above.)

The choice of D by both players is the PD's only *Nash equilibrium* (Nash, 1951): a set of strategies such that no player, knowing the strategy of the other(s), could improve their expected payoff by changing their own. (Any n-player game with a finite number of possible pure strategies has at least one Nash equilibrium if mixed strategies are permitted.) In the case of the PD, D is a *dominant strategy* for each player: a strategy that cannot be improved upon, whatever strategy the other player adopts.

If the two players know they are to play some fixed number $n > 1$ of successive rounds of the PD against each other, a dominant strategy no longer exists. For example, if player 1 is a mind-reader and knows that in a two-round game player 2 will play C in the first round, and in the second will copy whatever player 1 did in the first, player 1 should play C then D. In this situation, of course, Common Knowledge of Rationality does not hold, since player 2's intended plays are not instrumentally rational. If both players are instrumentally rational and CKR does hold, on the other hand, both will play D in each round. All pairs of strategies in Nash equilibrium lead to this outcome (Luce and Raiffa, 1957) [pp.98-99]. Such a strategy specifies a probability p, $0 \leq p \leq 1$ that the player will play C in any given round after any specific history of previous rounds. Suppose the two players have strategies that would *not* necessarily lead to both playing D on every round, and that each knows the other's strategy. If the interaction of their strategies could involve either playing C in the last (nth) round, that player could improve their expected payoff by planning to play D in that round unconditionally, so the two strategies cannot be in Nash equilibrium. If the interaction of strategies could *not* lead to a play of C in the nth round, the players need only consider the first $n - 1$ rounds — but the same reasoning will apply to the last of these, and can be iterated backward over any number of rounds.

However, the strategic situation changes again when the number of rounds to be played is not limited in advance — if for example there is a fixed probability w $(0 < w \leq 1)$ that the sequence will continue after each round. In this case, the 'folk theorem' for repeated games[2] applies. This states that any *individually rational outcome* — one giving each player at least the mean payoff below which they cannot be forced by other players — can be a Nash equilibrium in an indefinitely repeated game if w is high enough (how high it must be depends on the game and strategies concerned). Figure 1 shows the set of possible Nash equilibria

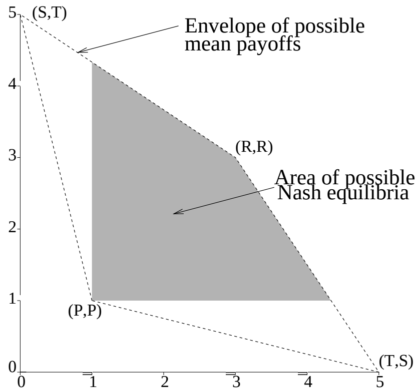

Figure 1. Nash Equilibria for the IPD

(within the envelope containing the possible outcomes over *all* pairs of strategies) for the indefinitely repeated PD.

From here on, 'IPD' (for *iterated prisoner's dilemma*) will be used for the indefinitely repeated version, and 'FRPD' (for *finitely repeated prisoner's dilemma*) for the version with a fixed number of rounds, while 'RPD' will be used to cover both. (The IPD when $w = 1$ implies an infinite sequence of games, an implausible but sometimes useful limiting case.) It is in the RPD that the secondary constraint on payoffs, $2R > T + S$, becomes necessary. If it does not hold, a second form of 'cooperation', in which each player plays C half the time, but the two coordinate so that each plays C when the other plays D, becomes as good as or better than both playing C every time. This complicates the strategic situation.

2.3. Social Dilemmas for Many Agents

The PD is a two-person social dilemma. The so-called *Tragedy of the Commons* (Hardin, 1968; Hardin, 1998) is a multi-person version: if all are free to graze as many cattle on a piece of land as they wish, it is always in each individual's interest to add another beast, but the result of all doing so is to ruin the pasture. As Monbiot (1994) notes, Hardin's name for this social dilemma is a complete misnomer. Historically, commons were used communally, but according to elaborate rules supported by systems of mutual surveillance; they disappeared only when enclosed by rich individuals with government support. What Hardin describes is an open access régime. The general name for a limited resource which many individuals can access, with or without rules to restrain overuse, is a *common-pool resource* or CPR.

Olson (1965), working within the economic theory of organisations rather than game theory, made another early contribution to the theory of multi-person social dilemmas. Olson argued that, in groups supposed to promote the common interests of their members (such as business cartels and trades unions), there will be a systematic tendency for the members to contribute too little to this common interest, or *public good* — in the sense that if all could be *forced* to contribute more, all would be better off. Olson argued that it will only be worthwhile (from a selfish standpoint, which he assumed to be general) for a particular member to contribute up to the level where a *unilateral* increase in their own contribution would be exactly balanced by the resulting increase in their *share* of the extra public good produced. The larger the group, the further the level of public good thus produced will tend to fall below a *Pareto-optimal* level. (A Pareto-optimal outcome to a game or socio-economic interaction is one to which no change can be made *without leaving at least one participant worse off.*) Olson argues that in order to operate effectively, large groups must employ coercion or selective incentives to oblige or persuade their members to contribute.

2.4. Empirical Evidence

Despite the assumptions implicit in neoclassical economics and classical game theory, and the explicit claims of economists such as Stigler (1981) and Mueller (1986), it is implausible that human beings are motivated solely by self-interest. Aside from everyday observation, there is considerable relevant work in experimental economics. This is aimed at investigating how far such considerations as fairness, group solidarity and envy produce behaviour at odds with the pursuit of self-interest. Generally, this work has found that other motives are indeed important. Most of it has placed the experimental subjects in a social dilemma of some kind, but some has used other experimental paradigms: Thaler (1992) [Ch.3] and Camerer (1997) summarise some results of this latter work.

The first experimental work on the PD itself — in fact concerning the FRPD — is reported in Flood (1958). The experimenters wished to discover whether people unfamiliar with the game would find the Nash equilibrium, as game theory suggests they should if rational (the two subjects knew that the game would be repeated 100 times). In fact, both played C rather more often than D. Their comments (given in Poundstone (1992) [pp.108-115], citing Flood (1952)) indicate that attempts to establish cooperation, and considerations of fairness, were important in determining the sequence of responses.

Although the Flood (1958) experiment used only a single pair of subjects, its results are qualitatively similar to those of most subsequent empirical studies of the FRPD, while even in one-shot PD experiments, levels of cooperation tend to be around 20% (Roth (1988), Routledge (1998)[p.109]). (There has been little experimental work on the IPD, perhaps because of the wide range of Nash equilibria in that game.) So far as the FRPD is concerned, a rational and selfish player might attempt to build a *reputation* for altruism (see above). Kreps et al. (1982) suggest such an explanation. However both Andreoni and Miller (1993) and Cooper (1996), comparing one-shot PD and FRPD, found C-moves in the one-shot PD that reputation building could not account for, and the latter found that neither pure altruism, pure reputation-building, nor a straightforward mix of the two could account for all their FRPD results.

Experiments on multi-agent social dilemmas concern the provision of public goods, or depletion of a limited resource — situations which are equivalent from a game theoretical viewpoint, but not necessarily so psychologically (Kahneman et al., 1982). For reviews, see van Lange et al. (1992), Ledyard (1995), and Gintis (2000a). Variations in both the individual payoffs, and the effect of an individual's actions on others' payoffs, influence results, suggesting mixed motives among the subjects. Communication concerning the experimental dilemma increases the level of cooperation. Ostrom et al. (1994) found that without communication, aggregate outcomes in finitely repeated common-pool resource decisions were not far from the (Pareto-suboptimal) Nash equilibrium (although *individual behaviour* was unpredictable). A single opportunity to communicate in advance made a considerable difference, and multiple opportunities to do so still more. Expectations that others will cooperate also raise the probability of cooperation, as does perceived ability to influence the overall outcome, while cooperation declines with group size up to about 7 or 8, but little if at all thereafter. Ledyard (1995) summarises the results of research as difficult to explain either as the outcome of general selfishness or of consistent altruism or feelings of group solidarity. Gintis (2000a), focusing on the fact that contributions in public goods games tend to start high and decline to near the Nash equilibrium, asks whether this could simply be a result of the subjects learning the game, but cites a number of studies telling against this, notably Andreoni (1988), who found that repeating the process with the same subjects restores initial levels of cooperation. Gintis suggests that 'public-spirited' subjects wish to retaliate against those acting selfishly, and in the experimental setup can only do so by withdrawing their own cooperation. Given an alternative, subjects punish selfishness in other ways, despite costs to themselves (Fehr and Gächter, 2000).

Field studies (Ostrom et al., 1994) confirm that frequently, although not always, groups finding themselves in real-world common-pool resource dilemmas are able to organise so as to restrain overuse. Key features of successful solutions appear to be a limit on who can draw from the resource at all, explicit rules for where, when and/or how the resource may be drawn on, active monitoring and sanctioning systems (generally run by the participants themselves), and graduated punishments for infractions.

Frank et al. (1993) suggest that economists are less likely than others to cooperate in social dilemmas, and that this may be due to their educational exposure to the *Homo economicus* model of human behaviour. Cadsby and Maynes (1998) indicate that this distinctively lower cooperativity persists in situations where there are *alternative* Nash equilibria. Perhaps those economists who regard self-interest as overwhelmingly more important than any other motivation are over-generalising from their knowledge of themselves and their colleagues?

2.5. Axelrod's RPD Tournaments

A key event in the history of both ABSS and PD research is the work reported and discussed in a series of papers by political scientist Robert Axelrod (Axelrod, 1980a; Axelrod, 1980b; Axelrod, 1981; Axelrod, 1984). Axelrod used the unusual research method of inviting entries to an FRPD 'tournament', in which each entry would play a sequence of 200 rounds against every other entry, against itself, and against a random strategy (playing C or D with equal probability). Recall that Nash equilibrium strategies in an FRPD game lead both players to play D in every round. Nevertheless, the winner of Axelrod's tournament was the *TIT FOR TAT* (TFT) strategy, entered by psychologist Anatol Rapoport. TFT plays C in round 1, and in all subsequent rounds plays whatever its opponent (or partner) has just played. (Table 2 in the appendix lists the main RPD strategies described in the text.) Playing itself, TFT achieves the R payoff in all rounds. Counterintuitively, it won the tournament although it can never score more than its opponent in a given FRPD game. Axelrod explains this as due to four properties: TFT is *nice* — never the first to play D, *provocable* — it retaliates if its opponent plays D, *forgiving* — it returns to playing C if the opponent does so, and *clear* — easily understood, so opportunistic strategies which 'probe' opponents for exploitability can quickly recognise its non-exploitability.

Axelrod held a second tournament after publicising the results of the first. Sixty-two strategies were entered, and each was pitted against itself, all the others, and the random strategy, in five games of varying lengths (these were the same for all pairs, and were fixed using a randomising method by Axelrod, but those entering strategies did not know what the lengths would be). TFT won again.

These two 'tournaments' fall somewhere between the categories of (unsystematic) empirical work on social dilemmas and ABSS, as simulation was used, but the set of strategies employed in the research was drawn from the intuitions of a self-selecting sample. However, work closer to 'mainstream' ABSS grew out of the second tournament, in which Axelrod found that most of the rules submitted fell into a small number of *types*. Most of the variation in strategies' performance could be accounted for in terms of their performance against these types.

Axelrod (1984)[Appendix A] calculated scores for six 'hypothetical tournaments'. These were statistical manipulations of the results of the second tournament, intended to indicate what would have occurred if the proportion of these types of strategy submitted had been different. TFT 'won' five of these, and 'came second' in the sixth. Axelrod (1984) [pp.49-52] also carried out what he calls an *ecological analysis* of the second tournament's strategies. This involved calculating the results of successive hypothetical tournaments, in each of which the notional proportion of the population using a strategy was determined by its success in the preceding round (so the proportions in the first of this series were determined by the results of the real tournament, the proportions in the second by the results of the first, etc.). Exploitative strategies did well in early rounds (although TFT held its lead), but lost ground as their 'victims' became rare. TFT did best, but did not eliminate the alternatives. A strategy's tournament score was an average of its expected score against all strategies in the population, weighted by their prevalence. This sequence of tournaments constitutes *numerical* simulation rather than ABSS, since neither players nor games were individually represented. Similar approaches were taken in many the studies discussed below.

In contrast, Axelrod (1987) used genetic algorithms (GAs) (Holland, 1992) to study the evolution of IPD strategies in a selective environment consisting of a set of eight 'representative' strategies drawn from the second tournament. This is most naturally viewed as an ABSS study, since each strategy tested was represented by a single member of a population of definite size. The *set* of strategies explored comprised all deterministic strategies able to take account of no more than the preceding three moves by both players: for each possible combination of its own and its opponent's last three moves, a given strategy's 'chromosome' specified C or D. Each member of the current population was tested against each of the eight representative strategies in a 151-move FRPD (the limit on memory length made this effectively equivalent to an IPD). Those that did best contributed most to the process of recombination and mutation producing the next generation. Most of the evolved strategies resembled TFT, and did about as well as TFT against the eight representatives. However, in 11 of 40 runs the median score for the final population exceeded TFT's score. In these 11 cases, the most successful rules, although able to establish reciprocal cooperation, were *not* nice: they began by playing D, and switched to C only if the opponent was not exploitable.

Axelrod also ran 10 simulations in which each of the 20 strategies in the current population was tested *against the other 19*. An initial trend away from cooperation, due to few reciprocating players existing, was succeeded by the spread of cooperation based on reciprocity.

Axelrod (1981) and Axelrod (1984) [Ch.3] also include mathematical analysis of the stability of TFT and other strategies. Axelrod defines a concept of *collective stability*: a strategy is collectively stable if and only if no other strategy can do better against it than it does against itself. Axelrod shows that TFT is collectively stable in the IPD if and only if $w \geq ((T-R)/(T-P))$ and $w \geq ((T-R)/(R-S))$. In general, a collectively stable strategy can be characterised as one which always defects on move n of an IPD game if its opponent's score over the first $n-1$ rounds is too great (see Axelrod (1981) for details and proof). This class is a broad one, and includes ALL-D, the strategy of playing D on every occasion. Axelrod (1984) also includes work on populations of IPD players which are *structured* in the sense that the probabilities that two given players will meet are not equal for all such pairs. This line of research is explored in section 6.

2.6. Classifying Research on Social Dilemmas

This subsection presents a summary of the ways in which the studies reviewed differ from each other — and conversely, the ways in which they could be grouped together. This summary is then related to the structure of the paper. Readers may find it useful to refer back to this summary to place particular studies or groups of studies in context.

1. In terms of method, we can distinguish empirical approaches, computational simulation, mathematical analysis, and discursive theoretical work. More than one of these may be used in a single study, and some methods, such as Axelrod's tournaments, are difficult to classify. Empirical work can be divided into experimental and field studies. Simulations may be numerical or agent-based, and the latter are distributed along a thick-thin continuum.

2. In discussing repeated games, we need a term for what is repeated: the minimal strategic interaction, in which each player makes a single decision. The standard term for this is the "stage game". This may be the PD, another 2-person game, or a multi-person

game. Within these categories, the exact payoffs used may differ between studies.

3. Given a stage game, an encounter between its two or more players may consist of a single occurrence of that game, a fixed number of repetitions, an indefinite number of repetitions (with some way to determine how many), or infinite repetition. All these possibilities have been discussed above for the case of the PD.

4. There may be contextual factors modifying the interactions between players. One possibility is “noise”, making either the players' actions, or their perceptions of others actions, unreliable. Another is the existence of computational costs that modify the payoffs.

5. Which pairs or sets of individuals take part in encounters may be determined in various ways: a round-robin tournament such as Axelrod's, random groupings among a population, or a schema involving spatial proximity or the agents' own choices, for example. (In the last case, these choices could be regarded as moves in a larger game.)

6. Holding all the factors above constant, studies may still differ in the set or space of strategies considered.

7. Agents may or may not be able to learn, altering their basic strategy according to a higher-level strategy. This may be implemented using a simple reinforcement schema or an ANN (artificial neural network), and may itself vary through second-level learning.

8. Schemas for assessing the success of strategies also differ. Many analytical studies look for strategies which can form part of a Nash equilibrium (of the stage game, a repeated game, or some meta-game. (A meta-game is a game derived from a simpler game by having one or more players choose a strategy for the latter in advance, and without the possibility of later deviation; one may think of these players as choosing a machine encoding the chosen strategy to play for them. The term originates with Howard (1971).) Axelrod's studies pioneered several other approaches: success in a single round tournament, in the repeated tournaments of an ‘ecological analysis’, or in an evolutionary process.

9. Similarly, studies (particularly evolutionary ones) may have different schemas for producing new strategies: none may be allowed to arise, or they may do so by simple mutation, or by use of a GA.

10. Studies differ in the underlying mechanism for producing or maintaining cooperation or altruism that they propose or investigate.

11. Finally, they differ in how far they focus on individual agents and the relative performance of the different strategies they may follow, how far on population-level phenomena such as the overall level of cooperative behaviour or strategic diversity, and how far (if at all) on the population's interaction with some larger system.

Section 3 describes underlying mechanisms that have been proposed as explanations for human cooperation and apparent altruism. The rest of the paper is structured primarily around the type of interaction between agents which studies investigate (the stage game, degree of repetition, and modifying contextual factors); and how the pairs or groups of agents taking part in each interaction are chosen. Sections 4-6 deal with 2-person interactions, mostly the RPD — the main exception is section 6.1, which focuses on the one-shot PD. Section 4 concentrates on criticisms of Axelrod's work on the RPD and developments from it, centred on notions of *evolutionary stability*, and on the relationship between ABSS and analytical work. Section 5 extends these themes with reference to the modifying contextual factors of noise and computational cost. Section 6 considers the PD (one-shot and repeated) in *structured* populations — those where some pairs of agents are more likely to play together than others. Section 7 turns to multi-agent social dilemmas, while 8 continues this theme, but concentrates on relatively thick simulations of *socio-ecosystems*: situations which embed a group of agents in an external environment that is affected by their actions and influences their payoffs. In general, concentration on ABSS work increases as the paper proceeds, and the situations considered become more complex and less abstract.

3. Theories of Human Cooperation and Altruism

Proposed scientific explanations for the occurrence of human cooperation and (apparent) altruism fall into three classes: eliminative, adaptive (the largest class), and constraint-based.

Eliminative explanations take two forms: minimising the frequency and extent of altruistic behaviour, and redefining 'altruistic'. Stigler (1981), Mueller (1986) and Binmore (1994) provide examples of the former. An example of the latter approach is the argument that parental care is not altruistic. As Cronin (1991) [p.264] puts it:

> But what should altruism include? Maternal care? Or is reproductive success so much a part of Darwinian self-interest that mothering counts as 'useful to the possessor'?

Such an approach is compatible with, indeed dependent on, other kinds of explanation. It draws attention to the importance of being clear whether 'altruism' is being used in something close to the everyday (and imprecise) sense, according to which much parental care is altruistic, or in some discipline-specific sense. In this paper, 'altruism', when used without qualification, means the unforced transfer of resources such as goods, money, time, or social prestige to another. *Resource altruism* will be used for any type of behaviour that would generally lead to an *unrequited* transfer of resources, regardless of the giver's intentions. *Motivational altruism* refers to behaviour *intended* to effect such a transfer, without expectation that it will be requited. Altruism in the simplest sense may belong to both these tighter categories, to neither, or to either one but not the other.

An *adaptive* explanation of some form of behaviour explains it in terms of the benefit it bestows on the behaving organism, or some entity associated with it. This is frequently linked to a selective explanation of the behaviour's persistence in terms of some form of environmental feedback which increases the probability that behaviours with favourable outcomes will occur again. The range of adaptive explanations of cooperative and altruistic human behaviour is wide.

1. Mutual dependence, or byproduct mutualism.
   A group of two or more agents can often achieve what none of them could achieve alone. Such cooperation may be entirely non-altruistic, and in considering it, attention shifts from *why* to *how* questions: how are shared goals determined, and actions coordinated? Byproduct mutualism will not be examined here.

2. Kin selection.
   The term *kin selection* was introduced by Maynard Smith (1964). Parental care (resource altruism towards offspring) is easily understood in evolutionary terms. Resource altruism *towards kin in general*, and not just towards offspring, can increase an organism's *inclusive fitness* (Hamilton, 1964): its contribution to the number of its own kin in the population, weighted by their relatedness. It can thus tend to increase the prevalence within the population of the alleles (gene variants) found in the altruistic organism and in particular, those favouring resource altruism toward kin.

   Inclusive fitness theory underlies a further sense of 'altruism', often used in evolutionary biology, which will be referred to here as *genetic altruism*: a type of behaviour is genetically altruistic if and only if it tends to *decrease* inclusive fitness. *If* natural selection is sufficiently effective in optimising behaviour, then genetically altruistic behaviour cannot persist in a population, above the level that can be maintained by random mutation (but see the discussion of group selection below for a complication). Cases of *apparent* genetic altruism therefore require explanation. Just as behaviour which is altruistic in the everyday sense may not be genetically altruistic, the converse also holds: for an infertile person to murder their entire family to gain their property would not generally be taken to exemplify altruism, but would be about as *genetically* altruistic as it is possible to be.

3. Reciprocal altruism.
Trivers (1971) argues that a tendency for unrelated individuals to help each other can be maintained in a population if giving help to an individual increases the probability of receiving it from that individual on future occasions. Trivers argues that reciprocal altruism should be favoured when opportunities for altruistic behaviour arise frequently, a given organism interacts frequently with the same small set of individuals (so those failing to reciprocate can be identified and refused further help), and pairs of individuals can render each other similar benefits. Reciprocal altruism may or may not be motivationally altruistic: Trivers' theory does not specify reciprocal altruists' motives. The point is that it is not *genetically* altruistic because it is not, on average, resource altruistic. When and how reciprocal altruism can persist is the focus of much of the work reviewed.

   Trivers argues that various features of human social relationships are due to the (genetic) costs and benefits of reciprocation. These include the complex mix of altruism and 'cheating' shown by most individuals, efforts to detect cheats and indignation against them, emotions such as gratitude, sympathy and guilt, friendship, developmental plasticity in the extent of altruism, and even generalised — i.e. not obviously reciprocal — altruism. It may be (genetically) advantageous to *appear* motivationally altruistic (or fair, or moral) and *one* way to appear so convincingly is to be so: restricting altruistic behaviour to cases where one is likely to be observed may entail risks and computational costs that outweigh the benefits gained.

   Many of Trivers's arguments are speculative. For example, he cites evidence (Trivers, 1971) [p.49] that:

> the scarcer the resources of the donor... the greater the tendency of the recipient to reciprocate.

This is explained in terms of selection favouring greater gratitude when the cost to the donor is high, but the opposite finding could have been put down to selection favouring greater reciprocation towards those best able to counter-reciprocate in future!

Trivers (1971) [pp.53-54] in effect attributes human altruism to a combination of socio-cultural and biological causes. In particular, cross-cultural *differences* in altruistic behaviour cannot be attributed to universal features of human biology, except in the sense that human developmental plasticity makes them possible.

4. The handicap principle.
   The handicap principle (Zahavi, 1975; Grafen, 1990b; Grafen, 1990a; Zahavi and Zahavi, 1997) concerns *signalling* between organisms, and in particular the circumstances in which *honest* signalling will arise and persist. The fundamental idea is that to be reliably honest, signals must be *costly* to the signaller. Initially it was proposed (Zahavi, 1975) to explain features such as the peacock's tail, which appears to increase predation risk, but attracts potential mates and intimidates rivals. Zahavi argues that such signals are costly for all individuals, but relatively more so for those otherwise less fit, so providing a reliable index of general fitness. Initially met with scepticism, the principle is now widely accepted as having a role in the evolution of behaviour, thanks to mathematical work (Grafen, 1990b; Grafen, 1990a; Siller, 1998; Grafen, 1998). Zahavi and Zahavi (1997) [Ch.12] propose that much altruistic behaviour in social species exemplifies the handicap principle: the altruist gains social prestige (and consequent advantages) by demonstrating that it is fit enough to give away resources. They briefly apply their approach to human altruism and moral behaviour, and note (Zahavi and Zahavi, 1997) [p.227] that the social scientist Veblen (1899) in some ways anticipated their explanation both of altruism and of many 'risky' and 'wasteful' human behaviours as prestige-seeking. (The Zahavis note that human altruism may be sincere, at the same time as being advantageous by tending to raise the altruist's prestige.)

5. Norms and social 'docility'
   Turning to explanations that are specifically cultural, much human social behaviour is influenced by imitation of others, persuasion, instruction or manipulation by others, or some combination of these factors. This is not just a matter of effects on overt behaviour: guilt, shame, sympathy, self-esteem, approval or disapproval of others

and their actions and the desire to impress others all influence behaviour, and are in turn socially influenced. In short, much human behaviour is influenced by *norms* specifying right or appropriate ways to behave (see Ullmann-Margalit (1977) and Conte and Gilbert (1995) [Part 4] for discussion of norms relevant to this review). Frequently, such norms enjoin altruistic or cooperative behaviour. However, to the extent that human altruism and cooperation result from the existence of norms, the existence and (partial) effectiveness of such norms itself requires explanation.

Simon (1997) [p.244] defines a related concept, *docility* as:

> a tendency to accept knowledge and advice that is transmitted through social channels.

He stresses that it does not imply passivity or tractability, and proposes that human beings genetically predisposed to docility are, in general, fitter than those who are not, as they acquire skills and knowledge that would not otherwise be available to them. Together with the notion of *bounded rationality* — that human beings are limited in knowledge and computational power, hence often unable to determine whether particular beliefs are true — Simon argues that the hypothesis of docility can explain altruism beyond what can be accounted for by kin selection or reciprocal altruism (Simon, 1997) [pp.244-245]:

> The society can now impose a 'tax' on this bonus to fitness by influencing docile persons... for example, to be brave soldiers or to limit the sizes of their families.

Closely connected with docility, if not a variety of it, is the human tendency towards *group identification*: adopting the interests of social groups (family, tribe, nation, organisation) as our own, and sometimes sacrificing individual welfare for them.

6. Meme selection. The term *meme* was coined by Dawkins (1976) [p.206], for what might be called *units of cultural transmission* or *cultural replicators*:

> Examples of memes are tunes, ideas, catch-phrases, clothes fashions... Just as genes propagate themselves in the gene pool... so memes propagate themselves in the meme pool by leaping from brain to brain via... imitation.

If this idea is sound, memes for altruistic or altruism-promoting behaviours might exist and spread. Blackmore (1999) [Ch.13] suggests that as altruists tend to be popular, and people tend to imitate those they like, any form of behaviour common among altruists has a selective advantage in the meme pool. However, genes are specific sequences of DNA bases, and are accurately replicated. 'Memes', if regarded as stored in the brains of culture-bearing organisms, have a physical basis which is not well understood (but may well be highly distributed, subject to interference by other memories, and different in each bearer), and are 'replicated' via complex, highly context-dependent processes of cultural transmission. There are serious problems with regard to whether transmissable aspects of culture are atomisable, and how cultural transmission works (Midgely, 2000; Ingold, 2000). However, the 'meme' idea does draw attention to the possibility that the spread of behaviour by imitation may occur for reasons which have nothing to do with advantage to the imitating agent.

7. Functionalism in social science, and group selection.
   There is a long tradition of attempting to understand (and justify) social phenomena in terms of their *function* in maintaining the social system they belong to, often linked with *organicist* views of social systems as similar to organisms. Functionalism of a sophisticated kind does not suppose that identifying the function of a social phenomenon explains its existence, as the following statement (Durkheim, 1964) [p.95, emphasis in original] illustrates:

   > *When, then, the explanation of a social phenomenon is undertaken, we must seek separately the efficient cause which produces it and the function it fulfils.*

   Nevertheless, Durkheim holds that social facts are *external* to any individual and cannot be reduced to facts about individuals — just as biological facts about a cell cannot be reduced to facts about the atoms it comprises. Functionalist views contrast with the *methodological individualism* of most economists, and many social scientists and philosophers, such as Popper (1966) [p91]:

   > the 'behaviour' and the 'actions' of collectives, such as states or social groups, must be reduced to the behaviour and to the actions of human individuals.'

   There are approaches to this 'micro-macro' problem in social science distinct from both functionalism and methodological individualism (Conte and Gilbert, 1995).

   In evolutionary biology, the functionalist/individualist dispute has a parallel in the dispute about *group selection*: the notion that natural selection sometimes acts at the level of groups of organisms,

> and in particular may increase the frequency of traits of individual organisms which favour social groups they belong to, even if these traits disadvantage the individuals bearing them. Despite the wide acceptance of arguments (Williams, 1966) that it will be important only in very unusual circumstances, some forms of group selection have recently gained considerable although by no means universal acceptance (Wilson and Sober, 1994; Sober and Wilson, 1998), among them the possibility that behaviour which is genetically altruistic *within a social group* may still be selected for *because of the success of groups in which such behaviour is more common.* Even some who dispute that most of the processes called 'group selection' are best described in that way agree that it may be important in the human case, because social norms can act to reduce intra-group fitness differences (Maynard Smith, 1998). Explanations of human altruism in terms of norms and of group selection may thus complement each other.

Constraint-based explanations of altruism have in common the idea that adaptive forces are of limited power: they can only work on the 'raw material' available, and only at a limited speed. Gould and Lewontin (1979) argue that evolutionary biologists have too readily assumed that natural selection is a highly efficient optimising agent. They suggest alternative kinds of explanation for particular features of organisms. The most relevant here is *genetic drift* — the random fixation of neutral or even unfavourable alleles in small populations.

Buss (1999) [p.20], like many other sociobiologists and evolutionary psychologists, argues that cultural change over the past 50,000 years or so has been too fast for natural selection to have kept pace. Altruistic tendencies that were adaptive in small, closely related groups during the Pleistocene might survive in modern mass societies, where an individual's social contacts are much more numerous and on average, much less closely related. While Gould and Buss have very different views about human evolution, they agree that historical constraints mean that non-adaptive traits may be important in modern human populations.

Another form of constraint-based explanation emphasises *trade-offs*; economics, evolutionary biology and computer science all use explanations of this kind. Calculating the best course of action in time to follow it requires computational resources that have associated costs, as explored in section 5. Simon's explanation of altruism in terms of docility and bounded rationality (Simon, 1997) has an important constraint-based aspect.

Human cooperation and altruism have been explained in a wide variety of ways. There is no reason, except perhaps parsimony, to assume that a single mechanism is wholly or largely responsible for the range of phenomena to be accounted for, and the true explanation may involve complex interactions between biology and culture, between adaptive processes and constraints, and between sub-individual, individual and supra-individual levels. However, the great majority of both analytical and simulation studies of social dilemmas have focused on adaptation, and specifically on reciprocal altruism.

4. Evolution and Stability in the Prisoner's Dilemma

In terms of the classifications of 2.6, this section concentrates on the RPD and on reciprocal altruism. Different ways of assessing the success of strategies are a major topic.

Axelrod's work, described in section 2.5, has been subjected to severe criticism by some game theorists, notably Ken Binmore (Binmore, 1994; Binmore, 1998b; Binmore, 1998a). Binmore (1998b) recognises Axelrod as a pioneer of *evolutionary equilibrium selection* — using evolutionary considerations to choose among multiple Nash equilibria — but says:

> to recognise Axelrod as a pioneer in evolutionary equilibrium selection is to endorse neither his claims for the strategy TIT-FOR-TAT, nor his unwillingness to see what theory can do before resorting to complicated computer simulations.

Binmore's criticisms concentrate on the proper relationship between simulation and mathematical modelling. The points at issue, subsequent work inspired by Axelrod's investigations, and relevant work in evolutionary game theory, thus require attention here.

Binmore (Binmore, 1994; Binmore, 1998b) regards Axelrod's claims for the performance of TFT as overblown: while TFT did better than any other single strategy in Axelrod's ecological analysis, it made up little more than 1/6 of the population. More fundamentally, he criticises the set of strategies used as arbitrary, preferring the work of Linster (1992). Linster used the 26 strategies (including TFT) that can be represented as one-state or two-state *Moore machines*: finite-state machines (with a distinguished starting state) in which each state has an associated output (the move C or D in the case of machines for RPD strategies), and transition arcs to the next state, each labelled with one of the possible inputs (the partner's possible moves). An arc may lead back to the same state, or to another, and both arcs may lead to the same state. (Figure 2 shows Moore machines for some RPD

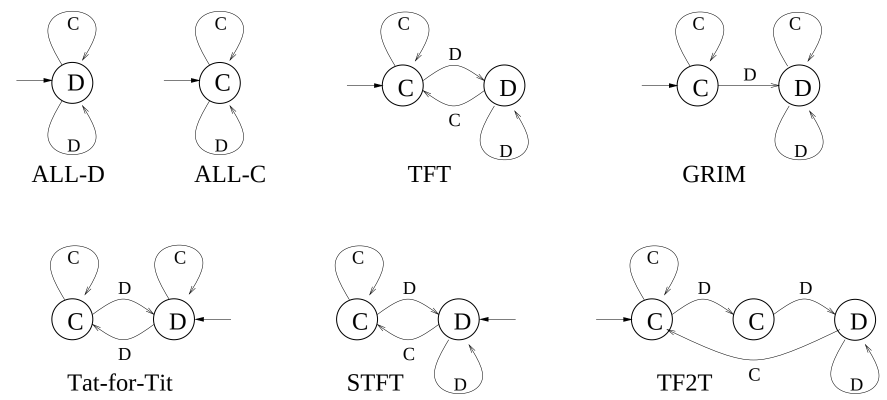

Figure 2. Moore machines for RPD strategies

strategies referred to in the text.) Given the sequence of inputs, a Moore machine is deterministic. ‘Nice’ strategies predominated in Linster’s final populations, but the most successful was ‘GRIM’, which plays C unless and until its partner plays D, when it switches permanently to D. Binmore also cites unpublished work by Probst (1996), who used GAs to evolve strategies expressible as finite automata with up to 25 states. TFT did well initially, but was then displaced by non-nice strategies. The simplest of these, ‘Tat-For-Tit’, begins with D, then switches from D to C or C to D whenever its opponent plays D.

Binmore’s most serious criticism of Axelrod, however, relates to the relationship between mathematical analysis and simulation. He accuses Axelrod of over-interpreting simulation results due to insufficient knowledge of game theory, and of ignoring game-theoretical commentary on his work. Binmore (1994) [p.200], citing Nachbar (1992), points out that Axelrod’s ecological analysis *could not* converge on anything other than a Nash equilibrium of the two-player game in which each player can adopt any of the 63 IPD strategies entered in the tournament, and that the results after 1000 iterations approximated a mixed strategy Nash equilibrium of this kind. Which such equilibrium results would depend on the proportions of the population each strategy began with (Axelrod started them off with equal proportions, which is an obvious choice, but has no particular *theoretical* justification).

However, the weakest point of Axelrod’s work, in theoretical terms, is that he underestimates the difference between his notion of a collectively stable strategy, and the *evolutionarily stable strategy* (ESS) concept developed by John Maynard Smith (Maynard Smith and Price, 1973; Maynard Smith, 1974; Maynard Smith, 1982). As Binmore (1994) [p.197] notes, Axelrod (1984) [p.217] is wrong in claiming that for ‘nice’ strategies, the two are equivalent.

Maynard Smith gives several distinct definitions of an ESS, which are not wholly consistent. Maynard Smith (1982) [p.10] says:

> An ESS is a strategy such that, if all the members of a population adopt it, then no mutant strategy could invade the population under the influence of natural selection.

Maynard Smith and Price (1973) [p.15] give an equivalent definition couched in terms of reproductive fitness. However, Maynard Smith (1982) [p.204] gives a version which removes a qualifying phrase, producing a significantly stronger formulation:

> An ESS... is a strategy such that, if all the members of a population adopt it, then no mutant strategy can invade.

Maynard Smith (1982) [p.5] stresses that in any analysis within evolutionary game theory, it is important to specify the set of possible strategies: so calling a strategy an 'ESS', for Maynard Smith, always involves at least the implicit specification of such a set.

Maynard Smith also gives criteria for *identifying* a strategy as an ESS, couched in terms of equations and inequalities specifying the payoffs (in terms of reproductive fitness) to different strategies. The most widely used, given in Maynard Smith (1982) [p.14], uses $E(A, B)$ for the expected payoff to a player of strategy A encountering a player of strategy B. For an infinite population, with asexual inheritance, and the fitness of individuals determined by one or more pairwise contests against randomly selected opponents, a (pure or mixed) strategy I is then an ESS if and only if, for any other strategy J, either:

1. $E(I, I) > E(J, I)$, or
2. $[E(I, I) = E(J, I)] \wedge [E(I, J) > E(J, J)]$.

That is for any J, I must either do better against itself than J does, or it must do as well, *and* do better against J than J does against itself. This fits with the stronger of the two formulations in terms of uninvadability rather than the weaker. The stronger, in line with most subsequent usage, will be taken as standard here. Maynard Smith (1982) [p.107] introduces a weaker notion, that of a *neutrally stable strategy* (NSS), which is one such that if almost all members of a population adopt it, there is no alternative strategy that would produce *greater* fitness if adopted by a small number of mutants or immigrants. This corresponds well to the weaker definition of an ESS given above. The mathematical criterion above can be weakened correspondingly, giving a way to identify an NSS:

1. $E(I, I) > E(J, I)$, or

2. $[E(I,I) = E(J,I)] \wedge [E(I,J) \geq E(J,J)]$.

For I to be collectively stable, it is necessary and sufficient that $E(I,I) \geq E(J,I)$ (a strategy is collectively stable if and only if it is a best reply to — in Nash equilibrium with — itself). Thus a collectively stable strategy need not be an NSS, let alone an ESS. Maynard Smith (1982) [pp.202-203] actually gives a supposed 'proof' that TFT is an ESS if w is high enough. This cites Axelrod (1981), which in fact shows only that TFT is collectively stable for sufficiently high w (see section 2.5). Under these circumstances, only nice strategies do as well against TFT as TFT does against itself. This means TFT is an NSS under these circumstances — no isolated mutant can do *better* than the TFT-playing majority — but it is not an ESS. In fact *no* nice strategy can be an ESS: any other nice strategy is a potential invader, because an ESS must have a selective advantage against any possible invader, but any two nice strategies will cooperate all the time (as pointed out by Selten and Hammerstein (1984), in response to whom Maynard Smith (1984) retracted the claim that TFT is an ESS). The difference is significant because a population in which all members are playing TFT can be invaded by players of other 'nice' strategies by *random drift*: the number of agents playing other nice strategies can increase just by chance, and the former state of the population will *not* be restored by selection, since all members of the population get the same payoffs. However, this can make the population vulnerable to invasion by other, non-nice strategies, under the influence of natural selection. Suppose, for example, that a 'mutation' to the strategy of always playing C ('ALL-C') occurred, and drifted to high frequency in the population. A mutant playing the opposite 'ALL-D' strategy of always defecting could then invade the population by exploiting the 'ALL-C' players.

Important theoretical results regarding evolution in the IPD appear in a series of three papers (Boyd and Lorberbaum, 1987; Farrell and Ware, 1989; Lorberbaum, 1994), which together demonstrate that there is *no* ESS for the IPD if w is high enough but < 1 (if w is sufficiently close to 0, ALL-D is an ESS). Boyd and Lorberbaum (1987) show that no *pure* strategy S_s (in which a player's choice of C or D always depends *deterministically* on the past history of the game) can be an ESS for high w values. The fundamental idea is that it is always possible to construct two strategies S_1 and S_2, where S_1 plays against S_s exactly as S_s does against itself, but does better against S_2. The existence of S_1 and the fact that it acts differently from S_s against S_2 actually suffice to show that no pure strategy is an ESS. The fact that S_1 and S_2 can always be chosen as required shows that invasion by a combination of alternative strategies is a real possibility. For example, in a largely-TFT population with small proportions of *Suspicious Tit for Tat* (STFT) — which defects in round one then behaves as TFT — and *Tit for Two Tats* (TF2T), which only defects after two consecutive defections by the opponent, TF2T has the highest mean payoff. Farrell and Ware (1989) extend the proof to all mixed strategies which are mixtures of a finite number of pure strategies, or equivalently, all those for which *some* sequence of pairs of moves cannot arise in contests between pairs of players both using the strategy. Lorberbaum (1994) uses a somewhat different approach to extend the proof to the remaining possibilities (the completely probabilistic strategies for which any such sequence can arise), so long as $w < 1$; here, the case in which $w = 1$ remains open.[3]

Tesfatsion and Ashlock (1998) include a brief reply to Binmore (1998b) from Axelrod, which does not directly address Binmore's charge that Axelrod pays insufficient regard to theoretical results, and elides the difference between the specific strategy TFT, and reciprocity in general. Axelrod's reply does, however, pinpoint an underlying difference in approach between Binmore and himself. Axelrod notes that standard game theory would direct the entrants to his first tournament, who knew the length of the game, to enter ALL-D. None did so, and were right not to, since no-one else did either. Axelrod's main interest is in how people *do* behave, rather than in how they *would* behave if they were completely rational (and assumed the same of those they interact with), or at the end of a hypothetical process of evolution or learning. However, Axelrod (1980a) [p.8] reveals that entrants to Axelrod's 'first' tournament were sent a description of a *preliminary* tournament in which TFT did well — and presumably knew that those devising other entrants would have the same information. There is no doubt that Axelrod's work itself stimulated mathematically-based research. However, his tournaments, and subsequent 'ecological analysis' and evolutionary simulations do not justify the claims made for TFT's 'robustness', and relevant results established by mathematical analysis must be taken into account if simulation-based studies are to be optimally designed, and correctly performed and interpreted.

Bendor and Swistak (1997) reinforce this point, proving results which in important respects justify Axelrod's intuitions, while underlining his errors and over-simplifications. Their work also draws attention to the complex relationship between ESSs or NSSs as discussed above, and the stability properties of *populations* using a variety of strategies. The ESS and NSS concepts specify the outcome of invasion by a *single* alternative strategy (drawn from some prespecified set); the fact that a strategy is an ESS within a given set of strategies does not mean that it will necessarily be able to resist invasion by two or more of the other members of the set simultaneously: Maynard Smith (1982) [pp.183-185], and Hofbauer and Sigmund (1998) [pp.75-76] provide examples.

Bendor and Swistak (1997) also stress the need to consider, in the context of evolutionary game theory, exactly how the relative success of strategies in terms of payoffs is linked to changes in their frequency in the population. This is particularly important in considering how evolutionary game theory relates to the study and simulation of social and cultural change.

By far the most extensively studied way to link payoffs and changes in frequency is the approach variously known as *Taylor-Jonker dynamics* (Taylor and Jonker, 1978), the *replicator dynamics* (Hofbauer and Sigmund, 1998) [Ch.7], or the *proportional fitness rule* (PFR) (Bendor and Swistak, 1997). In this approach, *changes* in the frequency of strategies in a population are proportional to the fitness of individuals using that strategy minus the mean fitness of the whole population. There are both discrete and continuous versions, expressed in difference and differential equations respectively. The former, which will suffice here, can be given the form:

$$p'_{S_i} = p_{S_i}\frac{V(S_i)}{\overline{V}}$$

where p_{S_i}, p'_{S_i} are the proportions of the population following strategy S_i in successive generations, $V(S_i)$ is the fitness of type S_i, and $\overline{V}$ is the population mean fitness. This is a deterministic approach, relying implicitly on the population being large enough to make stochastic effects unimportant.

As Bendor and Swistak (1997) observe, the replicator dynamics has a clear justification for biological applications of evolutionary game theory, so long as we can identify fitness with some linear function of expected game payoffs. (There are, however, complications when a sexually reproducing population is considered (Hofbauer and Sigmund, 1998) [Ch.22]). In social science applications, the justification for using the replicator dynamics is much less clear, despite the model of social learning proposed by Gale et al. (1995). Individuals may switch strategies as a result of individual experience, or learning from other members of the population and might (for example) use an 'imitate the winner' approach, switching only to the most successful strategy if they switch at all, so that some strategies yielding above-average payoffs fail to grow. Bendor and Swistak (1997) define a class of (deterministic) *evolutionary dynamics* to which the replicator dynamics belongs, but 'imitate the winner' dynamics would not: for any member of this class and any two strategies S_i, S_j:

$\frac{p'_{S_i}}{p_{S_i}} \geq \frac{p'_{S_j}}{p_{S_j}}$ if and only if mean payoff from S_i is at least equal to that from S_j.

(Evolutionary dynamics in this sense does not allow the proportions of strategies to change by drift, nor permit novel strategies to arise by mutation, recombination, or immigration.) They then derive results concerning PFR, and others concerning evolutionary dynamics in general, for iterated versions of a class of *games of cooperation.* This class includes the PD along with an infinite variety of other two-person games, many of which involve a social dilemma. All the results derived assume that the set of available pure strategies is fixed and finite, and that all available mixed strategies are mixtures of members of this set.

The most significant findings in Bendor and Swistak (1997) for our purposes[4] are:

1. Defining an *unbeatable* strategy as one giving an expected payoff at least as good as any other strategy provided its share of the population is sufficiently large, no *pure* strategy is unbeatable in any iterated *nontrivial* game, if w is set sufficiently high. (In a trivial game, each player has a single move which maximises their payoff whatever the other does.) This is a generalisation of the result of Boyd and Lorberbaum (1987).

2. Under the PFR, a population playing NSSs for the IPD can, for sufficiently high w, show *any* degree of cooperation.

3. Under the PFR, there are NSS strategies in the IPD which support any degree of cooperation greater than 0, but which are (unlike TFT) *infinitely exploitable* — meaning there is some other strategy which would outscore the exploitable strategy by amounts exceeding any specified limit, in a long enough game.

4. In any iterated game of cooperation, no strategy has a *minimal stabilising frequency* $< .5$ under any evolutionary dynamic. (The minimum stabilising frequency is the smallest frequency that, once reached, will always be maintained.)

5. For any iterated game of cooperation, under PFR, there is a nice and retaliatory strategy (such as TFT in the IPD) with a minimum stabilising frequency tending to .5 (from above) as w tends toward 1. Furthermore, with such a strategy entrenched, the population converges toward universal cooperation.

6. Any strategy meeting the criteria of the previous item must be *almost nice* — cooperate with itself almost all the time — and

almost retaliatory — set a finite limit on the extent to which any other strategy can outscore it.

Points 1-3 underline errors in Axelrod's conclusions: no strategy is an ESS for the IPD, and an NSS — the best achieveable — need be neither nice nor retaliatory as defined by Axelrod. The last three points, however, go some way toward justifying his intuition that nice and retaliatory strategies do have important advantages in terms of stability: taken together, they show that the strategies that achieve stability at the lowest frequency must at least be very close to being nice and retaliatory.

In conclusion, Axelrod's work, combining ABSS, analytical and an original (and controversial) form of empirical investigation, is both seriously flawed, *and* very fertile, giving a great impetus to analytical (and further ABSS) work on the RPD and related games. The analytical work described in this section depends on assumptions about the dynamics of competition between strategies that hold exactly only in highly idealised models of biological evolution (with infinite, asexually reproducing populations), and are much less plausible as good approximations in social than in biological contexts.

5. Dealing with Noise and Computational Costs

The work described in section 4 makes two implicit assumptions about the context in which RPD strategies compete: that the relative computational costs of following different strategies can be ignored, and that the environment is noiseless, so a player always moves as intended, and correctly perceives their partner's move. Changing either assumption makes a big difference — but the precise results depend on the exact change made. This section concerns the effects of noise and computational cost. The studies described use varying methods (analysis, numerical simulation, and ABSS), consider different ranges of strategies, and use different ways of assessing that success. Some allow the introduction of new strategies using a GA, others do not.

With regard to computational cost, the main line of investigation involves representing strategies by Moore machines. Most of this work is purely analytical. Papadimitriou and Yannakakis (1994) show that for the FRPD, cooperation can be stable between two players who probabilistically select finite automata to play for them. Specifically, a Nash equilibrium exists giving each player an expected score above $R - \epsilon$ for any $\epsilon > 0$ provided at least one player is limited to automata with a maximum number s of states, where s rises exponentially with the number of rounds in the FRPD.

For the IPD Rubinstein (1986), Abreu and Rubinstein (1988) and Banks and Sundaram (1990) treat Moore machines as strategies available to 'meta-players' of the IPD, who prefer 'less complex' machines among those giving the same expected return from the IPD. Different ways of assessing complexity turn out to give different sets of Nash equilibria for the meta-game. Binmore and Samuelson (1992) judge the superiority of one automaton over another first in terms of expected return, then in terms of their number of states, but give their work an interpretation in terms of social evolution — on the basis that people do not think deeply about their interactions, but are hosts for memes. They find that only repeated cooperation survives as a possible equilibrium outcome, and prove that certain Moore machines would be NSSs of the meta-game. However, they note that social evolution may involve frequent mutations, contrary to the standard assumption of evolutionary game theory that one mutation can be 'dealt with' before another arises, and find that long-term outcomes will depend on how mutations arise. Ho (1996), using an ABSS approach and a GA, finds that cooperation does not emerge from an initially defecting population if automata with more states are penalised, but does do so if the penalty is levied instead according to the frequency of switches between states. Both analytical and ABSS studies, therefore, show that the results of imposing costs on the complexity of strategies depend on how this is done.

Work on the effects of noise has been extensive. Most of it has concerned 'misimplementation' noise (making the wrong move), but 'misperception' noise (perceiving the partner's move inaccurately) has also been studied. Axelrod (1984) [pp.182-3] mentions a rerun of his second tournament with a 1% chance of a move being misperceived, which TFT won. Analysis, however (Molander, 1985) shows that in the IPD, TFT's performance against itself deteriorates to that of a pair of random players at any level of noise; a single error in an encounter between two TFT players sets off an unending chain of alternating defections which only a second error can end.

In a key paper, Boyd (1989) shows that the presence of misimplementation noise allows ESSs in the IPD, since:

> If individuals make mistakes, pure strategies can be evolutionarily stable because certain patterns of mistakes allow a strategy to be a uniquely best reply against itself.

A strategy can be the *unique* best reply to itself only if every finite sequence of moves by the two players has non-zero probability.

Specifically, ALL-D is not only an ESSs, but an 'invincible strategy': one that does better than any member of any set of strategies invading in small numbers. So, in some circumstances, is CTFT (Sugden, 1986).

This is like TFT except that the player shows 'contrition' by following an accidental play of D with two unconditional Cs. ALL-D is invincible at any non-zero error rate. For CTFT the maximum rate depends on payoffs and the value of w, but can be quite high.

In a series of papers (Nowak and Sigmund, 1989; Nowak, 1990b; Nowak, 1990a; Nowak and Sigmund, 1992; Nowak and Sigmund, 1993; Boerlijst et al., 1997), Martin Nowak and colleagues use mathematical analysis and numerical simulation to explore IPD contests between players with stochastic strategies, which they regard as biologically more plausible than deterministic ones. Initially (Nowak and Sigmund, 1989; Nowak, 1990b; Nowak, 1990a; Nowak and Sigmund, 1992) they investigated strategies which, after the first move, took account only of the partner's last move (not of the player's own). With both deterministic and stochastic strategies permitted, and with $w = 1$ (so the effect of the first move becomes negligible) there are no ESSs relative to this set of strategies (i.e. none that can repel, by selection pressure, an invasion by *any other member of the group*). Only strategies which always respond to C with C can be NSSs relative to the set. If deterministic strategies are excluded, however, (equivalent to introducing an irreducible minimum of misimplementation noise), the stochastic version of ALL-D (making occasional mistakes) is a unique best reply to itself (unlike any other strategy), and hence an ESS relative to the strategy set. Numerical simulations showed that most mixtures of 100 strategies from the set swiftly gave rise, in an Axelrod-type 'ecological' tournament, to populations dominated by strategies close to ALL-D. However, if at least one was close to TFT, this generally survived and displaced ALL-D, only to be displaced in turn by the strategy closest to GTFT ('generous Tit for Tat'), which responds to C with C, and responds to D with C 1/3 of the time.

Turning to strategies able to take account of the player's own last move as well as the partner's (Nowak and Sigmund, 1993), and allowing mutants to arise, strategies close to one called WSLC here[5] (switch from C to D or vice versa if and only if the opponent played D), usually dominated in the long run. This domination was preceded by a pattern of apparent stability interrupted by rapid shifts from C to D or the reverse, and between predominant strategies. If *misperception* noise is added, numerical simulations (Boerlijst et al., 1997) show that some ESSs, including ALL-D, can be invaded if there is a lower limit on the proportion of the population using the mutant strategy, as the mutants' scores against each other become significant. They describe a strategy resistant to both forms of noise: this is like WSLC except that after a defection by either player it returns to cooperation only after two rounds of mutual defection.

Others have taken work on strategies related to WSLC in other directions (Kraines and Kraines, 1993; Kraines and Kraines, 1995; Wu and Axelrod, 1995). Wu and Axelrod (1995) tested WSLC and a more generous version of it in the environment of Axelrod's 63 strategies, and with varying amounts of misimplementation noise, finding that both did poorly, while GTFT did better and CTFT best. Kraines and Kraines (1995) note that WSLC fails to adapt to different payoffs, unlike people (who cooperate much more if payoffs are $T = 10, R = 9, P = -1, S = -10$ than if they are $T = 10, R = 1, P = -1, S = -10$). They investigated a set of 'Pavlov learning strategies' P_n, where P_n is the *set* of sub-strategies $P(i,n) : 0 \leq i \leq n$ — $P(i,n)$ means 'play C with probability i/n — together with a procedure for switching between substrategies. A payoff of p units (where p may be positive or negative) causes an increase or decrease of p/n in the probability of playing the move just played (of course, this probability cannot become negative, nor exceed 1). Analysis showed that (with payoffs $T = 2, R = 1, P = -1, S = -2$) mutual cooperation always appears given long enough, and reappears after errors. P_2 and P_3 do better than P_1 (i.e. WSLC) against a wide range of strategies. In a population of the stochastic strategies P_n, faster learners will gradually replace slower and P_2 will eventually dominate, but those learning too much faster than most of the population will over-adapt.

Lomborg (1996) used a 67-round FRPD, but the deterministic strategies investigated remembered only the opponent's last three moves, making this effectively equivalent to an IPD. The work focused on misperception noise. A population of size 2^{20} was used, limited to containing 20 different strategies at a time. In each 'round', the expected payoffs of each strategy were calculated, and a percentage of the population switched to a mutated version of a randomly selected agent's strategy if that agent had done better. Starting from an ALL-D population, this procedure generally produced a meta-stable 'nucleus' of highly cooperative strategies and 'shield' of more cautious ones which repel invaders.

While some of the papers discussed above used numerical simulation, representing a population in terms of the proportions of its members following different strategies, none used an agent-based approach, representing individual agents and their interactions. Those using such an approach to investigate the dynamics of populations playing the RPD in the presence of noise generally combine the agent-based approach with use of a GA, as Axelrod (1987) did. Most of this work has appeared within the last few years.

Miller (1996) using a FRPD with payoffs $T = 5, R = 3, P = 1, S = 0$ and misperception noise, evolved a population of Moore machines limited to 16 states, using a GA with crossover and mutation (taken to represent imitation and innovation respectively). Each generation played a round-robin tournament. Average payoff fell, rose to a higher level than at the start, then leveled off for all conditions, but declined with increasing noise. Defection-reciprocity (returning D for D) was higher than cooperation-reciprocity. Noise made the survival of new strategies easier.

Macy (1996) challenges the use of GAs to represent processes such as imitation and instruction, and more broadly the idea that these processes will show dynamics similar to natural selection. His work employed an FRPD with misimplementation and misperception noise, and used GAs and ANNs to operationalise the distinction between hardwired and softwired strategies. Players remembered only the last two interactions with the current partner, and were paired at random. Use of the GA resulted in a highly cooperative equilibrium with brief spikes of defection. From either an initially random population or one composed of ALL-D players, WSLC came to predominate. Cooperation was increasingly stable (but at lower levels) with increasing noise. In contrast, WSLC did *not* emerge from the ANN model (where originally random weights were adjusted if the opponent defected). Instead, the ANNs became more likely to cooperate after they had themselves cooperated last time. Noise reduced the level of cooperation. Giving the ANNs the ability to learn from a more successful partner by imitation or instruction-like processes substantially increased the rate of cooperation, instruction by more than imitation.

Posch (Posch, 1997; Posch, 1999) used a GA to investigate the evolution of WSLC-like strategies for both IPD and other repeated two-player, two-strategy games, with 0%, 1% or 4% noise. The strategies were defined by either 3 or 4 parameters: whether the first move was C or D, the player's *aspiration level* (the payoff below which a switch of move would be considered), the probability of such a switch occurring if considered, and in some cases, the speed of change of the aspiration parameter. With fixed aspiration levels, cooperation dominated except at 4% noise, but the predominant strategies were usually close to GRIM: starting with C and having an aspiration level between the P and R payoffs. Allowing aspiration levels of agents to change during a game led to more cooperation (hence higher average payoffs).

Hoffmann and Waring (1998) used a GA to evolve Moore machines to play the FRPD. Without noise or complexity cost, the population switched repeatedly between ALL-D and mixtures of GRIM and TFT; adding either complexity costs or misperception noise left only ALL-D stable; misimplementation noise left no strategy entirely stable, but ALL-D did best.

The effect of noise on the dynamics of populations playing the RPD is hard to summarise, although WSLC-like strategies appear to be quite robust, particularly if modified toward the gradual learning typical of animals under a wide range of circumstances. The most illuminating recent paper in this subfield is analytical (Leimar, 1997). This uses a 'state-space' approach: this is equivalent to Moore machine terminology except that outputs and transitions between states can be probabilistic. Leimar also uses the concept of a 'limit ESS' (Selten, 1983): a strategy b is a limit ESS for a game with the set of strategies B, if there is an ESS $b(\eta^k)$ for each of an infinite series of games with misimplementation noise, $B(\eta^k)$, in which noise approaches 0 as k increases. Leimar's main results concern 'reflexive' strategies. During an encounter between two copies of such a strategy (whatever errors occur), any state x has a partner x^R, such that when one copy is in state x, the other is always in x^R. Non-reflexive strategies can be limit ESSs only for marginal parameter combinations (CTFT, WSLC and GRIM are all reflexive, but TFT is not). Of strategies which always play C on the first move and if the last encounter was CC, WSLC is a limit ESS if $w \geq (T-R)/(R-P)$, GRIM if $w > (T-R)/(T-P)$ — the greatest possible range, and CTFT when $[w > (T-R)/(R-S)] \wedge [w > (P-S)/(R-S)]$. TFT and GTFT are not limit ESSs, and Leimar has found no mixed strategy limit ESS. Future ABSS work on noisy games should take account of these and other results in Leimar (1997).

All the studies described in this section explore whether reciprocal altruism among completely selfish, computationally or cognitively limited agents could establish and maintain cooperation in the RPD, when each individual in the population is equally likely to encounter any other. There appear to be circumstances in which it could do so, and others in which it could not. As in the work considered in section 4, simulation and more specifically ABSS have played useful roles, but mathematical analysis has produced the most significant results.

6. The Prisoner's Dilemma in Structured Populations

This section examines the effects of removing the unrealistic assumption that all members of a population are equally likely to interact with each other. Most commonly, this is done by imposing some form of spatial structure on the population's interactions. The main alternatives to this are to give members of the population a choice as to whether they take part in a PD-type game with a randomly selected partner, or to go further and permit them actively to select their partners. This may be done on the basis of previous direct interactions with the potential

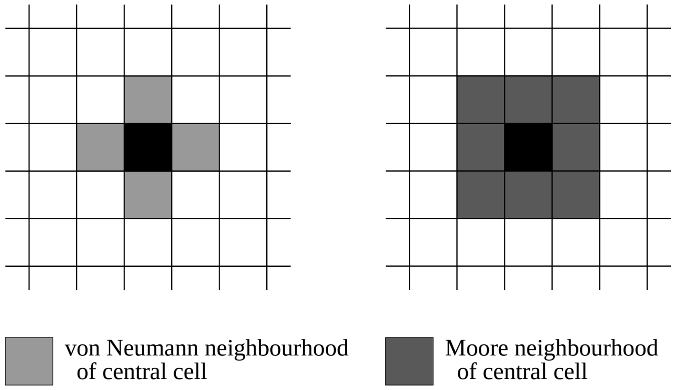

Figure 3. Cellular automata: von Neumann and Moore neighbourhoods

partners, or may rely on some form of 'reputation' based on interactions with third parties (see 7.1).

Adding spatial or social structure to the game-playing population generally makes the resulting system much less amenable to mathematical analysis. Although some analytical results are reported below, most of the work described relies on simulation, and most of this is, in a broad sense, agent-based rather than numerical. The 'agents' concerned are often very simple, particularly in studies of spatially structured populations playing the one-shot PD. Some of the studies described concern interactions which do not, strictly speaking, conform to the standard definitions of the PD and RPD; but all to some degree embody the conflict between immediate individual interest and the benefits of longer-term cooperation central to social dilemmas.

The investigation of spatially-embedded game-playing populations makes extensive use of cellular automata (CA) (von Neumann, 1966; Langton, 1984; Hegselmann, 1996): regular networks of identical finite automata (deterministic or stochastic), each linked to a small set of neighbours, and all, in general, updated synchronously.

6.1. Spatially Embedded One-Shot Prisoner's Dilemma

Adding spatial structure allows cooperation to maintain itself without giving agents strategies in which their moves are conditional on previous plays, as in the studies of the RPD discussed above.

Nowak and May (1992) and Nowak and May (1993) studied square arrays of cells, each occupied by a simple agent, set to cooperate or defect in one-shot PDs with its eight orthogonal and diagonal neighbours (this is a 'Moore neighbourhood'; the four orthogonal neighbours form a 'von Neumann neighbourhood') (see figure 3). After a round of simultaneous one-shot PD games over the entire array, the occupant of each cell copies its most successful neighbour (or if one interprets the system as modelling biological evolution rather than social learning, is replaced by a copy of that neighbour). The resulting system is a finite, deterministic CA. This work made use of non-standard payoffs, R being set at 1, both P and S at 0, and T at some value greater than 1. Hence the game played is not strictly a PD, since if a player knew their partner was going to play D, they would have no immediate reason to prefer either of their possible moves to the other. The authors say that setting P slightly above 0 made no qualitative difference to the findings (and neither did various changes to the neighbourhood used). Results depended on the value of T: with $T < 1.8$, 2-by-2 clusters of defectors (Ds) shrink, while with $T > 1.8$, they grow; with $T < 2$, 2-by-2 clusters of cooperators (Cs) grow, with $T > 2$, they shrink. The outcome, if $1.8 < T < 2$, is a pattern of 'fractal' appearance, with an asymptotic fraction of Cs of approximately 0.318. Largely static patterns in which networks of Ds form on a background of Cs form if $1.75 < T < 1.8$. These patterns retain any rotational and reflection symmetries present in the initial distribution of Cs and Ds.

In a similar study Lindgren and Nordahl (1994), using toroidal lattices with a von Neumann neighbourhood, report that a mix of cooperators and defectors gave five different dynamics depending on error and mutation parameters: homogeneous D, small islands of C, a stable percolation network of Ds on a C background, chaos (rapid and apparently unstructured fluctuations), and Cs forming organised linear structures.

The significance of the results of Nowak and May was challenged by Huberman and Glance (1993), and later by Mukherji et al. (1996), on the grounds that they depended on the synchronous updating of all cells, which these authors criticised as biologically unrealistic.

Huberman and Glance (1993) found that with one player (or cell) at a time updating their chosen move, and players receiving a score per unit time, adjusted continuously as the matrix varied, rapid universal defection resulted in conditions where Nowak and May found fractal patterns, although when the asynchronous updating was based on *delayed* scores, greater delays led to higher asymptotic levels of cooperation. In response, Nowak et al. (1994), while claiming that many biological systems have natural (e.g. annual) rhythms which make synchronous updating a reasonable approximation, found both synchronous and asynchronous updating schemes, in which cooperation appeared to persist indefinitely. Some used stochastic rather than discrete updating rules. Both discrete and continuous score adjustment allow Cs and Ds to coexist for some combinations of T values and stochasticity level, but the continuous case does not show the chaotic fluctuations of the discrete model. Games on random grids (neighbours being those within a certain radius) produce coexistence in rather static patterns (Nowak et al., 1994) [p.55]:

> irregularity and/or stochasticity simplifies dynamics, but does not change the basic observation that cooperators and defectors can coexist due to spatial effects.

Mukherji et al. (1996) report three modifications to the work reported in Nowak and May (1993), all leading to elimination of cooperation. Nowak et al. (1996) respond that in all three cases, cooperators can persist with T values below 1.8.

Both sides have valid points in this dispute: asynchronous updating does not abolish the effect of spatial structure, but does modify it considerably. Those using synchronous updating in models of cooperation, or other forms of social or ecological interaction, need to justify its use. A related point is made in a paper (Herz, 1994) covering a range of two-person games including the one-shot PD: the precise spatial structure used, even among two-dimensional lattices, can have important effects. The agents used switch their strategy if their payoff falls below a fixed threshold, rather than comparing with their neighbours. Spatially complex patterns emerge in the PD: neither C nor D is stable in a low-cooperation neighbourhood, both are stable in a high-cooperation neighbourhood, and either D or neither in intermediate neighbourhoods, depending on the threshold. For some other games, particularly where it is advantageous to adopt the opposite strategy to one's partner, the form of lattice has important effects: a hexagonal lattice or a square lattice with Moore neighbourhood (unlike the latter with von Neumann neighbourhood) produces 'frustration' phenomena — all pairs of neighbours cannot simultaneously adopt one of the preferred strategy combinations. Clearly, real organisms are unlikely to be arranged exactly on any form of regular lattice; researchers must beware of artefacts arising from such regularity.

Eshel et al. (1998) and Eshel et al. (2000) describe work combining analysis and agent-based simulation, similar in several ways to that of Nowak and colleagues. However, they worked primarily on the one-dimensional case, agents being arranged in a ring, which is more amenable than the two-dimensional to analytical attack. Eshel et al. (1998) analyses the interactions of 'Altruists', who provide benefits to their immediate neighbours to left and right (at a cost to themselves of less than half that benefit) and 'Egoists', who provide no such benefits. A member of either type will switch to the other if, on comparing the average payoff to Altruists and Egoists within its immediate neighbourhood (which includes itself and its left and right neighbours), it finds that the other type does better. Cost/benefit interactions are presumed synchronous around the ring, but neighbourhood comparison and consequent switching may be either synchronous or asynchronous.

A uniform ring of either Altruists or Egoists is stable. A large group of contiguous Egoists shrinks until only one or two remain, because those at the edge of the group see their immediate Altruist neighbour scoring more than the average of their own score and that of their Egoist neighbour. A contigous pair of Egoists surrounded by Altruists is stable, while a single Egoist can cause its two immediate neighbours to switch, but further comparison will then lead either neighbour to switch back. If all ring members compare performance with their neighbours synchronously, this will lead to indefinite switching between one isolated Egoist and three in a row; otherwise, the eventual result will be a contiguous pair. Groups of at least five Altruists always persist indefinitely, and groups of three or four may do so, depending on what flanks them. If the original configuration allows any Altruists to persist indefinitely, the ring will eventually contain at least 60% Altruists. In this case, moreover, the rarer Altruists are originally, the more there will eventually be. This is because a large group of contiguous Egoists shrinks until only a few remain, and the fewer Altruists there are initially, the fewer *groups* of Egoists there will be, since each must be separated from the next such group by some Altruists.

In Eshel et al. (2000) it is noted that the two-dimensional case resists analysis; so does the one-dimensional case if stochastic learning is combined with a wider neighbourhood than those immediately to either side. Simulations are used to explore conditions under which Altruists or Egoists resist invasion by the other type. A kind of 'cultural' conformity (in which changes of type will not occur if both immediate neighbours are of the same type as the agent itself) promotes cooperation, as does the ability to learn from a wider neighbourhood than that in which cost/benefit interactions occur.

Mitteldorf and Wilson (2000), investigating plant growth, look at the dynamics of altruism on a two-dimensional lattice with empty sites permitted. Simulations without empty sites showed that limited dispersal impeded the spread of altruists, as well as concentrating them in patches. With vacant sites, 'strong altruism' (altruism with an absolute rather than merely relative cost in reproductive terms) was able to spread. Periodic 'culling' (either uniform or clustered), and random failure of 'seeding' in a cell, appeared to support strong altruism in some cases, by allowing greater population densities in patches of altruists.

Another study involving empty sites, but with asynchronous updating, is Epstein (1997). Agents, which can accumulate 'wealth', move to a random site (within a maximum distance) on a lattice, and play a fixed C or D strategy against a random neighbour. If sufficient payoffs accumulate, an agent clones itself onto a neighbouring site; if accumulated payoffs become negative (the P and S payoffs are negative), the agent dies. Despite the agents' inability to recognise each other as individuals or types, spatial zones of cooperation emerge. Some parameter settings do lead to universal defection — for example, making all payoffs positive, but even here setting a maximum age for agents allows cooperation to persist. Epstein conjectures that cooperation acquires selective advantage as *external* selection pressures increase.

Killingback et al. (1999) look at the interactions between agents with a *range* of possible levels of altruism toward neighbours. Each agent occupies a lattice site, 'invests' at a specific level in each neighbour, and adopts the strategy of its most prosperous neighbour (including itself). Small mutations in strategy occur. On a two-dimensional lattice, clusters of agents with slightly higher levels of altruism can form by chance, and then expand and maintain themselves, leading to a gradual rise in investment levels. Stochastic updating (with the most successful neighbour having an 80% instead of 100% probability of being copied), asynchronous updating, random deletion of some cells in the lattice, and 10% of mutations creating very selfish individuals all speeded up the increase in investment but led to a slightly lower asymptotic value.

Finally, Sella and Lachmann (2000) use both analysis and simulation to study subpopulations of cooperators and defectors inhabiting sites on an infinite two-dimensional lattice: random PD interactions within a site determine fitness. The payoffs allow cooperators but not defectors to maintain homogeneous populations. Mutation and migration occur with low probability. The system can reach a dynamic steady state if 'birth' and 'death' rates of populations are equal — cooperation persists although every population of cooperators dies out.

The studies reviewed in this subsection indicate that embedding agents in a spatial context, so that they are certain or likely to interact with the same partners repeatedly, can substitute for the exogenously imposed repetition of the RPD. The emergence of cooperation is not dependent on synchronous updating, but either synchronous updating or spatial regularity may produce artefactual effects.

6.2. Spatially Embedded Repeated Prisoner's Dilemma

As with many other aspects of investigation into the PD, Axelrod (1984) [pp.158-168] pioneered the study of the spatially embedded RPD, and some of the details of his approach have been followed by other investigators. Axelrod placed copies of the entries to his second tournament into a two-dimensional lattice of cells, with edges joined to form a toroidal structure. The agent in each cell played an IPD game with each of its horizontal and vertical neighbours, then adopted the strategy of one of its most successful neighbours if (and only if) any achieved a higher average score than its own. This could be interpreted either as learning by imitation, or as competition to reproduce. In this setup, Axelrod's 'collective stability' implied a form of 'territorial stability': a 'collectively stable' strategy could not be overrun by the spread of a single player of a different strategy. However, while TFT still did quite well, other strategies which had not done well in his original tournament spread more successfully.

There is a small batch of studies using numerical simulation, and concentrating on the interactions between ALL-D and TFT strategies, which all use what can be called 'temporally distributed RPD': players keep track of one-shot PD encounters with different partners over time, and recognise them if re-encountered. Dugatkin and Wilson (1991) considers a 'metapopulation' of agents, divided into 'patches'. If the number of interactions is sufficiently large relative to patch size, ALL-D cannot invade TFT, but 'ROVERs' — ALL-D players able to move — can if travel cost is low enough. A follow-up study is considered in section 7.1. Other studies of similar type include Hutson and Vickers (1995) and Ferriere and Michod (1996), which apply analysis and numerical simulation to the interactions of TFT and ALL-D in a one-dimensional lattice of cells. The models differ in detail, leading to different conclusions. For example, Hutson and Vickers (1995) find that increased mobility always helps ALL-Ds, while Ferriere and Michod (1996) find that a TFT population can only be invaded by ALL-Ds with either very high or very low mobility.

Once a population is given a spatial structure, the distinction between the one-shot PD and the RPD is not clear-cut. Liebrand and Messick consider cooperation where decision-making is based on very simple heuristics (Messick and Liebrand, 1994; Messick and Liebrand, 1995; Liebrand and Messick, 1996b; Liebrand and Messick, 1996a). Their agents occupy cells of a Moore-neighbourhood toroidal lattice. The agent in a randomly selected cell, and a randomly selected neighbour, play a one-shot PD game, with the first cell (only) then updating its choice of move for next time it is selected. Three strategies were studied: TFT (but note that the partner next time need not be the same neighbour), a form of WSLC (a win being defined as a payoff at least equal to the mean of the last payoffs of itself and its neighbours), and a 'Win Cooperate Lose Defect' (WCLD) rule — equivalent to TFT in a two-player game — based on the idea that cooperation is more likely if the agent's self-esteem is high. For WSLC the only stable global state is uniform defection — for WCLD, uniform cooperation. Time to reach a more or less stable proportion of players intending to play C on their next move appeared linear in the number of players for WCLD, more than linear for the others, with levels of cooperation about 40% for WSLC, about 55% for WCLD, in which clustering of cooperators occurred. The authors conclude that cooperation is sustainable in large groups without universal mutual monitoring, when competitive social comparison is used to evaluate outcomes — a rather large claim on a limited basis.

Lindgren and Nordahl (1994), using toroidal lattices with a von Neumann neighbourhood, report that agents having a memory of one move and playing the IPD against their neighbours, with successful strategies spreading to neighbouring sites, produced complicated results as error rates and payoffs were altered. These included spiral waves, 'gliders' (small coherent moving structures), irregular wave fronts and patches. When gene duplication was added, allowing agents with longer memories to evolve, even more complicated histories were frequent. Asynchronous updating made little difference except that spiral waves vanished. Using the same environment Lindgren (1997) compares non-spatial models and spatial CA models across a range of payoff parameters, again allowing the evolution of more complex strategies in the course of a run. In the FRPD, cooperation persisted in the CA model (only) across a wide range of parameters. For the IPD, using finite memory strategies as in Lindgren and Nordahl (1994), misimplementation noise and mutations, the non-spatial system tended to pass through several metastable states dominated by agents with memories of increasing length, with generally cooperative but retaliatory types eventually predominating. In the CA model, either islands of ALL-C in a sea of ALL-D agents, or spiral waves of ALL-D, ALL-C and TFT occurred (depending on the payoffs used) before memory-2 strategies (those able to condition their move on the preceding two moves by each partner) appeared. When all strategies representable by finite automata were permitted (a much larger search space than finite memory strategies), ALL-D took over the non-spatial model more often than for finite memories, although cooperative strategies with error-correcting mechanisms sometimes did well. In the CA model, CTFT or strategies resembling it often won, and a single strategy generally dominated — which was not the case when only finite memory strategies are used.

Grim (1996) and Brauchli et al. (1999) both studied stochastic strategies with a memory only of the previous move, similar to those of Nowak and colleagues, on toroidal lattices with Moore neighbourhoods. Grim (1996) describes his work as a spatialisation of the stochastic model in Nowak and Sigmund (1992). He used a range of 121 strategies (probabilities of C following a C or a D from the partner independently taking any of the values $.01, .1, .2, .3, .4, .5, .6, .7, .8, .9, .99$). After each round of interactions, each cell adopted the strategy of the most successful of its neighbours (including itself). The strategy $< 0.99, 0.1 >$ (C following C with 0.99 probability, C following D with 0.1 probability) usually did best. Grim notes that in Nowak and Sigmund (1992), the more generous strategies that tended to win often fell to very low proportions of the population (in the case of $< 0.99, 0.3 >$, less than 1 in 10^{12}), before reviving. $< 1 - \epsilon, 2/3 >$, where ϵ is small but non-zero, seemed to have the highest score against itself among strategies invulnerable to invasion by small clusters. Brauchli et al. (1999), using strategies that could depend on the player's own last move as well as the partners and always holding $w < 1$, compared an approach similar to Grim's (but including random mutation), with a non-spatial alternative. In the latter, strategies likely to play C only when both players had just done so predominated. This contrasts with the findings of Nowak and Sigmund (1993), but the latter used many more generations, and set $w = 1$. Brauchli et al. (1999) found more cooperation in their spatial model; generous versions of WSLC were the most successful strategies.

The effects of varying the sizes of interaction and learning neighbourhoods in a one-dimensional setting found by Eshel and colleagues, described in section 6.1, contrast with those reported by Hoffmann and Waring (1996) for the IPD with $w = 1$. Players were limited to interaction with the nearest i neighbours on each side, and had learning neighbourhoods consisting of the nearest l on each side. Strategies were deterministic, the strategy space consisting either of ALL-C and ALL-D, or of these plus all 2-state Moore machines, encoded as bitstrings suitable for a GA. Each generation involved calculating the payoffs of a cell with each of its interaction neighbourhood, then picking two members of the learning neighbourhood (stochastically, but in a manner weighted to the more successful), and using the GA's crossover and mutation operators to produce a new agent. With ALL-C and ALL-D the only possibilities, there was convergence on one or other pure state, with greater localisation of both interaction and learning favouring ALL-C. With two-state machines permitted there was more cooperation (due to GRIM and to a lesser extent TFT), periodic shifts between high- and low-cooperation attractors, and again most cooperation with narrow interaction and learning neighbourhoods (i.e. low i and low l): low l increased cooperation for all values of i — localisation of learning aids clustering which helps cooperators — but for high l, low i seemed to inhibit cooperation. Low i and high l gave least cooperation. Eshel et al. (2000), in contrast, found that a learning neighbourhood wider than the interaction neighbourhood increased cooperation. Since Hoffmann and Waring's ALL-D versus ALL-C case is a very similar game to that used by Eshel, further investigation of the differences between the two models would be useful: possibly the different ways in which new agent strategies are chosen is the crucial factor.

The remaining studies described in this subsection further emphasise the importance of interactions between the forms of interaction and of learning or replacement used. Cohen et al. (1999) followed populations of agents through 2,500 learning periods (or generations, depending on whether the model is given a social or a biological interpretation), each involving a number of IPD games for each agent. They used two different strategy spaces (both involving only memory-1 strategies, but in one case deterministic, in the other stochastic), six processes for selecting partners, and three adaptive processes. The six selection processes were: use of a fixed two-dimensional lattice with von Neumann neighbourhood (2DK), or of a new lattice in each period (2DS), of a fixed random network in which each node had four neighbours (FRNE), of a fixed random network with nodes having different numbers of neighbours (FRN), of a system of arbitrary labels with a preference for playing similarly labelled agents (Tag), and of a random choice of partners (RWR). The three adaptive processes were imitation of the highest-scoring agent encountered, without error (Imit), imitation (with errors) of the highest-scoring agent encountered (BMGA), and imitation (with errors) of a random agent, if it had done better than the imitating agent (1FGA). The results confirmed the effectiveness of a two-dimensional space embedding in promoting cooperation, but FRNE and to lesser extent FRN also promoted it markedly. The selection procedures 2DS and RWR led to least cooperation (2DS was included to investigate whether the cooperation-enhancing effects of spatial embedding are due to repeated interaction or to neighbouring pairs sharing additional neighbours), but even here it could emerge with deterministic strategies and the 1FGA adaptive process. Tag usually led to intermediate levels of cooperation, but when combined with deterministic strategies and the Imit adaptive process was worst of all — leading to pure defection. The most robust adaptive process was 1FGA, while Imit, which led to complete loss of diversity, was mostly bad, but led to high levels of cooperation in a few cases. In summary, any form of 'context preservation' promoted cooperation, but there were complex interactions between strategy space, partner selection process, and adaptive process.

Kirchkamp (1996) investigated imitation of neighbours using a Moore-neighbourhood lattice, a strategy space including either just ALL-C and ALL-D, or these along with the two-state Moore machine strategies, and in the latter case an approach in which an agent played the same strategy against all opponents, but played an independent game with each neighbour. Both interaction and learning (copying successful neighbours) could be synchronous or asynchronous, although learning was always performed less frequently than interaction. Asynchronous learning eliminated cooperation — as in Huberman and Glance (1993) — but only when just the most recent payoffs were considered in assessing success. When 2-state strategies were allowed, more cooperation occurred, and strategies with different average payoffs survived indefinitely. 'Odd' strategies (such as 'Blinker, which alternates C and D without regard to outcomes) survived unless interaction was stochastic (in contrast to Huberman and Glance (1993)) as well as learning.

In further work (Kirchkamp, 1999; Kirchkamp, 2000), Kirchkamp took the unusual step of making the learning rules subject to adaptive processes in turn. In these experiments a new stage game was selected periodically (from a range including the PD) to force repeated adaptation. Repeated game strategies were represented by Moore automata with a maximum of 1, 2, 3 or 4 states. Learning involved comparison with a random neighbour. Agents' learning rules were characterised by parameters representing sensitivity to their own and the neighbour's payoffs, and general readiness to change. Learning rules were updated less often than strategies, but more carefully: agents looked at all their neighbours' rules, taking their relative success into account. The range of stage game parameters where cooperation dominated, shrank relative to simulations with fixed learning rules.

As might be expected from the effects of repetition and of spatial embedding in promoting cooperation in the PD, combining the two produces a high degree of cooperation in many circumstances. However, the stability of cooperation and of specific strategies, and other features of strategic dynamics such as the degree of diversity across space and time, appear to depend in complex ways on the range of permitted strategies, the spatial setup, and the precise procedures used for allowing agents to interact, new strategies to arise, and successful strategies to spread. This is not surprising, but makes it difficult to assess the significance of any particular result. However, since social interactions within populations of real human and animal agents do differ in all these ways and more, the diversity of dynamical phenomena found in the work reviewed may well be exceeded in real life.

6.3. Partner Choice

This subsection reviews studies giving agents some opportunity to select their partners in PD-type games. Usually this involves recognising previously encountered individuals, but there are other possibilities: individuals may bear 'tags' identifying them as belonging to particular groups, or agents may be spatially embedded, and simply given the opportunity to move in search of better payoffs.

Much of the work reported by Macy (1991) deals with a stochastic learning model of pairwise interactions between players who do not reason strategically, but respond to immediate outcomes. The work is summarised as follows:

> Simulations show how effects of network size, density, mobility and anonymity derive from a fundamental principle of collective action, that is, the need to reduce the number of choices that must be fortuitously coordinated in order to escape noncooperative equilibrium... Synchrony is the key to breaking out of social traps.

Macy hypotheses that if subjects cannot differentiate among interactants, the probability of lock-in to cooperation decreases exponentially with the number of agents. His simulations suggest that this probability increases when subjects choose interactants by mutual consent, and increases *further* when players choose a partner unilaterally, those with the greatest accumulated winnings getting first pick, as this speeds up the emergence of a stable set of pairings.

Ashlock et al. (1996) studied the evolution by GA of populations playing the FRPD, comparing random choice of partners with a model allowing choice and refusal of partners. In the latter, players in each generation keep running averages of payoffs in interactions with each other player. Players make offers to play, if any of the other agents score above their minimum tolerance level. Each player then accepts all tolerable offers made by others. The overall emergence of cooperation was faster than in random tournaments, and the initial dip in cooperation found in such tournaments was missing or shallower. Random choice populations tended to become homogeneous; choice and refusal populations stayed genetically diverse, but with all players having similar fitness. Cooperative ecologies emerged for almost every combination of parameters tried. The 'social network' developing depended on the minimum tolerance level. The apparently trivial choice between using a bubble-sort or a randomised-sort algorithm to select parents within the GA (the former favours incumbents over newcomers) altered the population's behavioural diversity, raising the question of how far the effects found depend on particular features of the implementation.

Crowley et al. (1996) studied the effects and evolution of the ability to recognise previously encountered individuals. Deterministic strategies with a limited memory capacity (between 1 and 5 moves) were evolved using a GA. Ability to recognise others was either absent, complete, or allowed to evolve (with an attached cost). A fixed, substantial amount of individual recognition maintained high levels of cooperation even at low pairing continuities, and a significant but limited recognition capacity evolved under selection. Without individual recognition, substantial cooperation arose only with the highest pairing continuity. Imposing a fitness cost on memory reduced both memory and cooperation.

Sherratt and Roberts (1998) studied the evolution by GA of 'generosity' (cost of cooperation) and 'choosiness'. In every generation, agents had repeated random encounters, with recognition occurring when individuals re-encountered each other. The factor used to multiply the cost to the donor to get the benefit to the recipient, k, was fixed for the duration of the run. With $k < 1$, as one would expect, no cooperation arose, while with $k > 1$ cooperation evolved, increasing with the expected number of re-encounters between individuals; the greatest generosity being found with $k \approx 1.2$. Choosiness generally rose, then declined as cooperation became general; abolishing choosiness destroyed cooperation.

Partner selection may be implemented using a spatial structure, as in Majeski et al. (1997) who found that levels of cooperation rose if agents were given a probability of moving to a new location when they got a non-positive total payoff in PD interactions with their neighbours. The lattice began with randomly distributed agents, and numbers were restricted by imposing a density-dependent cost of survival. Strategies were probabilistic, with memory of both player's moves in the preceding round. Agents had a limited lifespan, and reproduced, asexually but with mutation, when their accumulated payoffs were sufficient. Movement could be local (to an adjacent cell), global (to a random empty cell), or absent. In all three conditions, high levels of cooperation were reached, but were more stable in the movement conditions; they were achieved faster with local than with global movement.

Cooper and Wallace (2000) also gave partner selection a spatial context. The PD and other two-player, two-strategy games were played by $2N$ players placed in a double ring formation. Agents had four possible strategies (two stage game strategies, combined with staying put or moving to a random location), and after each play either learned by adopting the strategy that did best on average within an 'observation window' covering the pairs closest to them around the circle or, with probability ϵ, chose a strategy at random. If all players stayed put and ϵ was 0, members of any pair would always make the same choice, effectively becoming a single player. If the stage game was the PD, then partnerships and cooperation evolved even from a population of moving defectors, but the situation remained noisy.

Two studies aimed explicitly at exploring group selection are described in Hales (1998) and Hales (2001). In Hales (1998), which is aimed primarily at investigating group formation and stereotyping, agents in a one-dimensional world interact culturally by exchanging 'memes' encoding behavioural rules and arbitrary group markers, and economically though the PD. They apply stochastic repeated game strategies to all members of particular 'culturally tagged' groups rather than individuals. This produces distinct groupings with shared 'memes'. Hales (2001) switches the focus to genetic rather than cultural transmission, and to the one-shot PD. Agents occupy an abstract, multi-dimensional 'tag space', each having a string of bits constituting its 'tag', plus a one-bit PD strategy. Agents preferentially play those with an identical tag; the most successful agents reproduce asexually, but with mutation which can affect both tag and strategy. Both from a random strating point, and from initial universal defection, cooperative populations arise. Cooperative groups with the same tag can be invaded by mutant defectors; but such groups then decline relative to those lucky enough not to have been so invaded.

Zeggelink et al. (2000) used their 'Social Evolution Model' (SEM) to investigate whether reciprocal altruism can explain the formation of stable social groups. SEM agents (representing families) have a given probability of getting into distress in any period, and can ask help from others. Agents give help if the requesting agent has not received it more times than they have given it to the askee; those who have given more often are preferred if more than one qualifies. A group of any size can go extinct if all are in distress, and thus cannot help each other. Simulations were run with varying numbers of agents, and probabilities of getting into distress. To the authors' surprise, little formation of subgroups occurred, although some segmentation was evident when the probability of getting into distress was low, and hence the number of relationships formed was small.

To a considerable extent, partner selection reduces the 'dilemma' aspect of the PD, providing a third option in addition to cooperation or defection: non-interaction. Frequently such an option is indeed available to human or other animal agents. To judge by the studies reported here, the effects are comparable to those of imposing a spatial structure on the population (as already noted in relation to Cohen et al. (1999) above). Given the contrasting results of Hales and of Zeggelink et al., further work on the ability of partner selection to promote subgroup formation would be interesting.

6.4. The Preservation of Context

As noted in Cohen et al. (1999), any form of 'context preservation' promotes cooperation, and this can be seen as the main theme of section 6. The RPD itself is more conducive to cooperation than the one-shot

PD (and the IPD more so than the FRPD) for this reason, but the imposition of spatial structure, or the ability to recognise individuals or groups, shifts the balance further in the cooperative direction.

Most of the studies covered in the section employ ABSS; analytical techniques for spatial game theory are far less developed than in the non-spatial case (but see Dieckmann et al. (2000) [Part D] for current related developments). The studies differ in the stage game played, the form of context preservation studied, the range of strategies examined, the synchronous or asynchronous timing of interaction and strategy change, and the strategy change mechanisms used. Most primarily address the question of the circumstances under which reciprocal altruism can flourish by benefiting the agents practicing it. The exceptions are the recent studies which directly address the issue of group selection: (Dugatkin and Wilson, 1991; Hales, 1998; Hales, 2001; Mitteldorf and Wilson, 2000; Sella and Lachmann, 2000). All studies of structured populations, however, have some relevance to the group selection controversy: the contrast between local and global scales raises the possibility that strategies which are locally successful when most of their neighbours are following a different strategy, undermine their own success when this is no longer so, and thus fail globally. Conversely, strategies which have a hard time when mostly encountering agents playing other strategies may be able to spread once a core playing them has formed. Very similar contrasts of scale are central to the arguments of Wilson and Sober (1994) and Sober and Wilson (1998).

7. Multi-Agent Social Dilemmas

The focus now shifts to studies of multi-agent interactions. Of course, the work reviewed in the last two sections has concerned populations or societies of many agents, but the direct, payoff-generating interactions between agents have been pairwise; here, the strategies of larger groups of agents interact directly, as well as via mechanisms for strategy selection and transmission.

7.1. Indirect Reciprocity and Reputation

The studies in this subsection are intermediate between the two-agent and multi-agent cases: third parties learn about pairwise interactions, and may condition their own subsequent interactions with the members of the pair on what they learn — in particular, they may reward cooperative or altruistic behaviour and punish the reverse.

Boyd and Richerson (1989) studied a somewhat artificial form of indirect reciprocity, using analytical and numerical simulation methods. They assumed groups of size n to be sampled from a large population, with members of each group effectively arranged in a 'ring', each having an 'upstream' neighbour who could affect their payoff, and a 'downstream' neighbour, whose payoff they could affect. Possible strategies were ALL-D (never benefit the downstream neighbour), DTFT (do so if that agent benefited *their* downstream neighbour last time, and UTFT (do so if *you* were benefited last time). Likelihood of a group staying together for another round encouraged cooperation, as did assortative grouping (giving pairs with the same strategy an above-average chance of being in the same group), while large n strongly discouraged it. The authors conclude that 'be nice to those nice to others' (DTFT) is more helpful in maintaining cooperation than 'be nice if others are nice to you' (UTFT), and that the effect of indirect reciprocity depends strongly on the information available to agents. They note that there may be something about many interlinked reciprocation loops that this model does not capture.

Dugatkin (1992) incorporates the communication of *reputation* into his earlier study (Dugatkin and Wilson, 1991) of 'ROVER' agents: defectors moving about to exploit naive cooperators. 'Cultural transmission' of information about these 'con artists' was investigated using numerical simulation. Only if the likelihood of con artists being recognised depended on their frequency, did the low but non-zero frequency of such agents thought to be characteristic of human society result.

Yamagishi and Hayashi (1996) report a simulation study in which players could give (unilaterally) to another; all players were given the strategy of giving at their 'altruism level' to a (randomly selected) player as long as that player gave to others at a 'satisfactory' level, but different players had different definitions of 'satisfactory'. Altruism increased, and maintained a high level, only if selectivity was neither too strict nor too lax. An experimental study of six-person groups using this paradigm found that very high levels of giving developed (this is one possible Nash equilibrium of the game).

Both ABSS and analytical approaches are used in (Nowak and Sigmund, 1998a; Nowak and Sigmund, 1998b) in an attempt to specify conditions in which indirect reciprocity would maintain the stability of altruistic behaviour. The simulations involve agents deciding whether to give another agent a benefit b, which would impose a direct cost c, but would add 1 to the agent's 'image score', while refusing to help would subtract 1, within some preset limits. All agents were given strategies of helping when the potential recipient's score is at least some specific value k. In simulation runs, the value $k = 0$ became dominant. With mutation added, there was cooperation most of the time, with 'spikes' of defection due to drift towards unconditional cooperation. If only a small proportion of agents observed each interaction, cooperation could still be established, but required more interactions per generation, and was harder to establish for larger groups. Analytical techniques were applied to a simpler model, with only two image levels (depending on the last interaction) and two strategies, defectors and discriminators. In this model there is a threshold proportion of discriminators necessary for the establishment of cooperation, and a threshold number of rounds per generation necessary for its stability.

Dessalles (1999) used ABSS to explore closely related issues. The paper argues that conditions which permit reciprocal altruism to survive are quite restrictive, and proposes a variant of the handicap principle as an alternative. The model described is quite complex. Each encounter between two agents has an 'initiator', who sacrifices a certain amount to provide a (larger) benefit to the 'respondent'. The latter returns a fraction of the amount given (possibly greater than 1), again at a cost which is less than the benefit returned. In each generation, each agent has opportunities to be the initiator, and to be chosen as respondent. Generosity in each role is genetically determined, but agents have a limited ability to keep track of cooperative partners. The population is divided into groups, within which interaction and reproduction occur; there is some inter-group migration. It is argued that reciprocal altruism will be stable only when there is 'some highly profitable trade' between the parties, and they are cognitively sophisticated enough to discriminate between partners. Otherwise, cooperation levels will at best fluctuate. The author therefore turns to Zahavi's handicap principle explanation of altruism, but objects that while the advantages to the altruist of gaining prestige are clear, the advantages of *giving* prestige to an altruist are not. (This overlooks Zahavi's suggestion that altruistic behaviour is an *honest* signal of high fitness — and thus of ability to do well in antagonistic encounters. However, invocation of the handicap principle cannot in itself explain why altruism, and not some other expensive behaviour, should evolve.) Dessalles presents simulations both of a simple version of Zahavi's model, in which altruism is not stable, and a 'political' version, in which social groups divide into *coalitions*, competition between which is determined by the vitality of their leader. In this model, altruism is one possible sign of potential high performance as a leader (the model admittedly does not *predict* altruism, but is *compatible* with it).

Sen et al. (2000), in one of the few simulation studies of social dilemmas directly motivated by the potential problems of artificial multi-agent systems, consider a group of agents charged with delivery tasks, so organised that an agent can sometimes take on another agent's task at lower cost than it could do it itself (for example, if already committed to a journey that would take it near the delivery point). In the initial simulations, probabilistically reciprocating agents (with a greater probability of agreeing to help if the cost is small, and if benefits have been received from the requesting agent) outperformed selfish ones in the long run, but still wasted energy on them. 'Believing reciprocating agents', which use information from others about third parties, can cut this cost, but are vulnerable to exploitation by groups of 'collaborative lying selfish agents' who falsely vouch for each other. This problem can be avoided by 'Earned-Trust reciprocating agents', who only believe agents which have provided them with a net benefit.

The phenomena of indirect reciprocity and reputation have what can be seen as political or moral aspects, unlike the kinds of population structure explored in sections 5 and 6: agents, at least implicitly, make normative judgements about the behaviour of others even where this does not directly affect them. The empirical evidence reviewed in 2.4 strongly suggests that these phenomena are of considerable significance. Moreover, as Sen et al. (2000) indicate, this is an area where simulation is potentially of great practical interest for application-oriented MAS, with regard to the interaction between artificial agents with potentially conflicting interests; this can thus be expected to be a growth area for ABSS research over the next decade.

7.2. The N-Person RPD and Similar Games

In the N-person Prisoner's Dilemma (N-PD) as generally defined, all players have the same possible moves and payoffs, and decide simultaneously whether to play 'C' or 'D'. As in the two-person game, if the N-PD is played only once, a player will always get more for playing D than for playing C, but if all players play D, they will all get less than if all had played C. Repeated versions of the game can be defined analogous to those for the two-player game.

Boyd and Richerson (1988) used analysis and numerical simulation to investigate the N-IPD with payoffs $Bi/n - c$ for a play of C and Bi/n for a play of D, where n is the group size, i the number of players playing C, and B and c constants such that $B > c > B/n$. They looked at the possible strategies 'always defect' (U) and 'cooperate initially, then only if at least a players did so on the previous round' (T_a), using replicator-like dynamics (representation in the next generation being a monotonically increasing function of fitness). With random group formation in each generation, U can resist invasion by any T_a. If n is the group size, T_{n-1} can resist invasion by U (i.e., eliminate it) only if w is high enough (no other T_a can do so). The necessary value of w increases rapidly with n. When groups are formed assortatively (as kin selection and group selection explanations of resource altruism require), reciprocity will help to spread T_a if and only if $B(a+1) > c$. The authors admit results could change if payoffs did not increase linearly with the number playing C, or if assortativity were modelled differently, and that their model omits sanctions for defection and the internal structure of groups. However, the work does suggest that reciprocity and kin or group selection could interact synergistically to favour cooperation.

Bankes (1994) modelled players of the N-IPD as pools of 'memomes' each coding for a deterministic strategy with a memory of the preceding move. Behaviour on any given move was governed by a randomly selected memome, and hence had a stochastic element. A GA operated on the memomes *in each pool* — so memomes are definitely not memes. If the minimum size of cooperative coalition ensuring a non-negative payoff is K, quasi-stable cooperation resulted when $K \leq 3$. The number of agents in the game made much less difference. The most common strategy was a generalisation of TFT, cooperating if and only if at least a certain number did so last time (this number was often different for different agents).

Hauert and Schuster (1997) used numerical simulation to study the IPD and N-IPD with varying amounts of memory, and various T payoffs. Starting with a homogeneous population with random responses, mutant strategies were introduced periodically. Both increasing the number of agents in the N-IPD game, and increasing the T payoff, made establishing cooperation more difficult, but cooperative solutions exist with at least 5 agents (the maximum tried).

Watts (1999) [Ch.8] embedded the N-IPD in graphs ranging from random graphs to one-dimensional lattices. The emergence of cooperation when all agents use a generalised version of TFT like that of Boyd and Richerson (1988) (but modified so initial moves are random), is highly sensitive to the degree of clustering: how likely neighbouring agents are to have further neighbours in common. More clustering makes cooperation easier to maintain, but slower to spread. If all agents play an N-person WSLC (playing the same move against all neighbours simultaneously, and comparing scores against those of all neighbours to decide whether to switch move), the main effect of increased clustering is to slow the approach to the asymptotic cooperation level.

Joshi et al. (1998) suggest that the widespread use of technical trading rules in stock markets can be explained as the result of an N-PD: all the traders would be better off if all followed 'fundamental rules', based on the assumption that prices stay close to the true worth of stocks, but the existence of price trends, whatever their cause, makes it worthwhile for any individual agent to use any technical rules that can detect them. Their ABSS work with the 'Santa Fe artificial stock market' supported this claim.

Turning from the N-PD in the strict sense, Pedone and Parisi (1997) present one of the few ABSS studies to address kin selection directly. They argue that altruistic behaviours can emerge in social groups of related and therefore behaviourally similar individuals, but not in random social groups. Their (ANN-based) agents live, alone or in groups of four, in an environment where they must divide their time between collecting immediately available food, and increasing the future food supply. If sharing the environment with others, they cannot ensure that they, or agents which have benefited them, get preferential access to food they have produced — hence this study *excludes* reciprocity, unlike Boyd and Richerson (1988). The next generation is produced using asexual reproduction with mutation. Solitary agents, and those living with close kin evolved to be 'food producers' as well as collectors, but those living in unrelated groups did not.

This subsection covers only studies which consider multi-agent social dilemmas in something close to their simplest form, uncomplicated by consideration of norms (see 7.3), or of the effects of agents' actions on the world beyond the set of interacting agents (see section 8). Their main interest lies in their relevance to kin selection and group selection.

7.3. Norms and Punishment

Sociological and philosophical controversies about norms can scarcely be touched on in this review (see Ullmann-Margalit (1977) and Conte and Gilbert (1995) for relevant treatments). The studies described in this subsection deal only with the role of norm *enforcement* in regulating behaviour — not, for example, with the process by which norms are internalised by individuals or formulated within social groups.

Axelrod (1986) reports an agent-based simulation study of the N-FRPD, and of norms and 'metanorms' (norms concerning the formulation or, in this case, enforcement of norms). Agents received a payoff of 3 for defection, and -1 if any other agent defected (defection here may be thought of, for example, as cheating in an examination). Agents detecting defection received -2 if they punished it, while the agent punished received -9. Strategies were two-dimensional: agents differed in 'boldness' (propensity to defect) and 'vengefulness' (propensity to punish). Players defected when the (randomly determined) probability of detection fell below their boldness level. After 100 rounds, with strategy mutation between rounds, there were three different outcomes among five runs: in one agents evolved towards low boldness and high vengefulness, in two, both properties evolved to low values, and in the remaining two high boldness and low vengefulness evolved. Axelrod then added a 'metanorm' level, allowing the punishment of those who failed to punish others (with metavengefulness given the same level as vengefulness for each agent). Five runs with metanorms all led to the establishment of the norm (but this required initial vengefulness to be sufficiently high). A version in which agents were divided into 'blacks' and 'whites' (to investigate the dynamics of lynch law) showed that lower costs of inflicting punishment and greater numbers could both help a group maintain its ability to coerce another, but only if metanorms exist (otherwise group members 'freeride' — do not take part in the costly process of norm enforcement). This last aspect of the study is a warning not to automatically associate 'cooperativeness' or norm-following with morally correct behaviour.

Boyd and Richerson (1992) applied analytical and numerical simulation techniques to the exploration of a model very similar conceptually to that developed by Axelrod. They argue that if the long-term benefit of cooperation to an individual, in a group of size n, exceeds the cost of coercing the other $n - 1$ into cooperation by *retributive* punishment (or *sanctioning*, as opposed to simply withholding reciprocity from the group), strategies which cooperate and sanction noncooperators, strategies which cooperate only if sanctioned, and sometimes strategies which cooperate without sanctioning, will coexist. If the cost of being sanctioned is high enough, strategies that sanction not only noncooperators but also nonsanctioners (including those who fail to sanction nonsanctioners, and so forth) can be evolutionarily stable. Such strategies can also stabilise *any* individually costly behaviour, whether or not it benefits the group. The authors speculate that this kind of punishment may be an inherently diversifying force between groups, and may also stabilise between-group variation, possibly making group selection an important force in social evolution.

Yamagishi and Takahashi (1994) use the same algorithm and parameters as in Axelrod (1986) in an ABSS study with a GA, but suggest linkage[6] between cooperation and the sanctioning of defectors as an alternative to metanorms as a solution to the 'second-order social dilemma' produced by the costliness of sanctioning defectors: actors with low 'boldness' levels both cooperated and sanctioned defectors. Resulting levels of 'boldness' were very low (so levels of cooperation and of sanctioning were high), even when the cost of punishing was high. The authors also ask where linkage between aspects of behaviour is supposed to come from: instead of imposing linkage, they tried adding a (binary) gene for linkage to those for boldness and vengefulness (which had real values between 0 and 1): if the linkage gene was 'on', the boldness value overrode the vengefulness value. With a range of initial parameters, high levels of cooperation and roughly 50% linkage evolved. Analysis showed that when linkage is distributed independently of boldness and vengefulness, agents with linkage will do better than those without if and only if the average boldness plus average vengefulness exceeds 1. If the proportion carrying the linkage gene is initially low, selection will initially push boldness close to 1 and vengefulness close to 0, making the linkage gene more or less neutral in effect. If drift carries its frequency up from its initial low value, however, the cooperative punishers this produces may transform the situation, pushing levels of boldness down and vengefulness up.

Gintis (2000b) develops a mathematical model of the evolution of a genetically-based 'strong reciprocity', in which individuals both cooperate in social dilemmas, and punish defectors despite the cost of inflicting punishment. Gintis notes that if groups are likely to disband (in the terms used here, if w is low), then cooperation among self-interested agents cannot be maintained, and that the likelihood of a group continuing to exist will fall in times of crisis — exactly when cooperation is most necessary. Gintis (Gintis, 2000b; Gintis, 2000a) argues that simple reciprocity between self-interested agents cannot account for experiments showing that individuals will punish non-cooperators even at their own expense; this behaviour makes it possible to sustain cooperation in multi-person social dilemmas (Fehr and Gächter, 2000). (For anthropological evidence of sanctions against defectors see Boehm (1993).) The model assumes a population consisting of self-interested maximisers (who will cooperate to gain reciprocal benefits in good periods, but defect in bad periods if not coerced) and strong reciprocators. If this population is subject to intermittent crises, and is divided into groups (expulsion from which is costly, and is used as punishment by the strong reciprocators), there is a minimum fraction of strong reciprocators which will induce general cooperation in a group. This minimum depends on a number of model parameters. Bowles and Gintis (2000) numerically simulate a very similar model, aimed specifically at the kind of work-sharing described anthropologically within bands of foragers. They note the possible problem of an invasion of 'cheaters', who do not shirk (defect), but fail to punish those who do. They suggest this may be overcome by sanctioning non-sanctioners (Axelrod, 1986), by genetic linkage of the cooperative and punishing traits (such linkage is likely when two traits are individually disadvantageous, but advantageous together), or by cultural transmission.

An alternative (possibly complementary) mathematical model for group-beneficial altruistic behaviour in general and for sanctioning defectors in particular, is explored by Smith et al. (2000). The model, based on the handicap principle, shows that costly behaviours can be stabilised in a population consisting of 'high' and 'low' quality individuals, by providing an honest signal of value as a social partner, provided high-quality individuals are neither too common nor too rare, and the cost of giving the signal is sufficiently greater for low-quality individuals. Sanctioning defectors, as well as other altruistic acts, can play the role of costly signal. The model has a key point in common with the agent-based simulation model of Dessalles (1999): that it is the ability to choose social partners which explains why other individuals grant the signaller high status. The authors also ask why a signal of quality should be 'prosocial' (take the form of benefits to group members), and offer three possible reasons:

- Prosocial signals of quality may also indicate other attributes (e.g. motivational ones) of potential benefit to social partners.

- Prosocial signals may have high 'broadcast efficiency', as they attract an audience.

- The use of such signals may increase group-level fitness payoffs, as explored in (Gintis, 2000b; Bowles and Gintis, 2000).

Huberman and Glance (Huberman and Glance, 1998a; Huberman and Glance, 1998b) use a combination of ABSS and mathematical analysis to investigate the dynamics of cooperation when individuals with a finite time horizon are confronted with a social dilemma. Their analysis required approximations only true in the limit of large groups, and assumptions about the form of uncertainty present. Their focus is not on sanctions for defection, but on agents' *beliefs* about the effects of their own behaviour of that on others. Their model assumes that individuals decide (asynchronously) to cooperate if and only if they believe at least a critical proportion of agents are doing so; they believe their own actions influence others, but to a decreasing extent with increasing group size. There is then a group size below which only cooperation is stable, and one above which only defection is. Between these limits, the dynamics depend on the agents' beliefs: if they think their actions will be imitated 'to the extent that others are already behaving similarly' ('bandwagon agents') there is greater stability than if they believe that at high levels of cooperation their defection will make little difference. If individuals contribute in a way that fluctuates (as the authors argue will often be so), decreasing average payoffs can result despite most individuals contributing most of the time — the dynamics being dominated by rare bursts of defection.

Norm enforcement can promote altruistic behaviour, but maintaining the level of enforcement itself presents a social dilemma. Ways of overcoming this explored in the studies reviewed concentrate on ensuring that agents will police each other; studies of the *internalisation* of norms, which leads to self-enforcement, are lacking. Processes of norm formulation and revision are also absent. Conte and Gilbert (1995) explore much of the necessary conceptual groundwork for ABSS models of these aspects of norms. Sacco (1997) sketches an analytical approach to norm change following comparison with a more successful neighbouring society, but admits the need for detailed modelling of social and political organisation which he does not provide.

7.4. Beyond Reciprocal Altruism

In this section, adaptive explanations for cooperative and altruistic behaviour other than reciprocal altruism have come to the fore. Of the alternative explanations explored, those depending on reputation, and in particular on the handicap principle, resemble reciprocal altruism in important respects. First, preferential giving toward those with an altruistic reputation, and attempts to build up such a reputation, can be seen as an intelligent agent's refinements of simple reciprocal altruism, bringing additional channels of information into play. Second, the 'altruism' explained in terms of these mechanisms, while it may be *motivational* altruism as defined in section 3 (it may be produced by genuinely unselfish intentions), is not *resource* altruism, since the behaviour produced is assumed to be, on average, beneficial to the altruist as an individual. The work on kin selection (Pedone and Parisi, 1997), that on differences between random and assortative group formation (Boyd and Richerson, 1988) in the N-IPD, and those studies stressing the group-selective aspects of norms and punishment in section 7.3 (Axelrod, 1986; Gintis, 2000b; Gintis, 2000a; Bowles and Gintis, 2000) however, do involve resource altruism.

8. The Environmental Context: Simulations of Common-Pool Resource Dilemmas

The central theme of this section is the modelling of *socio-ecosystems*, involving interaction between socio-economic systems, and a wider environment. The ABSS studies described are considerably 'thicker' than any others reviewed. The emphasis shifts from the success of individual strategies for those within a social dilemma, toward the choice of management strategies for those trying to control system-level effects.

Rossi et al. (1997), using agent-based simulations of reinforcement learning, argue that the amount of cooperation by self-interested but adaptive and boundedly rational agents in the Cournot duopoly (see below) and CPR (common-pool resource) social dilemmas may be explained by the shape of individual payoff surfaces, and how these are coupled in the structure of the game. The Cournot duopoly can be regarded as a continuous version of the RPD or N-RPD. A small number of agents take part, generally interpreted as firms which together make all of a product in a market. The more they produce in total, the lower the price, but it pays each to produce more until the cost of additional production equals the return from selling it. The Nash equilibrium is for higher production than the 'collusion point' at which profits would be maximised. The agents in Rossi et al. (1997) play a Cournot duopoly or CPR game repeatedly, adjusting their decision variable (quantity produced in the Cournot game, investment in the common resource asset in the CPR game) between rounds. The probabilities of moving the decision variable up or down are subject to reinforcement learning, so if a move in one direction is followed by increased payoff, a second move in the same direction becomes more likely. In Cournot game experiments, individuals tend to be erratic, with aggregate amount produced below the Nash equilibrium, but above the collusion point; the simulations reported show similar patterns, with cooperation decreasing as returns to defection grow relative to those to cooperation, and as the number of agents increases. CPR simulations similarly reflect the results of experimental studies, producing significant cooperation with six agents and relatively low temptation to defect, but less as either numbers or temptation increase, and some cases in which exploitation levels exceed the Nash equilibrium. The authors argue that these results can be explained by the adaptive but myopic nature of their agents. For example, if there are only two firms in the duopoly game and both decrease their output at once from the Nash equilibrium, both will receive higher profits, even though each would have got more by maintaining the same level of output. Both are therefore likely to reduce output again, until they reach the collusion point, when profits will fall and a shift in the opposite direction becomes likely.

Lansing and Kremer's (Lansing and Kremer, 1994; Lansing, 2000) model of a Balinese agricultural socio-ecosystem does not fit neatly into a social dilemma framework, but applies ABSS to a closely related problem. In the highlands of Bali, crops are grown on rain-irrigated terraces by farmers' cooperatives called subaks. These coordinate cropping schedules through a network of water temples, each serving a group of subaks, and related by ritual ties. If all subaks planted simultaneously, water would run short, but coordinated planting by neighbours, and hence coordinated fallow periods, enhance anti-pest measures. Lansing and Kremer first modelled the production characteristics of the system under different conditions (cropping schedules, rainfall, and pest infestation). This model was validated using empirical data. A second model then tested whether the observed pattern of cropping schedules could have emerged without central direction. Starting subaks with randomly assigned schedules, and allowing them to copy more successful neighbours, gave rise to spatio-temporal patterns of cropping qualitatively similar to those observed. The situation is not quite a social dilemma, because of asymmetries between the subaks. Of two subaks on the same watercourse, the upstream member of the pair should always prefer synchronised planting, since this will tend to reduce pest damage to its crops, and will not reduce the water available for them. Whether the downstream subak would prefer synchronised or non-synchronised planting, however, depends on whether water stress or pests would cause more damage. From the point of view of the upstream subak, it is always 'cooperative' of the downstream subak to synchronise planting, but for the downstream subak, 'cooperation' from the upstream subak may well mean *non*-synchronous planting. Adding more pests to the fields would therefore tend to bring neighbouring subaks into harmony about what they want, and into synchrony. Lansing (2000) appears to identify synchrony and cooperation, but this is not justified by the structure of the subaks' incentives as he describes it. The asymmetries involved in many environmental problems related to watersheds, and the mix of similarities and differences between these problems and CPR situations, merit further ABSS research.

Antona et al. (1998) report on a fairly abstract model of the exploitation of common pool resources, constructed using the CA-like Cormas (Common-Pool Resources and Multi-Agent Systems) environment. Renewable resource dynamics are simulated on a two-dimensional grid. There are four types of agents (collectors, consumers, and two types of intermediate trader), all assumed instrumentally rational. The study explored the likely effectiveness of different approaches to conserving the resource. Two approaches to modelling economic interactions were implemented: 'general equilibrium', with a nominal 'auctioneer' setting global prices based on aggregate supply and demand; and direct inter-agent trading, with all agents located on the grid, and travel costs paid by the seller. Two types of management instrument were studied in these frameworks: imposition of a global quota (with harvesting stopped when it is filled), and the imposition of taxes, which could be levied at different stages between collector and consumer. Simulations were calibrated to produce realistic results. Prices were higher and quantities produced lower in the decentralised system (which could be due to transport costs, or more probably to the pattern of exchanges); the quota system reduced the amount taken and raised prices at later stages (consonant with expectations), while taxes made little difference — although a range of different regimes remains to be tried.

Rouchier et al. (2000) compare four Cormas ABSS models of the use of common resources. SHADOC models irrigation systems in the middle Senegal valley. JuMel simulates the seasonal movement of cattle herders, and their relations with farmers from whom they rent grazing land, in North Cameroon. Djemiong models forest hunting of the blue duiker antelope in an Eastern Cameroonian village of that name; hunters place traps on forest paths. CommonForest is a more abstract model of herders in a partially forested landscape (the forest being the resource, vulnerable to overgrazing).

In SHADOC, collectives as well as individuals are represented as agents. Water and common infrastructure constitute a shared resource, and farmers belong to a variety of organisations set up to grant loans, manage common infrastructure and share out water. Preliminary model simulations revealed social networks as more important to system viability than the 'collective water attribution rule' used. SHADOC was then designed with this in view: the 'Agent' class defines several attributes (a kitty, loans, credit, a mailbox, a list of rules, and the existence of a criterion of satisfaction) which are inherited by the 'Farmer' and 'Group' classes. Each Farmer agent represents internally the rules of each Group it belongs to, and acts accordingly. Farmers belong to different Groups for credit, water sharing and pumping management. At the end of every growing season, each agent assesses its success, and if its criterion of satisfaction is not met, enquires about the results of other agents in its network, and depending on how its 'meta-rule' for learning operates, may adopt the rules of either the agent with the best result, or the one with the most similar rules to its own. Simulations are based on patterns of rules, and a few environment parameters, used to test the effects of different simulated coordination mechanisms: they indicate that there is no good or bad set of collective rules or of individual behaviours: viability may depend on the *coherence* between sets of rules at individual and collective levels.

JuMel simulations indicate that if individual herders maintain ongoing relationships with specific villages and farmers, the grazing resource is better preserved than if they seek to minimise their costs each year. Djemiong simulations suggest that it is not enough for individual hunters to avoid overusing particular paths: coordination within groups of hunters using the same set of paths is needed to avoid declining duiker populations. In both these cases (in contrast to SHADOC), different agents *disagreeing* in their perceptions of the environment, and acting on these different perceptions, is crucial to maintaining the common resource. In CommonForest, the agents gain differing individual perceptions of how much forest remains as they move around, but also share these perceptions to build a community-level model. If agents limit their grazing in the light of this shared model, in effect agreeing on actions *not* to be taken, the resource is better preserved, and more evenly distributed spatially, than if they act on their individual models. However, in contrast to SHADOC, they do not need to coordinate their perceptions or actions beyond this point.

The conclusion drawn is that the optimal relationships between individual and collective perceptions, and action rules, depend on the nature of the common resource. If actions must be coordinated to produce the resource, common perceptions are needed. If agents must be distributed to avoid overuse, either varying perceptions of which places are best, or agreement on what places to avoid are required.

Thébaud and Locatelli (2001) describe another Cormas model focusing on conventions for sharing limited but renewable resources — in this case, driftwood on a beach. The situation modelled is based on an example from Sugden (1989), in which a convention allowed individuals to collect wood into a pile, mark it as owned, and remove it later, secure in the knowledge that it would not be taken meanwhile. The work focuses on how access rules become established in such cases. The basic model includes no conventions. Agent behaviour may be either 'aggressive' or 'conciliatory', with a given agent having a fixed probability of being aggressive. If an aggressive and a conciliatory agent simultaneously approach a piece of wood (whether in a pile or not), the former gets it (for a payoff of 1); two conciliatory agents have equal chances of getting it, for a mean payoff of .5; while two aggressive agents come into conflict, generating payoffs of -1 for each (what happens to the wood is not specified). Under these conditions, there were no marked differences in total payoff across levels of aggressiveness. In a 'peer pressure' model, agents avoid a wood pile above a certain threshold size if the owner is in sight (the presumption is that the owner would punish theft). In one variant, pile owners themselves steal others' piles if the owner is not visible; in another, they do not. In the first case, complete compliance only occurred if the agents could see the whole beach; in the second, full compliance was reached in a time which grew with the size of pile required and shrank with the range of vision. In two 'imitation' models, agents tend to adopt the strategy (to respect existing piles or not) either of agents owning large piles which they passed close to, or of a majority of visible agents. Respect for existing piles generally became established in the first case; in the second, it was favoured by low pile size threshold, and short visual range.

Carpenter et al. (1999) describe simulations of a lake subject to phosphorus flow from agriculture. The simulation system can be set up to allow the user to take part in the ongoing flow of events. The ecosystem is multistable, moving among domains of attraction distinguished by levels of phosphorus in the lake. The aggregate behaviour of agents determines the pollutant input rate; pollution lowers the value of water for irrigation and tourism. The agents have heterogeneous beliefs and access to information, and form expectations about pollution dynamics, markets and/or the actions of managers. For a wide range of scenarios, irregular oscillations among ecosystem states and agent behaviours result.

The paper describes three minimal models, each admittedly an oversimplification — the authors hope that the models' strengths and weaknesses may be complementary. Polluter agents are assumed to choose their actions stochastically, probabilistically preferring options with higher (time-discounted) returns. In the 'market manager' model, the user, playing the part of the market manager, sets the price of high-quality information, and the preference intensity of agents for more profitable options (these can also be fixed, or could be evolved using a GA, for example). In the 'governing board' model, the board regulates the pollution level, and members (who may be 'environmentalist' or 'individualist' in their policy stance) stand for re-election periodically. The 'land manager' model incorporates soil levels of phosphorus (a relatively slowly changing variable) and different land use practices. The user sets pollution targets, determining the regulations and incentives concerning land use. All three models show occasional outbreaks of high phosphorus levels in the water, resulting in economic performance declines, and a subsequent fall in pollution levels. The land manager model, incorporating the 'slow variable' of soil phosphorus level, has the most potential to reduce outbreaks. The authors end the paper by listing the limitations of the models described, which include the need to address a wider range of time scales on the ecological side, and to have more capable agents on the socio-economic side.

A companion paper (Janssen and Jager, 1999) focuses on the 'Myths of Nature' which human agents may be modelled as following in simulations of natural resource management. (These 'Myths' are deliberately caricatured sets of beliefs about how natural and human systems interact, differing in how fragile natural systems are seen to be, and how far human activities need to be limited.) The three sets of agents modelled are named 'hierarchists' (nature is moderately robust, but some centralised controls are required), egalitarians (nature is fragile, human intervention must be minimised), and individualists (nature is resilient and requires little if any protection from human activities). According to the authors, each style works well when the system in question conforms to its worldview, but badly otherwise. Simulations described involve a population of fifty agents, each with one of the three management styles corresponding to these worldviews, and able to learn using a GA (changing their perspectives in the face of contrary evidence), but ignorant of the importance of phosphorus levels in the mud. Distribution of the three perspectives among the agents at any time determines overall management style and hence influences the flow of phosphorus into the lake. A mix of the three perspectives appears to work better than any one of the three in isolation in the face of ignorance: despite imperfect learning, adaptive management maintains the resilience of the system.

The model described in Deadman (1999) and Deadman et al. (2000) is unique among those described here as a simulation of experimental studies, specifically the 'baseline CPR experiments' of Ostrom et al. (1994) [pp.105-120], which were designed within the framework of 'Institutional Analysis and Development' (IAD). In these experiments, a group of subjects could individually and without communication (but with feedback about their own performance and the group's) invest tokens in two 'markets': Market 1 produced a constant return, while market 2 had CPR characteristics, with returns per token falling as more tokens were invested. Individual decisions were highly erratic, with agents apparently often using rules of thumb leading to fluctuating bids. Aggregate results approximated the Nash equilibrium when token availability was low, but not when it was high. Deadman and colleagues' simulations gave agents an array of strategies — based on whether Market 2 returns had increased or decreased in previous rounds, on comparing returns from Markets 1 and 2 (as some of Ostrom's subjects reported in exit interviews), or on comparing the agent's own performance with group performance. Agents could switch to the strategy that would have netted them most, every third round. Group performances were similar to those reported in Ostrom et al. (1994).

In Deadman et al. (2000) parallels are drawn between ABSS and the IAD framework: the latter considers the individual an important unit of analysis, with four features: preferences, information processing capabilities, selection criteria, and resources. The agents used in the simulations are described in the same terms. They have stable preferences and are self-interested, but have bounded rationality, their selection criteria being implemented as collections of condition-action heuristics. A further set of simulations modelled experiments of Ostrom et al. (1994) on communication between the participants. Classical game theory says this should make no difference, as the participants are not bound by any agreements they make and their bids remain private. Ostrom, however, found that one-time communication after 10 rounds produced around 3/4 of the maximum return in the next five rounds, followed by a decline. With communication after each round of the second set of 10 rounds, groups earned very near the maximum with low token availability, 62% to 79% in the high availability case (Ostrom et al., 1994) [p.155]. The corresponding simulations used two approaches. In the first, after every five rounds each agent advises every other about the bid that had given it the best return, and each agent then tries one of these bids for at least one round. In the second approach, bids are submitted to a central authority which calculates which would be best if universally adopted, and this is then adopted by all (but again, agents can later switch strategy). In the first case, agents eventually locked in to a common bid — but much more slowly than experimental subjects. This bid was frequently suboptimal but never worse than the Nash equilibrium, and never involved underinvesting. In the second case lock-in was much quicker, and frequently near the optimal level, but often a few agents changed strategy shortly afterwards, causing a drop in group returns before the next communication period.

Jager et al. (2000) describe their simulations as a 'comparison of *Homo economicus* and *Homo psychologicus* models' of human behaviour in two different 16 agent, 50-year, non-spatial simulations of ecological-economic setups incorporating a CPR dilemma. All agents in a run were of one of these two types. Setup 1 had fishing as the only activity, setup 2 also had mining. Agents have needs for subsistence, leisure, identity (defined as wealth relative to similar agents) and freedom (defined as absolute wealth) — these two identifications are surely questionable. The 'need satisfaction' of an agent is a weighted product of the satisfaction of these needs. *H. economicus* agents always optimise on this product, *H. psychologicus* agents embody the 'consumat approach', a 'meta-theory', intended to combine different theories of decision-making. An agent may use *deliberation* if unsatisfied but certain of the outcomes of behaviour (optimising like *H. economicus*), *repetition* of previous activity if satisfied and certain, *social comparison* if unsatisfied and unclear (comparing behaviour with that of similar agents and selecting the behaviour with the best return), or simple *imitation* of others if satisfied and unclear. The agents never explicitly consider the dilemma, nor each others' interests.

In the fishing-only model, fish stocks decline faster with *H. economicus* agents (they all implicitly foresee the decline, but act to maximise their own need satisfaction), leading to a wealth crash, while the *H. psychologicus* agents scarcely deplete fish stocks (through 'inefficiency', not altruism or forethought). Both types end up poor, but *H. economicus* poorer. With mining permitted, the *H. economicus* model shows much less decline in fish stocks, and oscillation between time spent mining and fishing as all agents make the same decision at the same time. *H. psychologicus* agents show a gradual shift from fishing to mining as pollution causes falling catches. When different agents are given different fishing and mining abilities, greater economic inequality results; the oscillating pattern in *H. economicus* runs disappears, while the fishing to mining transition takes longer for *H. psychologicus*. When innovation is introduced (in a fishing-only models, all agents having equal ability), with agents 'inventing' new fishing methods during deliberation and learning from each other, the depletion of fish speeds up in the *H. economicus* case.

Another fishery model, intended to show how fishermen could encode beliefs about a renewable resource (i.e. fish) is described in Weisbuch and Duchateau-Nguyen (1998). The paper discusses:

> ...the dynamics of the society, represented by economic and cultural variables, coupled to the fishery represented by fish abundance.

Individual agents are not represented, so (uniquely in this section) it is not an ABSS model, but it does model a socio-ecosystem. A single ANN represents the fishermen's learning about the environment. The learning method employed uses linear predictors based on what has happened in the last 70 time steps. This tends to generate crises every 70 time steps, as information about the last crisis is lost, but the crises can be avoided by adjusting the learning procedure.

A bottom-up approach to explaining the emergence of institutions such as beliefs, norms, relationships, property rights, and agencies, based on networks of automata representing agents, is described in Weisbuch (2000). An example ABSS model, more fully described in Weisbuch et al. (1996), concerns the choice of polluting or non-polluting technology. Economic choices involving pollution relate to uncertainty with respect to social costs, and to the emergence of cooperation among actors. In the model, pollution costs feed back into the agents' decisions; if agents were (in conventional economic terms) rational and cooperative they would pay for non-polluting equipment up to the pollution cost, but a model based on bounded rationality (and unadulterated self-interest) indicates that they will agree only to pay a fraction of this. The model agents, located on a two-dimensional lattice of square cells, have no global view of real pollution costs, but notice decreased utility due to local pollution. The local nature of information exchange and pollution diffusion makes agent opinions vary across time and space. An agent's initial estimates of the utility of the competing polluting and non-polluting products is drawn from publicly available but uncertain information, and will favour the cheaper, polluting technology. These estimates, which determine the agents' choice of technology, are repeatedly updated to take into account the opinion of those in their Moore neighbourhood already using each brand. This opinion includes a negative effect proportional to local pollution, irrespective of the brand the neighbour agents have bought themselves. Pollutant concentration is calculated using partial differential equations. The relative weights given to past information and neighbours' current opinions can be varied. The most influential parameter in the model is the ratio of differences in prior utilites to maximum pollution cost. When this is low, the system evolves towards domination by nonpolluting equipment; when high, to partial domination by polluting equipment (between 75% to 100% of the market). Maximum accepted extra cost grows in proportion to the gradient of pollution cost. Transition between the two regimes is abrupt. Coexistence regions show a regular pattern of cheaters and cooperators. The existence of stable mixed configurations was a surprise to the researchers, and considering the findings of, for example, Herz (1994), described in 6.1, it seems worth investigating whether it could be an artefact of the use of a square grid. Weisbuch (2000) concludes that systems of the type modelled (non-linear elements interacting via entangled loops) can be expected to display resilience, in the sense that when a parameter is varied, a stepwise response is to be expected. Precision in prediction would require knowledge of parameters that is not available, but the dynamical properties reported are robust with respect to a wide range of parameters and for both regular lattices and random networks.

The relatively 'thick' models reviewed in this section are difficult to summarise and relate to each other, in part because of their very specificity. The collection of studies described by Rouchier et al. (2000) is of particular interest precisely because they are directly comparable with each other, and suggest ways in which real-world situations involving social dilemmas can usefully be classified. All the studies do have in common a shift of attention away from the variation of strategy across a population of agents caught in a social dilemma, towards the system-level consequences of changes in management policies, institutional arrangements, or features common to all the agents in the population. While some do make use of heterogeneity among their agents, the possibility of agents either learning or being replaced if their strategies give poor results, is given little attention except in Weisbuch (2000). In several of the other models described, it would be interesting to explore the consequences of an invasion by a small number of 'mutant' agents, behaving in a way contrary to established practice and providing larger immediate returns. Under what circumstances can an established system for regulating a CPR situation repel such an invasion? Conversely, it would be interesting to add external or internal constraints on self-interest to the model of Weisbuch (2000).

9. Implications and Applications

This paper has focused on attempts to improve scientific understanding of social dilemmas, particularly the strategies which participants in those dilemmas may adopt, and the dynamics of populations adopting different ranges of strategies in various circumstances. The wider significance of an improved understanding of social dilemmas has been dealt with only in a piecemeal fashion. This section summarises some of the broader scientific implications of such an improvement in understanding, and some possible applications.

The significance of research on social dilemmas to social scientists interested in collective action, norms, organisations and a wide range of other issues is clear. As shown particularly in sections 2 and 3, issues related to social dilemmas, cooperation and altruism arise in economics, sociology, political science, anthropology and other branches of social science. They are also of great significance to evolutionary biology, ethology, behavioural ecology and kindred fields. In both the biological and the social sciences, however, the ways in which cooperation between autonomous entities can arise and maintain itself have wider significance than has yet been suggested.

In evolutionary biology, as Maynard Smith and Szathmáry (1995) argue, many of the most important developments in the history of life, such as the initial evolution of cells, the appearance of eukaryotes, and the origins of multicellularity, have involved the emergence of fundamentally new and more complex *kinds* of biological entity through the cooperation of pre-existing simpler types, which in some sense have to forego the most immediately advantageous, 'selfish' ways of behaving, to secure the good of the whole, and hence their own longer-term advantage. Although there has been some work within the Alife field on symbiosis (Bull and Fogarty, 1995) and on the growth of complexity in evolution more generally (Nehaniv and Rhodes, 2000), it is still far from clear how these major transitions came about, how much they have in common, or how likely they (or some similar sequence of transitions) were once life had begun.

In the social sciences, the suggestion by evolutionary psychologists such as Buss (1999) that the levels of altruism toward strangers observable in modern societies are due to cultural change outpacing natural selection, whether correct or not, draw attention to a fundamental problem: how is it that societies made up of thousands, millions, or even billions of people work at all? Our ancestors presumably lived in societies made of at most a few hundred individuals until a few millennia ago. How are we as individuals able to navigate our way through mass societies, and collectively to keep them functioning? There is evidence (Fagan, 1990) that our ancestors have conducted long-distance trade for several tens of millennia; and the loss of important technologies among small populations — such as the disappearance of boat-building and fire-making techniques among native Tasmanians (Flannery, 1994, Ch.24) — suggests that being part of such extensive cultural networks has considerable survival value. Such networks would clearly be facilitated by a propensity to interact peaceably with strangers on an occasional basis. However, the routine exchange of goods and information with large numbers of strangers have no pre-urban parallel. Our ability to circumvent social dilemmas in such circumstances through communication and institution-building are surely crucial in allowing mass societies to work, and a better understanding of this ability could greatly improve our grasp of social and economic history.

Beyond its purely scientific significance, there are two main areas of application for research on social dilemmas: in the design of social policies and institutions, and in the development of MAS and related kinds of computer software, particularly in relation to the Internet.

The studies of section 8 indicate that in the first of these areas, substantial work is already underway, particularly with regard to the interaction between relatively small societies and their local ecosystems. As such work moves toward application, social dilemmas tend to be modelled as aspects of multi-faceted socio-ecosystems rather than in isolation, but there is no sharp division between basic and applied modelling work. ABSS models of social or socio-ecological systems involving social dilemmas may be used in several ways within application-driven work on social policy and related areas:

- Modelling aspects of how the system *currently* works (as in the more situation-specific Cormas models of section 8 (Rouchier et al., 2000)).

- Modelling aspects of how the system *could* work, under different policy measures, or after a change in climatic or economic conditions (Berger, 2001).

- Modelling how participants in the system *think* it works. Making ABSS models that reflect the views of a range of different participants may be of use in participatory approaches to planning (Downing et al., 2001).

- Modelling processes of participatory planning or negotiation to resolve a social dilemma. So far as we are aware, this has not yet been attempted.

It was noted in section 1 that several recent major MAS texts make little or no mention of ABSS. More broadly, as noted by Conte et al. (1998), the two apparently closely related fields of ABSS and application-oriented MAS have had little influence on each other. These authors attribute this lack of fruitful dialogue to:

- The different origins of the two fields (largely from within social science and from within distributed artificial intelligence (DAI) respectively), and the accompanying tendency of ABSS to draw on a range of social science disciplines beyond the neoclassical economics and game theory which inform application-oriented MAS.

- The concentration of ABSS on investigating theoretical hypotheses and of application-oriented MAS on software engineering problems.

- Their differing emphases on the interactions between large number of simple agents (ABSS) and smaller numbers of more complex ones (application-oriented MAS).

Conte et al. (1998) do note that both fields have drawn significantly on game theory, but we note a significant difference even here. Application-oriented MAS researchers who use game theory, such as Sandholm (1999), assume agents to be self-interested, and are primarily interested in whether there are stable, non-manipulable mechanisms which [p.203]:

> motivate each agent to behave in the desired manner.

As we have seen, empirical work on social dilemmas has uncovered evidence that real-world agents are not consistently self-interested, and analytical and simulation studies have provided possible explanations for this. Social dilemmas research has also given considerable attention to the *dynamics* of systems involving social dilemmas, frequently focusing on systems with features such as spatial structure or the ability to generate novel strategies (Lindgren and Nordahl, 1994) which increase dynamical complexity, while to MAS researchers such as Wooldridge and Jennings (1998), complex dynamics are something to avoid, since they are difficult to control.

Given this divergence of approaches, can we expect research on social dilemmas of the kinds described in this paper to contribute much to application-oriented MAS? We believe that there is indeed considerable relevance, at least to that subset of MAS applications in which the individual agents represent real-world entities with potentially opposing interests.

In MAS applications where this is not so, the use of multiple agents is simply a distinctive approach to distributed problem-solving: all agents share a common goal or goals, the main question to be resolved between them is the most efficient division of labour, and social dilemmas research does not appear relevant. An example of this class of application is provided by work on robot soccer (Veloso et al., 1998). Here, each agent has a separate robot 'body' and sensory input, but all members of a 'team' coordinate their activities toward a shared objective. (Note that in soccer teams of human players, by contrast, the team members have individual interests that may be opposed to those of their fellows, in addition to the team's common interest in defeating the opposing team.) Another example is provided by the work of Deen (1997) on 'Cooperating Knowledge-Based Systems' (CKBS). Deen (mistakenly in our view) regards much work in DAI and application-oriented MAS as aiming to simulate human social behaviour, an aim he explicitly rejects. He says [emphasis in original]:

> All our agents are *implicitly* cooperative, as they are so designed. In some tasks, agents may have to compete, but this competition takes place within a cooperation framework. We shall therefore use the term cooperation to include competition as well.

In our view, something close to this approach is actually very widespread in application-oriented MAS. For example, even when agents are involved in 'negotiations' or other quasi-social scenarios, they are frequently assumed to be 'honest' about their requirements and preferences. For example, Wooldridge and Jennings (1994), discussing cooperative problems solving, say that negotiation:

> may also involve agents lying, or being cunning and devious, though we shall not consider such cases here.

There is nothing wrong with imposing such limitations, so long as it is clear that in any real social situation, or any situation where artificial agents have been designed or programmed by different individuals or organisations, the possibility of deceit must be taken into account. Frequently, it can be proved that lying can be beneficial within a particular domain, if agents have genuinely independent and hence possibly conflicting goals. Fischer and Müller (1996), for example, distinguish 'cooperative' and 'competitive' cases within their domain (transportation); the former involves the truck fleet of a single company while the latter involves those of multiple companies. They show that in the competitive case lying can be advantageous, and attempt to find negotiation protocols which will overcome this, admitting, however, that:

> Developing such protocols for practical applications is an area of future work.

Davidsson (2000) draws a useful distinction between 'closed' and 'open' agent societies. A closed agent society is one devised by a single team of software developers to implement a complex software system; an open agent society is one to which anyone with Internet access can add agents. Intermediate ('semi-open') societies are also possible, controlled by institutions which may accept or reject applications to join. Although a closed society could include agents representing multiple, independent real-world entities, research on social dilemmas of the kinds discussed in this paper is most relevant to open and semi-open agent societies, in which agents cannot be assumed to be honest. As Davidsson notes, open and semi-open agent societies may include agents placed there for malicious purposes such as theft. If such societies are to function in such a way that honest users will be willing to place agents in them and delegate decision-making power to those agents, ways of deterring or detecting malicious agents will be required. To build our understanding of how open agent societies work, Davidsson suggests using large-scale simulations, modelling societies of artificial agents *plus* the persons or organisations owning them. Although he does not mention research into social dilemmas, much of the work described here appears relevant. Points of particular interest are:

- The contribution which *maintaining context* (i.e., allowing agents to interact repeatedly with the same partners) makes to the stability of cooperation in a social dilemma. This suggests that the managers of semi-open agent societies, and the designers of agents intended for open agent societies, should encourage repeated interactions between partners.

- The importance of *reputation* — the ability of agents to pass on information about interactions to third parties. Both this and the preceding point stress the importance of being able to *reidentify* agents: fortunately, the use of public key cryptography (Sipser, 1997, p.374) should be of great assistance here.

- The work of Ostrom et al. (1994) on CPR dilemmas suggests some additional institutional measures that can stabilise cooperation: a limit on those allowed to take part (favouring semi-open rather than open societies), explicit rules, provision for mutual surveillance and graduated sanctions for infractions.

- It may be possible to persuade some agent designers to incorporate some principles of 'fair' or 'ethical' behaviour into their agents — including a willingness to sanction infractions of these principles even at cost to the sanctioner. Agent designers, like other human beings, may actually accept such principles, and it may indeed pay to act fairly if enough other agents are doing so.

- Once agents are able to learn (by trial and error, or by imitating other agents), the extensive research on learning social dilemma strategies will become relevant.

- A form of group selection between agent societies may also operate within the wider context of the Internet: those with institutional arrangements able to protect honest agents from malicious ones will themselves gain a reputation in the world of agent-using individuals and organisations, and will tend both to increase in size and to be copied.

- We should remain aware that stable cooperation in a social dilemma can serve the interests of a dominant group rather than those of justice (Axelrod, 1986).

Finally, it is worth mentioning a possible (semi-open) testbed for some of these ideas: online role-playing games. These have become increasingly complicated, popular and profitable, with the largest apparently involving up to 500,000 people (Dodson, 2002). The players of such a game have (within the game) independent and sometimes opposed goals, while (generally) sharing the goal of keeping the game going. There is a considerable online literature on the problems of managing such games, much of which draws attention to problems reminiscent of social dilemmas. Simpson (1999) notes that players often find and exploit 'bugs' which allow players to gain advantages in unintended ways, but that as knowledge of a bug spreads, its use can cause the game to deteriorate (he cites a case where a loophole allowing players to counterfeit gold caused runaway inflation). A survey by Yee (2001) suggests that a significant proportion of players would 'hack' the game (i.e. cheat by interfering with the software) if they could. Currently, a player is either online or inactive, but it would seem feasible to design 'placeholder' agents able to take at least some possible actions on behalf of their owners while the latter are offline. Such agents would need to incorporate negotiating skills, and a degree of understanding of social interactions, including the social dilemmas implicit in these games.

Such games are thus a potential application area for social dilemmas research. They are also potential research tools. Simpson (1999) argues that such games have, in a significant sense, real economies: the players often trade within the game, and devote time and real-world money to acquiring game 'assets'. There are already entirely independent (and rather simple) agents designed by the owners of the game ('non-player characters' or NPCs), which ensure a degree of continuity (Simpson, 1999) as players go online and offline. It might be possible to use more sophisticated NPCs to test ways of maintaining (or disrupting) cooperation between allies within the game, and of maintaining the more basic level of cooperation necessary to keep the game going.

There are, however, reasons for caution here. It might be thought that such a game, with its lack of real-world consequences, is a safe experimental environment. However, these games are not an entirely self-contained world:

- Players do build real social relationships with other players, which may extend outside the game (Yee, 2001).
- Dodson (2002), notes that within-game assets have now (to the disapproval of the software's owners) begun to be traded for real money.
- The large-scale games are commercial enterprises, and their owners are likely to be wary of permitting research that could in any way threaten their monetary interests.
- Yee (2001) reports that a *majority* of players of a particular game, EverQuest, consider themselves addicted, and relays accounts of players spending many hours a day on it that they feel they should be devoting to other activities.

This last point is somewhat alarming. There are surely possibilities for researchers to set up and run such games, with the experimental subjects getting free entertainment in return for their time. However, researchers thinking of doing so should consider whether they might in effect be supplying a highly addictive drug.

10. Conclusions

The study of social dilemmas focuses attention on problems with the concept of a 'rational agent', when that agent is operating within a social context. Finite computational powers mean that agents cannot be 'rational' in the classical game-theoretic sense. Furthermore, real-world agents do not in general appear to be guided by the straightforward egotism of classical game theory players, and many of the empirical, analytical and simulation studies reviewed indicate that selective forces cannot be relied upon to eliminate motivational and behavioural quirks such as a concern for fairness, nor even to ensure that all societies of agents will show the same quirks. The complexities revealed by research on social dilemmas are thus highly relevant both to biological and social scientists, policy makers, and the designers of any MAS in which agents may represent real-world individuals or organisations.

What has ABSS contributed to this research, and where and how can it best contribute in future? In this area of social science, as in others such as innovation and imitation, organisations, and the dynamics of markets, ABSS is entering fields where there is already a great deal of work using other approaches. Insufficient attention to this work has lead to serious flaws in ABSS studies, and to an overconcentration of work in the (admittedly important) area of reciprocal altruism.

However, ABSS has already made a considerable contribution to social dilemmas research. Axelrod's early work, which is based largely on simulation and to a considerable extent on ABSS, is admitted even by his severest critics to have been highly significant, despite its flaws, particularly in stimulating research on the (biological and social) evolution of strategies in the RPD. With regard to the RPD in its simplest form, where any two agents are equally likely to meet, there appears little now to be gained from simulation-based studies of yet one more strategy or class of strategies, such as are still appearing (O'Riordan, 2000). At least, these should take full account of the main analytical results described here (Boyd and Lorberbaum, 1987; Farrell and Ware, 1989; Lorberbaum, 1994; Bendor and Swistak, 1997; Leimar, 1997).

The situation is considerably different when the PD and other social dilemmas in spatially and/or socially structured populations are considered: here, analytical techniques have up to now been much less useful, and most simulation studies belong to ABSS in the broad sense used here. However, the survey of section 6 suggests that researchers planning further ABSS work in this area need to beware of artefactual effects arising from the use of synchronous updating and regular lattices. There is a particular need for studies that will aid our understanding of how previous results (for example, of the effects of localising interaction and learning, or on the effects of synchrony) relate to each other. Studies of what happens when the structure of a population *changes* while social interaction continues — for example, by an increase in its size, or the formation of new neighbourhood or social links between agents — would be highly relevant to the kinds of social change that have occurred over recent millennia and are perhaps now happening faster than ever. There is room, and need, for both analytical and ABSS approaches to social interaction in structured populations, and for intelligent coordination of the two — as in other fields, such as percolation theory, where analytical results are hard to come by (Newman and Ziff, 2001). Simulation may well suggest where interesting results might be found, and what those results should be.

With regard to work on reputation and norms, ABSS has played an important part in revealing the complexities that arise when agents condition their strategies on the behaviour of their partners outside the direct interactions agent and partner have shared. There is also much room for further work here, particularly with regard to spatially or socially structured populations, and to the formation, transmission and internalisation of norms.

The importance of maintaining a 'good' reputation (as a cooperator and/or as a non-sucker) *among your neighbours and associates* is a fundamental feature of human life. There is a dearth of studies on norms and punishment in spatially structured populations, and of models in this area foregrounding the structure and dynamics of social networks, although expulsion from a locality or social group is an obvious form for punishment to take, and spatial, organisational and subcultural variation in norms is an obvious phenomenon in human societies.

While there have been many studies of norm enforcement in relation to social dilemmas, these have largely neglected norm formation, transmission and internalisation (Thébaud and Locatelli (2001) is a beginning here). Studies of norm internalisation would need model agents with representations of internal psychological mechanisms such as guilt and self-esteem. Currently, there is little ABSS work on social dilemmas using agents which reflect even the gross cognitive and motivational features of human beings (or for that matter other social animals), such as trial-and-error learning (as opposed to simple adoption of another agent's strategy), the effect of emotional state on attitude to risk, and the existence of *desires* to reward or punish others according to their behaviour, even at one's own expense. These all appear likely to be important in accounting for phenomena related to social norms and punishment. Comparisons of social dilemma model dynamics when agents possess or lack one of these features could be illuminating.

All the norm-related studies described assume that the payoffs from violating a norm, and the punishment for doing so, are directly comparable. In practice, this is frequently not the case: the former may be monetary, and the latter take the form of disapproval or ostracism, for example. While these *may* result in monetary loss, the expected amount of such loss is not readily calculable by the agents concerned, nor is it plausible that this is the only reason people dislike these forms of punishment. Multi-criterial decision procedures are likely to be necessary in modelling norm internalisation and consequent behaviour with any degree of realism. This would make it possible to *endogenise* the cost of violating norms, in terms of the agent's potential to reproduce or to be imitated — that is, to make this cost emerge from interactions within the model and vary over time, rather than take an externally fixed value. How far these aspects of norms can usefully be modelled in the abstract, and how far their exploration requires situation-specific, 'thick' ABSS studies remains to be seen, but an attempt to explore their significance in stabilising cooperation in the broadest possible terms would be worthwhile.

ABSS studies of norm formation and transmission require the modeling of political and cultural power within societies: the power to *change the rules*. Some of the studies in section 8 make initial moves in this direction. Studies of what happens when a socio-ecosystem that has successfully 'tamed' a social dilemma is subject to internally generated pressure requiring a revision of norms (by an increase in the number of agents, for example), or to an externally imposed change in norms, would be both theoretically intereresting, and highly relevant to current issues in environmental policy and politics.

With regard to 'thick' social dilemma simulations of socio-ecosystems, there is less competition for ABSS from other non-empirical methods than is the case with the more abstract areas considered above. However, the complexity of thick models, and the need to make links to empirical data, make them very demanding in terms of time and effort. The results may be both difficult to interpret — because of the large number of parameters a complex model contains — and rather narrowly applicable. One possible response to these problems is to assemble a set of models, built using the same or similar tools and methods, which are both sufficiently similar to be compared and sufficiently different to generate illuminating contrasts (Rouchier et al., 2000). Another is to precede any attempt at a 'thick' simulation by building and exploring a number of relevant 'thin' simulations (and/or analytical studies), in order to clarify which aspects of a specific situation to be modelled are likely to be most important in understanding its dynamics. Such an approach is advocated by Polhill et al. (2001). It should be possible to combine these approaches, producing collections of related 'thick' ABSS models, building on a common foundation of simpler models.

The judicious use of thin simulations might perhaps help to prevent the over-interpretation of results from thick ones. For example, it is sometimes stressed (Conte and Gilbert, 1995; Gilbert and Troitzsch, 1999) that complex systems with agents capable of learning, planning, and other cognitive activities as basic elements are very different from those in which the basic elements are simpler — because, for example, such agents are capable of recognising and responding to emergent patterns in the system. This is quite true, but understanding *how* such systems are distinctive requires understanding of how complex effects can result from interactions between much simpler elements. If a modeller uses highly sophisticated agents in an ABSS model, without also using simpler agents to model the same kind of social interaction, this may obscure what kinds of cognitive sophistication are and are not essential to that type of interaction.

Moreover, given the current limitations of AI models of cognition (Hofstadter and The Fluid Analogies Research Group, 1995) it is certain that the most 'sophisticated' agents in ABSS models will remain at best crude caricatures of human cognitive capacity. This certainly does not mean it is not worth studying ABSS models with complex agents, but when any interesting effect is found in such cases, attempts should be made to replicate it with simpler ones, and with ones that implement that type of complexity in different ways. Given the fundamentally social nature of human beings, this might in fact help to determine which theories of human cognition are plausible.

There is considerable room for ABSS (and other simulation and analytical methods) to be applied to explanations of human cooperation and altruism other than reciprocal altruism: kin altruism, the handicap principle, group selection, meme selection, and historical constraints. A few studies have been reviewed which deal with all of these topics except the last, but very few in relation to those focused on reciprocal altruism; yet all are at least plausible contributors to human cooperative and altruistic behaviour.

The adaptive forms of explanation — those other than explanation in terms of constraints — can be classified according to whether they involve resource altruism or not. Those that do not — reciprocal altruism, the handicap principal, and those norm-based explanations that stress the advantages to the individual of enforcing norms and of being seem to follow them — all rely on mechanisms that *blunt the social dilemma*, making altruistic behaviour a form of enlightened self-interest. Dilemma-blunting explanations contrast with *level switching* explanations (in terms of kin selection, group selection or meme selection), where an individual's altruism benefits larger or smaller structures associated with it. Docility (Simon, 1997) and some other norm-based approaches are intermediate, suggesting advantages both to individuals and to their social groups (the explanation in terms of docility, since it links these levels via individuals' bounded rationality, also has a constraint-based element). Multi-level models, in which 'agents' of different kinds exist, compete and cooperate at different scales, would seem a natural arena for ABSS research. In the human case, agents above the individual level — organisations such as firms, states, and political parties — clearly do engage in both cooperation and competition, while individuals not only belong to them, but are aware of doing so, and of at least some of the resulting advantages and disadvantages, and may seek to create or destroy them. Neither ABSS nor analytical techniques yet approach the ability to deal with social dilemmas in terms which do justice to human cognitive and political reality, although relevant multi-level models are now being developed, as section 8 showed.

Among the theoretical issues relevant to the study of social dilemmas using ABSS approaches, the relationships between genetic transmission, cultural transmission, and individual learning is both crucial, and very difficult. Many of the most important analytical findings in the area of the evolutionary stability of strategies depend (Bendor and Swistak, 1997) on models of transmission and selection such as replicator dynamics, which have their limitations even within biological contexts. Similarly, many ABSS studies of strategic dynamics depend on GAs. The transfer of these approaches to social contexts is highly problematic, and ABSS could be used in developing and comparing theories of different kinds of cultural transmission: imitation, pedagogy and advice-taking for example.

Whether cultural transmission can be treated as involving discrete chunks of information similar to genes, as the memetic model assumes, is itself doubtful (Midgely, 2000; Ingold, 2000), and the way in which individual learning transforms, recombines and selects what is transmitted culturally is one of the grounds for doubt. Moreover, human culture involves the construction of technology and of *institutions* (of which norms and organisations are two important types). These exist outside the individual mind, they both constrain and enable individual behaviour, and they in part emerge in unintended ways from social interactions, but are also the subject of deliberate design and modification. Even if a memetic model is valid, these institutions and technologies, themselves products of culture, change the way in which cultural transmission takes place (consider the effect of printing, television, and the Internet). Whether this kind of feedback can be scientifically understood using ABSS, or any other methods, is an open question.

Appendix

The tables of abbreviations in this appendix are not intended to be exhaustive. They concentrate on widely used terms relating to the Prisoner's Dilemma (tables 1 and 2); other abbreviations (table 3) are not listed if they are used only within a single short piece of text.

Table I. Basic Terminology for the PD and RPD

Term	Meaning or Explanation	Section of First Use
PD	Prisoner's Dilemma	2.2
C	Cooperate, or Cooperator: one of the two possible moves in the PD, or the player of that move	2.2
D	Defect or Defector : one of the two possible moves in the PD, or the player of that move	2.2
R	Reward: PD payoff if both players play C	2.2
P	Punishment: PD payoff if both players play D	2.2
T	Temptation: PD payoff for playing D when opponent plays C	2.2
S	Sucker's payoff : PD payoff for playing C when opponent plays D	2.2
RPD	Repeated PD	2.2
FRPD	Finitely Repeated PD: PD repeated some fixed number of times, known in advance to players	2.2
IPD	Iterated PD: PD repeated an indefinite number of times	2.2
w	Probability that another round of PD will follow the current round in the IPD	2.2
N-PD	N-person Prisoner's Dilemma	7.2
N-IPD	N-person IPD	7.2
N-FRPD	N-person FRPD	7.3

Table II. Named Strategies for the RPD

Name	Other Names	Description	Sections
Random		Play C or D, with probability 1/2	2.5
ALL-D		Play D on every move.	2.5, 4, 5, 6.2, 7.1
ALL-C		Play C on every move.	4, 5, 6.2
TFT	Tit-for-Tat	Play C on initial move; then play what partner played on previous move.	2.5, 4, 5, 6.2
GRIM		Play C unless and until partner plays D; play D thereafter.	4, 6.2
Tat-for-Tit		Play D on initial move; then switch moves whenever partner plays D.	4
STFT	Suspicious Tit-for-Tat	Play D on initial move; thereafter what partner played on previous move.	4
TF2T		Play C on first two moves; thereafter play D if and only if partner played D on last two moves.	4
WSLC	Win-Stay Lose-Change, Pavlov	Play randomly on initial move; then switch moves whenever partner plays D.	5, 6.2
CTFT	Contrite Tit-for-Tat	Only relevant if errors can occur. Like TFT, but follows an accidental D with two unconditional Cs.	5
Stochastic ALL-D		Only relevant if errors can occur. ALL-D with Cs occurring as errors.	5
GTFT	Generous Tit-for-Tat	Play C on initial move; then play C if partner played C; if partner played D, play D with probability 2/3.	5
BLINKER		Play C and D alternately	6.2

Table III. Abbreviations other than those for PD/RPD terms

Abbreviation	Abbreviated Term	Section of First Use
ABSS	agent-based social simulation	1
MAS	multi-agent systems	1
Alife	artificial life	1
AI	artificial intelligence	1
CKR	common knowledge of rationality	2.1
CPR	common pool resource	2.3
GA	genetic algorithm	2.5
ANN	artificial neural network	2.6
ESS	evolutionarily stable strategy	4
NSS	neutrally stable strategy	4
PFR	proportional fitness rule (otherwise replicator dynamics or Taylor-Jonker dynamics)	4
CA	cellular automaton (or automata)	6
DAI	distributed artificial intelligence	9

Notes

[1]See the appendix for tables of abbreviations.

[2]So called (Binmore, 1998b) because of its more or less simultaneous discovery by several researchers.

[3]When $w < 1$, the result of a contest between two strategies can be given in terms of the *expected* payoffs. When $w = 1$ these generally become infinite, and the *limit of means of payoffs* must be used instead.

[4]Given here in the terminology established above, which is somewhat different from theirs.

[5]'Win-Stay Lose-Change', also called 'Pavlov'.

[6]This has nothing to do with linkage in the geneticist's sense.

Acknowledgements

This work was funded by the Scottish Executive Environment and Rural Affairs Department, to whom we express our thanks for their support. We thank Jonathan Beecham and two anonymous reviewers for their helpful and stimulating comments on earlier versions of this paper.

References

Abreu, D. and A. Rubinstein: 1988, 'The Structure of Nash Equilibrium in Repeated Games with Finite Automata'. *Econometrica* **56**, 1259–1281.

Andreoni, J.: 1988, 'Why Free Ride? Strategies and Learning in Public Goods Experiments'. *Journal of Public Economics* **37**, 291–304.

Andreoni, J. and J. H. Miller: 1993, 'Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence'. *The Economic Journal* **103**, 570–585.

Antona, M., F. Bousquet, C. LePage, J. Weber, A. Karsenty, and P. Guizol: 1998, 'Economic Theory of Renewable Resource Management: A Multi-Agent System Approach'. In: J. S. Sichman, R. Conte, and N. Gilbert (eds.): *Multi-Agent Systems and Agent-Based Simulation: First International Workshop MABS'98*. Berlin, pp. 61–78, Springer.

Ashlock, D., M. D. Smucker, E. A. Stanley, and L. Tesfatsion: 1996, 'Preferential Partner Selection in an Evolutionary Study of Prisoner's Dilemma'. *BioSystems* **37**, 99–125.

Axelrod, R.: 1980a, 'Effective Choice in the Prisoner's Dilemma'. *Journal of Conflict Resolution* **24**, 3–25.

Axelrod, R.: 1980b, 'More Effective Choice in the Prisoner's Dilemma'. *Journal of Conflict Resolution* **24**, 379–403.

Axelrod, R.: 1981, 'The Emergence of Cooperation among Egoists'. *American Political Science Review* **75**, 306–318.

Axelrod, R.: 1984, *The Evolution of Cooperation.* Basic Books.

Axelrod, R.: 1986, 'An Evolutionary Approach to Norms'. *American Political Science Review* **80**, 1095–1111.

Axelrod, R.: 1987, 'Evolution of Strategies in the Iterated Prisoner's Dilemma'. In: L. Davis (ed.): *Genetic Algorithms and Simulated Annealing*, Vol. 24 of *Research Notes in Artificial Intelligence*. London: Pitman, pp. 32–41.

Bankes, S.: 1994, 'Exploring the Foundations of Artificial Societies: Experiments in Evolving Solutions to Iterated N-player Prisoner's Dilemma'. In: R. A. Brooks and P. Maes (eds.): *Artificial Life IV*. Cambridge MA, pp. 337–342, MIT Press.

Banks, J. S. and R. K. Sundaram: 1990, 'Repeated Games, Finite Automata, and Complexity'. *Games and Economic Behavior* **2**, 97–117.

Bendor, J. and P. Swistak: 1997, 'The Evolutionary Stability of Cooperation'. *American Political Science Review* **91**(2), 290–307.

Berger, T.: 2001, 'Agent-based Spatial Models Applied to Agriculture: a Simulation Tool for Technology Diffusion, Resource Use Changes and Policy Analysis'. *Agricultural Economics* **25**, 245–260.

Binmore, K. G.: 1994, *Game Theory and the Social Contract Volume 1: Playing Fair*. MIT Press.

Binmore, K. G.: 1998a, *Game Theory and the Social Contract Volume 2: Just Playing*. Cambridge, Mass.: MIT Press.

Binmore, K. G.: 1998b, 'Review of "The Complexity of Cooperation" by Robert Axelrod'. *Journal of Artificial Societies and Social Simulation* **1**(1). Online journal, at http://www.soc.surrey.ac.uk/JASSS/JASSS.html.

Binmore, K. G. and L. Samuelson: 1992, 'Evolutionary Stability in Repeated Games Played by Finite Automata'. *Journal of Economic Theory* **57**, 278–305.

Blackmore, S.: 1999, *The Meme Machine*. Oxford: Oxford University Press.

Boehm, C.: 1993, 'Egalitarian Behavior and Reverse Dominance Hierarchy'. *Current Anthropology* **34**(3), 227–254.

Boerlijst, M. C., M. Nowak, and K. Sigmund: 1997, 'The Logic of Contrition'. *Journal of Theoretical Biology* **185**, 281–293.

Bowles, S. and H. Gintis: 2000, 'The Evolution of Reciprocal Preferences'. Working paper available online at http://www.univ.oit/umas.edu/~gintis/papers.html.

Boyd, R.: 1989, 'Mistakes Allow Evolutionary Stability in the Repeated Prisoner's Dilemma Game'. *Journal of Theoretical Biology* **136**, 47–56.

Boyd, R. and J. P. Lorberbaum: 1987, 'No Pure Strategy is Evolutionarily Stable in the Repeated Prisoner's Dilemma'. *Nature* **327**, 58–59.

Boyd, R. and P. J. Richerson: 1988, 'The Evolution of Reciprocity in Sizable Groups'. *Journal of Theoretical Biology* **132**, 337–356.

Boyd, R. and P. J. Richerson: 1989, 'The Evolution of Indirect Reciprocity'. *Social Networks* **11**, 213–236.

Boyd, R. and P. J. Richerson: 1992, 'Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups'. *Ethology and Sociobiology* **13**, 171–195.

Brauchli, K., T. Killingback, and M. Doebeli: 1999, 'Evolution of Cooperation in a Spatially Structured Population'. *Journal of Theoretical Biology* **200**, 405–417.

Bull, L. and T. C. Fogarty: 1995, 'Artificial Symbiogenesis'. *Artificial Life* **2**(3), 269–292.

Buss, D. M.: 1999, *Evolutionary Psychology: The New Science of the Mind*. Boston, Mass.: Allyn and Bacon.

Cadsby, C. B. and E. Maynes: 1998, 'Choosing between a Socially Efficient and Free-Riding Equilibrium: Nurses versus Economics and Business Students'. *Journal of Economic Behavior and Organization* **37**, 183–192.

Camerer, C. F.: 1997, 'Progress in Behavioral Game Theory'. *Journal of Economic Perspectives* **11**(4), 167–188.

Carpenter, S., W. Brock, and P. Hanson: 1999, 'Ecological and Social Dynamics in Simple Models of Ecosystem Management'. *Conservation Ecology* **3**(2), article 4. Online journal, at http://www.consecol.org/.

Chattoe, E.: 1996, 'Why Are We Simulating Anyway? Some Answers From Economics'. In: K. G. Troitzsch, U. Mueller, N. Gilbert, and J. E. Doran (eds.): *Social Science Microsimulation.* pp. 78–104, Springer.

Cohen, M. D., R. L. Riolo, and R. Axelrod: 1999, 'The Emergence of Social Organization in the Prisoner's Dilemma: How Context-Preservation and other Factors Promote Cooperation'. Working Paper 99-01-002, Santa Fe Institute.

Conte, R. and N. Gilbert: 1995, 'Introduction'. In: N. Gilbert and R. Conte (eds.): *Artificial Societies: The Computer Simulation of Social Life.* London, pp. 1–15, UCL Press.

Conte, R., J. S. Sichman, and N. Gilbert: 1998, 'MAS and Social Simulation: A Suitable Commitment'. In: J. S. Sichman, R. Conte, and N. Gilbert (eds.): *Multi-Agent Systems and Agent-Based Simulation: First International Workshop MABS'98.* Berlin, pp. 1–9, Springer.

Cooper, B. and C. Wallace: 2000, 'The Evolution of Partnerships'. *Sociological Methods and Research* **28**(3), 365–381.

Cooper, R.: 1996, 'Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games'. *Games and Economic Behavior* **12**, 187–218.

Cronin, H.: 1991, *The Ant and the Peacock: Altruism and Sexual Selection from Darwin to Today.* Cambridge, UK: Cambridge University Press.

Crowley, P. H., L. Provencher, S. Sloane, L. A. Dugatkin, B. Spohn, L. Rogers, and M. Alfieri: 1996, 'Evolving Cooperation: The Role of Individual Recognition'. *BioSystems* **37**, 49–66.

Darwin, C.: 1968, *The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* Harmondsworth, UK: Penguin, 1st edition. Originally published 1859, John Murray, London.

Davidsson, P.: 2000, 'Emergent Societies of Information Agents'. In: M. Klusch and L. Kerschberg (eds.): *Cooperative Information Agents IV: The Future of Information Agents in Cyberspace, 4th International Workshop, CIA 2000 Proceedings.* Berlin, pp. 143–153, Springer.

Dawkins, R.: 1976, *The Selfish Gene.* Oxford: Oxford University Press, 1st edition.

Deadman, P. J.: 1999, 'Modelling Individual Behaviour and Group Performance in an Intelligent Agent-Based Simulation of the Tragedy of the Commons'. *Journal of Environmental Management* **56**, 159–172.

Deadman, P. J., E. Schlager, and R. Gimblett: 2000, 'Simulating Common Pool Resource Management Experiments with Adaptive Agents Employing Alternate Communication Routines'. *Journal of Artificial Societies and Social Simulation* **3**(2). Online journal, at http://www.soc.surrey.ac.uk/JASSS/JASSS.html.

Deen, S. M.: 1997, 'A Database Perspective to a Cooperation Environment'. In: P. Kandzia and M. Klusch (eds.): *Cooperative Information Agents First International Workshop, CIA'97 Proceedings.* Berlin, pp. 19–41, Springer.

Dessalles, J.-L.: 1999, 'Coalition Factor in the Evolution of Non-Kin Altruism'. *Advances in Complex Systems* **2**(2), 143–172.

Dieckmann, U., R. Law, and J. A. J. Metz (eds.): 2000, *The Geometry of Ecological Interactions*, Cambridge Series in Adaptive Dynamics. Cambridge UK: Cambridge University Press.

Dodson, S.: 2002, 'Lords of the Ring'. *The Guardian* pp. 21/03/2002, online supplement, pp.1–3.

Doran, J. E. and N. Gilbert: 1994, 'Simulating Societies: An Introduction'. In: J. E. Doran and N. Gilbert (eds.): *Simulating Societies: The Computer Simulation of Social Phenomena*. London, pp. 1–18, UCL Press.

Downing, T. E., S. Moss, and C. Pahl-Wostl: 2001, 'Understanding Climate Policy Using Participatory Agent-Based Social Simulation'. In: S. Moss and P. Davidsson (eds.): *Multi-Agent-Based Simulation: Second International Workshop MABS 2000*. Berlin, pp. 198–213, Springer.

Dugatkin, L. A.: 1992, 'The Evolution of the "Con Artist"'. *Ethology and Sociobiology* **13**, 3–18.

Dugatkin, L. A. and D. S. Wilson: 1991, 'Rover: A Strategy for Exploiting Cooperators in a Patchy Environment'. *American Naturalist* **138**, 687–701.

Durkheim, E.: 1964, *The Rules of Sociological Method*, . New York: Free Press, 8th edition. Translated by Sarah A. Solovay and John H. Mueller, edited by George E. G. Catlin.

Epstein, J. M.: 1997, 'Zones of Cooperation in the Demographic Prisoner's Dilemma'. Working Paper 97-12-094, Santa Fe Institute.

Eshel, I., D. K. Herreiner, L. Samuelson, E. Sansone, and A. Shaked: 2000, 'Cooperation, Mimesis and Local Interaction'. *Sociological Methods and Research* **28**(3), 341–364.

Eshel, I., L. Samuelson, and A. Shaked: 1998, 'Altruists, Egoists and Hooligans in a Local Interaction Model'. *American Economic Review* **88**(1), 157–179.

Fagan, B. M.: 1990, *The Journey from Eden: The Peopling of our World*. London: Thames and Hudson.

Farrell, J. and R. Ware: 1989, 'Evolutionary Stability in the Repeated Prisoner's Dilemma'. *Theoretical Population Biology* **36**, 161–166.

Fehr, E. and S. Gächter: 2000, 'Cooperation and Punishment'. *American Economic Review* **90**(4), 980–994.

Ferber, J.: 1999, *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence*. Harlow, UK: Addison-Wesley.

Ferriere, R. and R. E. Michod: 1996, 'The Evolution of Cooperation in Spatially Heterogeneous Populations'. *American Naturalist* **147**, 692–717.

Fischer, K. and J. P. Müller: 1996, 'A Decision-Theoretic Model for Cooperative Transportation Scheduling'. In: W. V. de Velde and J. W. Perran (eds.): *Agents Breaking Away: 7th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, MAAMAW '96*. Berlin, pp. 177–189, Springer.

Flannery, T. F.: 1994, *The Future Eaters: An Ecological History of the Australasian Lands and People*. Sydney: Reed Books.

Flood, M. M.: 1952, 'Some Experimental Games'. Technical Report RM-789-1, RAND Institute.

Flood, M. M.: 1958, 'Some Experimental Games'. *Management Science* **5**, 5–26.

Frank, R. H., T. Gilovich, and D. T. Regan: 1993, 'Does Studying Economics Inhibit Cooperation?'. *Journal of Economic Perspectives* **7**(2), 159–171.

Gale, J., K. G. Binmore, and L. Samuelson: 1995, 'Learning to be Imperfect: the Ultimatum Game'. *Games and Economic Behavior* **8**, 56–90.

Gibbons, R.: 1997, 'An Introduction to Applicable Game Theory'. *Journal of Economic Perspectives* **11**(1), 127–149.

Gilbert, N. and K. G. Troitzsch: 1999, *Simulation for the Social Scientist*. Buckingham, UK: Open University Press.

Gintis, H.: 2000a, 'Beyond *Homo economicus*: Evidence from Experimental Economics'. *Ecological Economics* **35**, 311–322.

Gintis, H.: 2000b, 'Strong Reciprocity and Human Sociality'. *Journal of Theoretical Biology* **206**, 169–179.

Gould, S. J. and R. C. Lewontin: 1979, 'The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme'. *Proceedings of the Royal Society of London B* **205**, 281–288.

Grafen, A.: 1990a, 'Biological Signals as Handicaps'. *Journal of Theoretical Biology* **144**, 517–546.

Grafen, A.: 1990b, 'Sexual Selection Unhandicapped by the Fisher Process'. *Journal of Theoretical Biology* **144**, 473–516.

Grafen, A.: 1998, 'A Note in Response to S. Siller's Comments'. *Journal of Theoretical Biology* **195**, 417–418.

Grim, P.: 1996, 'Spatialization and Greater Generosity in the Stochastic Prisoner's Dilemma'. *BioSystems* **37**, 3–17.

Hales, D.: 1998, 'Stereotyping, Groups and Cultural Evolution: A Case of 'Second Order Emergence'?'. In: J. S. Sichman, R. Conte, and N. Gilbert (eds.): *Multi-Agent Systems and Agent-Based Simulation: First International Workshop MABS'98*. Berlin, pp. 140–155, Springer.

Hales, D.: 2001, 'Cooperation without Memory or Space: Tags, Groups and the Prisoner's Dilemma'. In: S. Moss and P. Davidsson (eds.): *Multi-Agent-Based Simulation: Second International Workshop MABS 2000*. Berlin, pp. 157–166, Springer.

Halpin, B.: 1999, 'Simulation in Sociology'. *American Behavioral Scientist* **42**(10), 1488–1508.

Hamilton, W.: 1964, 'The Genetical Evolution of Social Behavior: Parts I and II'. *Journal of Theoretical Biology* **7**, 1–16 and 17–52.

Hardin, G.: 1968, 'The Tragedy of the Commons'. *Science* **162**, 1243–1248.

Hardin, G.: 1998, 'Extensions of "The Tragedy of the Commons"'. *Science* **280**, 682–683.

Hargreaves Heap, S. P. and Y. Varoufakis: 1995, *Game Theory: A Critical Introduction*. London: Routledge.

Hauert, C. and H. G. Schuster: 1997, 'Effects of Increasing the Number of Players and Memory Size in the Iterated Prisoner's Dilemma: A Numerical Approach'. *Proceedings of the Royal Society of London B* **264**, 513–519.

Hegselmann, R.: 1996, 'Understanding Social Dynamics: The Cellular Automata Approach'. In: K. G. Troitzsch, U. Mueller, N. Gilbert, and J. E. Doran (eds.): *Social Science Microsimulation*. pp. 282–306, Springer.

Henrich, J.: 2000, 'Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining Among the Machiguenga of the Peruvian Amazon'. *American Economic Review* **90**(4), 973–979.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath: 2001, 'In Search of Homo Economicus: Behavoiral Experiments in 15 Small-Scale Societies'. *American Economic Review* **91**, 73–78.

Herz, A. V. M.: 1994, 'Collective Phenomena in Spatially Extended Evolutionary Games'. *Journal of Theoretical Biology* **169**, 65–87.

Ho, T.-H.: 1996, 'Finite Automata Play Repeated Prisoner's Dilemma with Information Processing Costs'. *Journal of Economic Dynamics and Control* **20**, 173–207.

Hobbes, T.: 1914, *Leviathan*. London: J M Dent and Sons. Originally published 1651.

Hofbauer, J. and K. Sigmund: 1998, *Evolutionary Games and Population Dynamics*. Cambridge University Press.

Hoffmann, R. and N. Waring: 1996, 'The Localization of Interaction and Learning in the Repeated Prisoner's Dilemma'. Working Paper 96-08-064, Santa Fe Institute.

Hoffmann, R. and N. C. Waring: 1998, 'Complexity Cost and Two Types of Noise in the Repeated Prisoner's Dilemma'. In: G. D. Smith, N. C. Steele, and R. F. Albrecht (eds.): *Artificial Neural Nets and Genetic Algorithms: Proceedings of the First International Conference in Norwich, UK, 1997*. Vienna, pp. 619–623, Springer-Verlag.

Hofstadter, D. and The Fluid Analogies Research Group: 1995, *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. New York: HarperCollins.

Holland, J. H.: 1992, *Adaptation in Natural and Artificial Systems*. MIT Press, 2nd edition.

Howard, N.: 1971, *Paradoxes of Rationality: Theory of Metagames and Political Behavior*. Cambridge MA: MIT Press.

Huberman, B. A. and N. S. Glance: 1993, 'Evolutionary Games and Computer Simulations'. *Proceedings of the National Academy of Science, USA* **90**, 7716–7718.

Huberman, B. A. and N. S. Glance: 1998a, 'Beliefs and Cooperation'. In: P. A. Danielson (ed.): *Modeling Rationality, Morality, and Evolution*. Oxford: Oxford University Press, Chapt. 11, pp. 210–235.

Huberman, B. A. and N. S. Glance: 1998b, 'Fluctuating Efforts and Sustainable Cooperation'. In: M. Prietula, K. Carley, and L. Gasser (eds.): *Simulating Organizations*. Cambridge MA: MIT Press, Chapt. 5, pp. 89–103.

Hutson, V. C. L. and G. T. Vickers: 1995, 'The Spatial Struggle of Tit-for-Tat and Defect'. *Philosophical Transactions of the Royal Society of London B* **348**, 393–404.

Ingold, T.: 2000, 'Evolving Skills'. In: H. Rose and S. Rose (eds.): *Alas, Poor Darwin: Arguments Against Evolutionary Psychology*. London: Jonathan Cape.

Jager, W., M. A. Janssen, H. J. M. De Vries, J. De Greef, and C. A. J. Vlek: 2000, 'Behaviour in Commons Dilemmas: *Homo economicus* and *Homo psychologicus* in an Ecological-Economic Model'. *Ecological Economics* **35**, 357–379.

Janssen, M. and W. Jager: 1999, 'An Integrated Approach To Simulating Behavioural Processes: A Case Study Of The Lock-In Of Consumption Patterns'. *Journal of Artificial Societies And Social Simulation* **2**(2). Online journal, at http://www.soc.surrey.ac.uk/JASSS/JASSS.html.

Johnson, P. E.: 1999, 'Simulation Modelling in Political Science'. *American Behavioral Scientist* **42**(10), 1509–1530.

Joshi, S., J. Parker, and M. A. Bedau: 1998, 'Technical Trading Creates a Prisoner's Dilemma: Results from an Agent-Based Model'. Working Paper 98-12-115E, Santa Fe Institute.

Kahneman, D., P. Slovic, and A. Tversky: 1982, *Judgement Under Uncertainty*. Cambridge UK: Cambridge University Press.

Killingback, T., M. Doebeli, and N. Knowlton: 1999, 'Variable Investment, the Continuous Prisoner's Dilemma, and the Origin of Cooperation'. *Proceedings of the Royal Society of London B* **266**, 1723–1728.

Kirchkamp, O.: 1996, 'Spatial Evolution of Automata in the Prisoner's Dilemma'. In: K. G. Troitzsch, U. Mueller, G. N. Gilbert, and J. E. Doran (eds.): *Social Science Microsimulation*. Berlin: Springer, Chapt. 15, pp. 307–358.

Kirchkamp, O.: 1999, 'Simultaneous Evolution of Learning Rules and Strategies'. *Journal of Economic Behavior and Organization* **40**, 295–312.

Kirchkamp, O.: 2000, 'Evolution of Learning Rules in Space'. In: R. Suleiman, K. G. Troitzsch, and G. N. Gilbert (eds.): *Tools and Techniques for Social Science Simulation.* Berlin: Physica-Verlag, Chapt. 10, pp. 179–195.

Kliemt, H.: 1996, 'Simulation and Rational Practice'. In: R. Hegselmann, U. Mueller, and K. G. Troitzsch (eds.): *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View.* Kluwer, Chapt. 2, pp. 13–28.

Kraines, D. and V. Kraines: 1993, 'Learning to Cooperate with Pavlov: An Adaptive Strategy for the Prisoner's Dilemma with Noise'. *Theory and Decision* **26**, 47–79.

Kraines, D. and V. Kraines: 1995, 'Evolution of Learning among Pavlov Strategies in a Competitive Environment with Noise'. *Journal of Conflict Resolution* **39**, 439–466.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson: 1982, 'Rational Cooperation in the Finitely Repeated Prisoner's Dilemma'. *Journal of Economic Theory* **17**, 245–252.

Langton, C. G.: 1984, 'Self-Reproduction in Cellular Automata'. *Physica D* **10**, 134–144.

Lansing, J. S.: 2000, 'Anti-Chaos, Common Property, and the Emergence of Cooperation'. In: T. A. Kohler and G. J. Gumerman (eds.): *Dynamics in Human and Primate Societies*, Santa Fe Institute Studies in the Sciences of Complexity. Oxford University Press, pp. 207–223.

Lansing, J. S. and J. N. Kremer: 1994, 'Emergent Properties of Balinese Water Temple Networks: Coadaptation on a Rugged Fitness Landscape'. In: C. G. Langton (ed.): *Artificial Life III.* pp. 201–223, Addison-Wesley.

Ledyard, J. O.: 1995, 'Public Goods: A Survey of Experimental Research'. In: J. H. Kagel and A. E. Roth (eds.): *Handbook of Experimental Economics.* Princeton NJ: Princeton University Press, pp. 111–194.

Leimar, O.: 1997, 'Repeated Games: A State Space Approach'. *Journal of Theoretical Biology* **184**, 471–498.

Liebrand, W. B. G. and D. M. Messick: 1996a, 'Computer Simulations of Sustainable Cooperation in Social Dilemmas'. In: R. Hegselmann, U. Mueller, and K. G. Troitzsch (eds.): *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View.* Kluwer, Chapt. 1, pp. 235–247.

Liebrand, W. B. G. and D. M. Messick: 1996b, 'Game Theory, Decision Making in Conflicts and Computer Simulations: A Good-Looking Triad'. In: K. G. Troitzsch, U. Mueller, N. Gilbert, and J. E. Doran (eds.): *Social Science Microsimulation.* pp. 211–236, Springer.

Lindgren, K.: 1997, 'Evolutionary Dynamics in Game-Theoretic Models'. In: W. B. Arthur, S. N. Durlauf, and D. A. Lane (eds.): *The Economy as an Evolving Complex System II*, Vol. Proceedings Volume XXVII of *Studies in the Sciences of Complexity.* Reading, Mass., pp. 337–367, Addison-Wesley.

Lindgren, K. and M. G. Nordahl: 1994, 'Evolutionary Dynamics of Spatial Games'. *Physica D* **75**, 292–309.

Linster, B. G.: 1992, 'Evolutionary Stability in the Infinitely Repeated Prisoner's Dilemma Played by Two-state Moore Machines'. *Southern Economic Journal* **58**, 880–903.

Lomborg, B.: 1996, 'Nucleus and Shield: The Evolution of Social Structure in the Iterated Prisoner's Dilemma'. *American Sociological Review* **61**, 278–307.

Lorberbaum, J. P.: 1994, 'No Strategy is Evolutionarily Stable in the Repeated Prisoner's Dilemma'. *Journal of Theoretical Biology* **168**, 117–130.

Luce, R. D. and H. Raiffa: 1957, *Games and Decisions: Introduction and Critical Survey.* London: John Wiley.

Macy, M. W.: 1991, 'Learning to Cooperate: Stochastic and Tacit Collusion in Social Exchange'. *American Journal of Sociology* **97**(3), 808–843.

Macy, M. W.: 1996, 'Natural Selection and Social Learning in Prisoner's Dilemma'. *Sociological Methods and Research* **25**(1), 103–137.

Majeski, S., G. Linden, C. Linden, and A. Spitzer: 1997, 'A Spatial Iterated Prisoner's Dilemma Game Simulation with Movement'. In: R. Conte, R. Hegselmann, and P. Terna (eds.): *Simulating Social Phenomena*, No. 456 in Lecture Notes in Economics and Mathematical Systems. Berlin: Springer, pp. 161–167.

Maynard Smith, J.: 1964, 'Group Selection and Kin Selection'. *Nature* **201**, 1145–1147.

Maynard Smith, J.: 1974, 'The Theory of Games and the Evolution of Animal Conflict'. *Journal of Theoretical Biology* **47**, 209–221.

Maynard Smith, J.: 1982, *Evolution and the Theory of Games*. Cambridge, UK: Cambridge University Press.

Maynard Smith, J.: 1984, 'Game Theory Without Rationality'. *Behavioral and Brain Sciences* **7**(1), 117–125.

Maynard Smith, J.: 1998, 'Review of "Unto Others: The Evolution and Psychology of Unselfish Behavior"'. *Nature* **393**, 639–640.

Maynard Smith, J. and G. R. Price: 1973, 'The Logic of Animal Conflict'. *Nature* **246**, 15–18.

Maynard Smith, J. and E. Szathmáry: 1995, *The Major Transitions in Evolution*. W.H. Freeman.

Messick, D. M. and W. B. G. Liebrand: 1994, 'Computer Simulations of the Relation between Individual Heuristics and Global Cooperation in Prisoner's Dilemmas'. In: U. Schulz, W. Albers, and U. Mueller (eds.): *Social Dilemmas and Cooperation*. Berlin: Springer-Verlag, pp. 327–340.

Messick, D. M. and W. B. G. Liebrand: 1995, 'Individual Heuristics and the Dynamics of Cooperation in Large Groups'. *Psychological Review* **102**(1), 131–145.

Midgely, M.: 2000, 'Why Memes?'. In: H. Rose and S. Rose (eds.): *Alas, Poor Darwin: Arguments Against Evolutionary Psychology*. London: Jonathan Cape, Chapt. 5, pp. 67–84.

Miller, J. H.: 1996, 'The Coevolution of Automata in the Repeated Prisoner's Dilemma'. *Journal of Economic Behavior and Organization* **29**, 87–112.

Mitteldorf, J. and D. S. Wilson: 2000, 'Population Viscosity and the Evolution of Altruism'. *Journal of Theoretical Biology* **204**, 481–496.

Molander, P.: 1985, 'The Optimal Level of Generosity in a Selfish, Uncertain Environment'. *Journal of Conflict Resolution* **29**(4), 611–618.

Monbiot, G.: 1994, 'The Tragedy of Enclosure'. *Scientific American* **270**(1), 140.

Moss, S.: 2001, 'Messy Systems — The Target for Multi Agent Based Simulation'. In: S. Moss and P. Davidsson (eds.): *Multi-Agent-Based Simulation: Second International Workshop, MABS 2000*. Berlin, pp. 1–14, Springer.

Mueller, D. C.: 1986, 'Rational Egoism vs. Adaptive Egoism'. *Public Choice* **51**, 3–23.

Mukherji, A., V. Rajan, and J. R. Slagle: 1996, 'Robustness of Cooperation'. *Nature* **379**, 125–126.

Müller, J. P., M. P. Singh, and A. S. Rao (eds.): 1999, *Intelligent Agents V: Agent Theories, Architectures and Languages*, No. 1555 in Lecture Notes in Artificial Intelligence. Berlin: Springer-Verlag.

Nachbar, J.: 1992, 'Evolution in the Finitely Repeated Prisoner's Dilemma'. *Journal of Economic Behavior and Organization* **19**, 307–326.

Nash, J.: 1951, 'Non-cooperative Games'. *Annals of Mathematics* **54**, 286–295.
Nehaniv, C. L. and J. L. Rhodes: 2000, 'The Evolution and Understanding of Hierarchical Complexity in Biology from an Algebraic Perspective'. *Artificial Life* **6**(1), 45–67.
Newman, M. E. J. and R. M. Ziff: 2001, 'A Fast Monte Carlo Algorithm for Site or Bond Percolation'. Working Paper 01-02-010, Santa Fe Institute.
Nowak, M.: 1990a, 'An Evolutionarily Stable Strategy may be Inaccessible'. *Journal of Theoretical Biology* **142**, 237–241.
Nowak, M.: 1990b, 'Stochastic Strategies in the Prisoner's Dilemma'. *Theoretical Population Biology* **38**, 93–112.
Nowak, M. and K. Sigmund: 1989, 'Oscillations in the Evolution of Reciprocity'. *Journal of Theoretical Biology* **137**, 21–26.
Nowak, M. and K. Sigmund: 1992, 'Tit for Tat in Heterogeneous Populations'. *Nature* **355**, 250–252.
Nowak, M. and K. Sigmund: 1993, 'A Strategy of Win-Stay, Lose-Shift that Outperforms Tit-for-Tat in the Prisoner's Dilemma Game'. *Nature* **364**, 56–58.
Nowak, M. and K. Sigmund: 1998a, 'The Dynamics of Indirect Reciprocity'. *Journal of Theoretical Biology* **194**, 561–574.
Nowak, M. and K. Sigmund: 1998b, 'Evolution of Indirect Reciprocity by Image Scoring'. *Nature* **393**, 573–577.
Nowak, M. A., S. Bonhoeffer, and R. M. May: 1994, 'More Spatial Games'. *International Journal of Bifurcation and Chaos* **4**(1), 33–56.
Nowak, M. A., S. Bonhoeffer, and R. M. May: 1996, 'Robustness of Cooperation: Reply'. *Nature* **379**, 126.
Nowak, M. A. and R. M. May: 1992, 'Evolutionary Chaos and Spatial Games'. *Nature* **359**, 826–829.
Nowak, M. A. and R. M. May: 1993, 'The Spatial Dilemmas of Evolution'. *International Journal of Bifurcation and Chaos* **3**(1), 35–78.
Olson, J.: 1965, *The Logic of Collective Action*, Vol. CXXIV of *Harvard Economic Studies*. Cambridge MA: Harvard University Press.
O'Riordan, C.: 2000, 'A *Forgiving* Strategy for the Iterated Prisoner's Dilemma'. *Journal of Artificial Societies and Social Simulation* **3**(4). Online journal, at http://www.soc.surrey.ac.uk/JASSS/JASSS.html.
Ostrom, E., R. Gardner, and J. Walker: 1994, *Rules, Games and Common Pool Resources*. Ann Arbor: University of Michigan Press.
Papadimitriou, C. H. and M. Yannakakis: 1994, 'On Complexity as Bounded Rationality'. In: *Proceedings of the Twenty-Sixth Symposium on Theory of Computation (STOC-94)*. pp. 726–733, ACM.
Pedone, R. and D. Parisi: 1997, 'In What Kinds of Social Group Can 'Altruistic' Behaviors Evolve'. In: R. Conte, R. Hegselmann, and P. Terna (eds.): *Simulating Social Phenomena*, No. 456 in Lecture Notes in Economics and Mathematical Systems. Berlin: Springer, pp. 195–201.
Polhill, J. G., N. M. Gotts, and A. N. R. Law: 2001, 'Imitative Versus Non-Imitative Strategies in a Land Use Simulation'. *Cybernetics and Systems* **32**(1-2), 285–307.
Popper, K. R.: 1966, *The Open Society and its Enemies, Volume II*. London: Routledge and Kegan Paul, 5th edition.
Posch, M.: 1997, 'Win Stay-Lose Shift: An Elementary Learning Rule for Normal Form Games'. Working Paper 97-06-056, Santa Fe Institute.
Posch, M.: 1999, 'Win Stay-Lose Shift Strategies for Repeated Games — Memory Length, Aspiration Levels and Noise'. *Journal of Theoretical Biology* **198**, 183–195.

Poundstone, W.: 1992, *The Prisoner's Dilemma*. New York: Doubleday.

Probst, D.: 1996, 'On Evolution and Learning in Games'. Ph.D. thesis, University of Bonn.

Rossi, A., M. Warglien, and E. Zaninotto: 1997, 'Cooperation as Illusory Hill-Climbing: Co-adaptation and Search in Social Dilemmas'. In: R. Conte, R. Hegselmann, and P. Terna (eds.): *Simulating Social Phenomena*, No. 456 in Lecture Notes in Economics and Mathematical Systems. Berlin: Springer, pp. 169–178.

Roth, A. E.: 1988, 'Laboratory Experimentation in Economics: A Methodological Overview'. *The Economic Journal* **98**, 974–1031.

Roth, A. E.: 1995, 'Bargaining Experiments'. In: J. H. Kagel and A. E. Roth (eds.): *Handbook of Experimental Economics*. Princeton NJ: Princeton University Press, pp. 253–348.

Rouchier, J., F. Bousquet, O. Barreteau, C. L. Page, and J.-L. Bonnefoy: 2000, 'Multi-Agent Modelling and Renewable Resource Issues: The Relevance of Shared Representations for Interacting Agents'. In: S. Moss and P. Davidsson (eds.): *Multi-Agent-Based Simulation: Second International Workshop MABS 2000*. Berlin, pp. 181–197, Springer.

Routledge, B. R.: 1998, 'Economics of the Prisoner's Dilemma: A Background'. In: P. Danielson (ed.): *Modelling Rationality, Morality and Evolution*, Vol. 7 of *Vancouver Studies in Cognitive Science*. Oxford, UK: Oxford University Press, pp. 92–118.

Rubinstein, A.: 1986, 'Finite Automata Play the Repeated Prisoner's Dilemma'. *Journal of Economic Theory* **39**, 83–96.

Sacco, P. L.: 1997, 'On the Dynamics of Social Norms'. In: C. Bicchieri, R. Jeffrey, and B. Skyrms (eds.): *The Dynamics of Norms*, Cambridge Studies in Probability, Induction, and Decision Theory. Cambridge UK: Cambridge University Press, Chapt. 3, pp. 47–65.

Sandholm, T. W.: 1999, 'Distributed Rational Decision Making'. In: G. Weiss (ed.): *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge MA: MIT Press.

Sella, G. and M. Lachmann: 2000, 'On the Dynamic Persistence of Cooperation: How Lower Individual Fitness Induces Higher Survivability'. *Journal of Theoretical Biology* **206**, 465–485.

Selten, R.: 1983, 'Evolutionary Stability in Extensive 2-Person Games'. *Mathematical Social Sciences* **5**, 269–363.

Selten, R. and P. Hammerstein: 1984, 'Gaps in Harley's Argument on Evolutionarily Stable Learning Rules and in the Logic of "tit for tat"'. *Behavioral and Brain Sciences* **7**(1), 115–116.

Sen, S., A. Biswas, and S. Debnath: 2000, 'Believing Others: Pros and Cons'. In: *Proceedings, Fourth International conference on MultiAgent Systems — ICMAS-2000*. Los Alamitos, Ca, pp. 279–285, IEEE Press.

Sherratt, T. N. and G. Roberts: 1998, 'The Evolution of Generosity and Choosiness in Cooperative Exchanges'. *Journal of Theoretical Biology* **193**, 167–177.

Siller, S.: 1998, 'A Note on Errors in Grafen's Strategic Handicap Models'. *Journal of Theoretical Biology* **195**, 413–417.

Simon, H. A.: 1997, *Models of Bounded Rationality Volume 3: Empirically Grounded Economic Reason*. Cambridge, Mass.: MIT Press.

Simpson, Z. B.: 1999, 'The In-game Economics of Ultima Online'. Online at http://www.totempole.net/uoecon/uoecon.html. Presented at Computer Game Developer's Conference, San Jose, CA; Mar 2000.

Sipser, M.: 1997, *Introduction to the Theory of Computation.* Boston, MA: PWS Publishing Company.

Smith, E., S. Bowles, and H. Gintis: 2000, 'Costly Signalling and Cooperation'. Working paper, available online at http://www.univ.oit/umas.edu/~gintis/papers.html.

Sober, E. and D. S. Wilson: 1998, *Unto Others: The Evolution and Psychology of Unselfish Behavior.* Cambridge MA: Harvard University Press.

Stigler, G. J.: 1981, 'Economics or Ethics'. In: S. M. McMurrin (ed.): *The Tanner Lectures on Human Values, Vol. 2.* Cambridge, UK: Cambridge University Press, pp. 145–191.

Sugden, R.: 1986, *The Economics of Rights, Co-operation and Welfare.* Oxford: Basil Blackwell.

Sugden, R.: 1989, 'Spontaneous Order'. *Journal of Economic Perspectives* **3**(4), 85–97.

Taylor, P. and L. Jonker: 1978, 'Evolutionarily Stable Strategies and Game Dynamics'. *Mathematical Biosciences* **40**, 145–156.

Tesfatsion, L. and D. Ashlock: 1998, 'A Friendly Joust of the Minds'. *Complexity* **3**(4), 5–6.

Thaler, R. H.: 1992, *The Winner's Curse: Paradoxes and Anomalies of Economic Life.* The Free Press.

Thébaud, O. and B. Locatelli: 2001, 'Modelling the Emergence of Resource-sharing Conventions: An Agent-based Approach'. *Journal of Artificial Societies and Social Simulation* **4**(2). Online journal, at http://www.soc.surrey.ac.uk/JASSS/JASSS.html.

Tokoro, M. (ed.): 1996, *ICMAS-96: Proceedings, Second International Conference on Multi-Agent Systems.* Menlo Park, California: AAAI Press.

Trivers, R. L.: 1971, 'The Evolution of Reciprocal Altruism'. *Quarterly Review of Biology* **46**, 35–57.

Tucker, A. W.: 1950, 'A Two-Person Dilemma'. mimeo, Stanford University.

Ullmann-Margalit, E.: 1977, *The Emergence of Norms.* Oxford: Oxford University Press.

van Lange, P. A. M., W. B. G. Liebrand, D. M. Messick, and H. A. M. Wilke: 1992, 'Social Dilemmas: The State of the Art 1: Literature Review'. In: W. B. G. Liebrand, D. M. Messick, and H. A. M. Wilke (eds.): *Social Dilemmas: Theoretical Issues and Research.* Oxford, UK: Pergamon Press, pp. 3–28.

Veblen, T.: 1899, *The Theory of the Leisure Class: An Economic Study of Institutions.* New York: Macmillan.

Veloso, M., P. Stone, and K. Han: 1998, 'The CMUnited-97 Robotic Soccer Team: Perception and Multiagent Control'. In: K. P. Sycara and M. Wooldridge (eds.): *Proceedings of the Second International Conference on Autonomous Agents.* pp. 78–85, ACM.

von Neumann, J.: 1966, *The Theory of Self-Reproducing Automata.* University of Illinois Press. Edited by A.W. Burks.

von Neumann, J. and O. Morgenstern: 1944, *Theory of Games and Economic Behavior.* Princeton NJ: Princeton University Press.

Watts, D. J.: 1999, *Small Worlds: The Dynamics of Networks between Order and Randomness*, Princeton Studies in Complexity. Princeton, NJ: Princeton University Press.

Weisbuch, G.: 2000, 'Environment and Institutions: A Complex Dynamical Systems Approach'. *Ecological Economics* **34**, 381–391.

Weisbuch, G. and G. Duchateau-Nguyen: 1998, 'Societies, Cultures and Fisheries from a Modeling Perspective'. *Journal of Artificial Societies and Social Simulation* **1**(2). Online journal, at http://www.soc.surrey.ac.uk/JASSS/JASSS.html.

Weisbuch, G., H. Gutowitz, and G. Duchateau-Nguyen: 1996, 'Information Contagion and the Economics of Pollution'. *Journal of Economic Behavior and Organization* **29**, 389–407.

Weiss, G. (ed.): 1999, *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence.* Cambridge MA: MIT Press.

Williams, G. C.: 1966, *Adaptation and Natural Selection.* Princeton NJ: Princeton University Press.

Wilson, D. S. and E. Sober: 1994, 'Reintroducing Group Selection to the Human Behavioral Sciences'. *Behavioural and Brain Sciences* **17**, 585–654.

Wooldridge, M. and N. Jennings: 1998, 'Pitfalls of Agent-Oriented Development'. In: K. P. Sycara and M. Wooldridge (eds.): *Agents '98: Proceedings of the Second International Conference on Autonomous Agents.* New York, pp. 385–391, ACM Press.

Wooldridge, M. and N. R. Jennings: 1994, 'Formalizing the Cooperative Problem Solving Process'. In: M. Klein (ed.): *Proceedings of the 13th International Workshop on Distributed Artificial Intelligence (IWDAI-13).* pp. 403–417.

Wu, J. and R. Axelrod: 1995, 'How to Cope with Noise in the Iterated Prisoner's Dilemma'. *Journal of Conflict Resolution* **39**, 183–189.

Yamagishi, T. and N. Hayashi: 1996, 'Selective Play: Social Embeddedness of Social Dilemmas'. In: W. B. G. Liebrand and D. M. Messick (eds.): *Frontiers in Social Dilemmas Research.* Berlin: Springer-Verlag, pp. 363–384.

Yamagishi, T. and N. Takahashi: 1994, 'Evolution of Norms without Metanorms'. In: U. Schulz, W. Albers, and U. Mueller (eds.): *Social Dilemmas and Cooperation.* Berlin: Springer-Verlag, pp. 311–326.

Yee, N.: 2001, 'The Norrathian Scrolls: A Study of EverQuest (version 2.5)'. Online at http://www.nickyee.com/eqt/report.html.

Zahavi, A.: 1975, 'Mate Selection: A Selection for a Handicap'. *Journal of Theoretical Biology* **53**, 205–214.

Zahavi, A. and A. Zahavi: 1997, *The Handicap Principle.* Oxford, UK: Oxford University Press.

Zeggelink, E. P. H., H. de Vos, and D. Elsas: 2000, 'Reciprocal Altruism and Group Formation: The Degree of Segmentation of Reciprocal Altruists Who Prefer "Old-Helping-Partners"'. *Journal of Artificial Societies and Social Simulation* **3**(3). Online journal, at http://www.soc.surrey.ac.uk/JASSS/JASSS.html.