REVIEW ARTICLE

Spatial and temporal knowledge representation

Antony Galton

Received: 12 September 2008 / Accepted: 16 April 2009 / Published online: 12 May 2009 © Springer-Verlag 2009

Abstract Knowledge Representation (KR) originated as a discipline within Artificial Intelligence, and is concerned with the representation of knowledge in symbolic form so that it can be stored and manipulated on a computer. This article surveys that part of KR that is concerned with the representation of space and time, with particular reference to the use of such representations in geographical information science.

Keywords Knowledge representation \cdot Spatial KR \cdot Temporal KR

What is knowledge representation?

Knowledge Representation (KR) originated as a subfield of Artificial Intelligence (AI). In the early days of AI, it was sometimes imagined that to endow a computer with intelligence it would be sufficient to give it a capacity for pure reasoning; it quickly became apparent, however, that the exercise of intelligence inevitably involves interaction with an external world, and such interaction cannot take place without some kind of knowledge of that world. Thus it became clear that part of the quest for AI must involve the development of methods for endowing computer systems with knowledge. This in turn brought to the fore the question of how such knowledge is to be represented within the computer. This question can be approached in many different ways, but one can broadly distinguish between approaches which seek to discover, and thereby emulate, the forms in which knowledge is represented in the human brain, and those which take their inspiration from the external forms of representation used by humans to encode their knowledge, notably language, mathematics, and formal logic. The term Knowledge Representation, when used in the AI context, is generally taken to refer to approaches of the latter kind rather than the former, which are regarded as more within the province of Cognitive Science. We thus find that KR is characterised in the literature in terms that emphasise the quest for explicit symbolic representations of knowledge that are suitable for use by computers:

Knowledge representation and reasoning is the area of Artificial Intelligence (AI) concerned with how knowledge can be represented symbolically and manipulated in an automated way by reasoning programs. (Brachman and Levesque 2004)

... knowledge representation (the study of how to put knowledge into a form that a computer can reason with) ... (Russell and Norvig 2003)

Any intelligent declarative system will need to know an awful lot about the environment in which it is situated; knowledge representation research studies the problem of finding a language in which to encode that knowledge so that the machine can use it. (Ginsberg 1993)

Communicated by: H.A. Babaie

A. Galton (⊠) University of Exeter, Exeter, UK e-mail: apgalton@ex.ac.uk



Knowledge Representation is no longer, however, exclusively the preserve of AI, as witness the following remark by Sowa:

[T]oday, advanced systems everywhere are performing tasks that used to require human intelligence ... As a result, the AI design techniques have converged with techniques from other fields, especially database and object-oriented systems. (Sowa 2000)

In particular, Knowledge Representation is closely allied with Formal Ontology, which is concerned with the systematic enumeration and classification of the various kinds of entity represented within a given conceptualisation of the world, together with an account of their properties and relationships. Having started life as a discipline within Philosophy, Formal Ontology has become an important strand within information systems research (Guarino 1998; Smith 2004), with particular application to the problems of maintaining coherence and consistency when combining large bodies of knowledge from different sources, as happens ever more frequently with the expansion of the World Wide Web (Colomb 2007).

In this article I shall be particularly concerned with the use of KR techniques in encoding the spatial and temporal aspects of knowledge, as these aspects are fundamental to the use and development of spatial information systems in geography and the Earth sciences. Before turning to a review of these techniques, however, a few more general remarks are in order.

It should be emphasised that the term 'knowledge' implies much more than just facts, information, or data. These things can only constitute knowledge to the extent that they are situated in a context provided by some general understanding of the aspects of the world which they pertain to. To give a very simple example, a child might be able to give the correct answers to questions such as 'What is the capital of France?' and 'What country is Paris in?', and to that extent may be said to be in possession of certain facts. But does the child *know* that Paris is the capital of France? Suppose, on questioning, he proved to have no understanding of, for example, the difference between a country and a city, or the role played in the life of a country by its capital. In that case we might dismiss his supposed 'knowledge' as nothing more than parroting. Even so, it is not totally without value: imagine two children, one with a very good understanding of geography, politics, economics, etc, but a very poor memory for specific facts such as which city is the capital of which country; and another, like an idiot savant, who has an encyclopaedic command of all those specific facts, but is totally lacking in any such understanding. Together, they could make a very good team, although it would no doubt be most satisfactory for the two sets of skills to be combined in one person. Currently, computer systems resemble the second child much more than they do the first, and the whole enterprise of KR may be seen as the quest to endow them with something of the capacities of the first child. Representation of knowledge thus involves representing understanding too, typically in the form of some general model within which the facts can be represented and brought into relation with one another. Knowledge Representation as a discipline is thus more concerned with formulating such models than with the collection of individual items of knowledge, although the ultimate purpose of such models is, of course, to support that knowledge.

In the nature of the enterprise, KR as it has been practised in AI has developed strong affinities with disciplines such as philosophy and linguistics, where the nature of what exists, what we can know about it, and how that knowledge is represented are of key importance. These relationships are sometimes regarded as surprising by those whose expectations are that a computer-based subject should be more naturally allied with engineering and the natural sciences. In fact, during its short history, AI has proved receptive to influences from a wide range of disciplines, including mathematics, linguistics, philosophy, psychology, physics, and engineering, and AI researchers include amongst their ranks practitioners qualified in all these disciplines. This is also true, to some extent, of KR as a subfield of AI, but here it must be admitted that within KR there has historically been a strong emphasis (some would say bias) on the use of techniques based on formal logic.¹ In this article I shall not make any use of formal logic notations, but it should be emphasised that much of the KR literature is inaccessible to readers without at least a nodding acquaintance with the concepts and notations of first-order logic. A good introduction for the general reader is Hodges (2001).

The use of formal logic as a basic language for knowledge representation is motivated by the consideration that one of the main purposes for representing knowledge in AI is to enable reasoning about it, and logic, as the science of inference and deduction, is a natural tool to use for this. Reasoning in logic can take various



¹Not that this has been without controversy, with criticisms directed both at a perceived over-reliance on formal logical techniques (McDermott 1987) and at the underlying philosophy of representationalism (the assumption that internal representations of the world are a prerequisite for intelligent behaviour) (Brooks 1991).

forms, notably *deduction* (i.e., deriving logically valid consequences for some collection of initial statements presented as premises) and *consistency checking* (i.e., determining whether or not some set of statements can all be simultaneously true). These two processes are intimately related in that the inference of conclusion C from some list of premises P_1, P_2, \ldots, P_n will be valid if, and only if, the statements P_1, P_2, \ldots, P_n , not-C are inconsistent. Widely used formal techniques include Natural Deduction, Semantic Tableaux, and various forms of Resolution, which are particularly suitable for automated deduction and form the basis for the logic-programming language Prolog (Bratko 2009).

Such methods provide basic tools for determining the validity of inferences or the consistency of knowledge bases, but because they are extremely general they are often unsuited, as they stand, to the practical requirements of specific knowledge representation problems. In the development of expert systems, for example, various more specific forms of reasoning have been employed, notably: rule-based methods, in which knowledge is encoded in the form of 'if ... then ...' rules which can be used to establish conclusions, either in a goal-driven way (backward chaining) or a datadriven way (forward-chaining); model-based reasoning, which seeks to go beyond mere rules by encoding an understanding of the domain in the form of explicit structural models; and case-based reasoning, which uses an explicit corpus of known solutions on which to base solutions to new, but related problems. In addition, there are hybrid systems which may combine features of all of these. For more details on such methods, see Luger (2002). Although reasoning is clearly an important aspect of knowledge representation (indeed the major conference series in this area, generally abbreviated KR, has the full title 'Principles of Knowledge Representation and Reasoning'), the main emphasis in this survey will be on questions of representation as being both more immediately accessible to the novice and in a certain sense more fundamental.

Another emphasis of KR has been on what is known as 'commonsense knowledge'; this is the kind of knowledge that we humans routinely deploy in our day-do-day existence when we are not engaged in tasks that require technical knowledge and skills that have been acquired through specialist training (Hobbs and Moore 1985). Thus, for example, the knowledge of the physical world that is required is not what is presented

within the academic discipline of physics, but rather the 'naïve physics' that is shared by all humans by virtue of their intuitive understanding of how physical objects behave under the ordinary circumstances encountered in everyday life—for example, the knowledge that unsupported objects fall towards the ground, that unpowered moving objects normally come to rest eventually, that solid objects cannot occupy overlapping locations simultaneously, that liquids can be poured, and so on (Hayes 1979, 1985; Davis 2008).

In keeping with this emphasis on common sense rather than specialist or technical understanding, it has been the view of many practitioners of KR that this commonsense understanding is largely *qualitative* in nature. With reference to time and space, this has the implications that the standard mathematical models that have been of such great service in the natural sciences may be less appropriate for the purposes of emulating the spatial and temporal aspects of commonsense knowledge, and this in turn has sometimes led to conflict or misunderstanding (cf. Shoham 1988; Naur 1989).

Some aspects of everyday knowledge that have been found to be particularly problematic—and which have therefore engaged what might be thought to be a disproportionate amount of the attention of KR researchers—include *vagueness*, *uncertainty*, and *granularity*.

The problem of vagueness (or indeterminacy Burrough and Frank 1996) is that whereas much of ordinary language is imprecise in various ways, the symbolic systems devised for representing in a computer the knowledge encoded in such language are by nature precise. An example of this that has been much studied in the context of GIScience concerns the application of place names to geographical regions. Many place-names seem to be somewhat indeterminate in their spatial reference. For example, terms such as 'Central London' or 'the west of England', which are often used in everyday life, do not correspond to any precisely delineated geographical regions, since it is in some measure indeterminate just which locations they cover; yet most geographical information systems and spatial databases are only able to assign a precisely determined spatial region as the referent of such a name. This problem has been studied in relation to such examples as forests (Bennett 2001b), mountains (Fisher et al. 2004; Smith and Mark 2003), and town centres (Montello et al. 2003), and various technical solutions have been proposed for how to represent vagueness formally, e.g., using fuzzy set theory (Wang and Hall 1996), 'egg-yolk' theory (Lehmann and Cohn 1994; Cohn and Gotts 1996a, b),



²Note for the initiated: Strictly, if they are unsatisfiable. The text tacitly assumes we are working in a sound and complete logic, for which consistency and satisfiability coincide.

rough sets (Bittner and Stell 2001; Vögele et al. 2003; Worboys 1998), supervaluation semantics (Bennett 2001a; Kulik 2000; Kulik 2001; Bittner and Smith 2003), and anchoring (Galton and Hood 2005; Hood and Galton 2006).

Uncertainty is different from vagueness in that whereas the latter involves the intrinsic indeterminacy of certain terms, the former is concerned with the limitations in our knowledge. The location of Central London is vague rather than uncertain, because there simply is no fact of the matter as to where the boundaries of Central London are—it is not as if we could discover these by unearthing new facts. On the other hand, the location of Archimedes' tomb is uncertain, not vague: it must have had a precisely definable location, but we do not know for sure exactly where it was. Approaches to uncertainty include probabilistic methods (Pearl 1988) and various forms of non-monotonic reasoning (Antoniou 1997).

Granularity is different again, and concerns the level of detail with which information is recorded. Most obviously, this could be a matter of resolution: e.g., a map which records the presence or absence of a particular species of plant in each 2 km square has a finer granularity than one which is based on 10km squares. Similarly, a road map which only shows the main trunk roads has a coarser granularity than one which shows all the minor roads in addition. And a map of the USA showing the state boundaries only has a coarser granularity than one which also shows the county boundaries. This is not just a matter of scale; in general the larger the scale of a map, the finer its granularity is likely to be, but this correlation is not exact. An approach to a formal treatment of granularity in AI is proposed by Hobbs (1985). Granularity affects both time (Euzenat and Montanari 2005) and space (Schmidtke and Woo 2007) and the two in combination (Bittner 2002; Stell 2003). Granularityrelated problems arise in a specific form in cartography in relation to map generalisation, that is, the task of reducing the level of detail in a map while preserving the correct relationships amongst the features that remain (João 1998).

Temporal knowledge representation

The standard conception of time in the natural sciences is determined by the way in which times are specified and measured. Measurement of time means the measurement of the temporal duration of an event or interval; specification means locating an event within the historical time series. To some extent these can be accomplished independently, but in practice the latter is dependent on the former. The duration of a temporal interval is reckoned by counting the repetitions of some standard cyclical process and its subdivisions; the result is expressed as a real (in practice always rational!) number of units. The choice of unit employed (years, days, seconds, milliseconds, etc) depends on the context of the measurement, but it is usual to regard the second as the basic unit of time in terms of which other units may be defined. Locating an event in the time series means in the first instance locating it relative to other events. This can be done in a purely qualitative fashion, by noting e.g., that event E_1 occurred after event E_0 and before E_2 , for some suitable choice of reference events E_0 and E_2 ; and this can be made quantitative by specifying further the duration of one or both of the intervals between E_0 and E_1 and between E_1 and E_2 . In order to arrive at an 'absolute' system of temporal location, one must agree on a standard reference event relative to which all other events can be located. In the Christian calendar used in most western societies, the reference event is the notional time of Christ's birth (in reality a matter of considerable uncertainty). In cosmological contexts, events may be referred to the time of the Big Bang (also uncertain); in geology and archaeology, it is common to refer dates to the present ("B.P."); the obvious disadvantage of thisthat the date of an event changes continuously—is negated when considering events for which the uncertainty in their temporal location exceeds the time scale within which the B.P. dating system is in use. Abstracting from these considerations, we can say time is modelled as a linear series of instants isomorphic to some portion of the real number line as regards both its metric and ordering properties. An advantage of this is that it enables time to be handled mathematically in well-known and well-established ways; and to some extent this fits in with our everyday use of clocks and calendars as providing precise—or precise enough—numerical indices of time.

None the less, it might be argued that our most basic notions of temporality are essentially qualitative: that the idea that one event preceded another is conceptually more fundamental than the idea that the temporal separation of the events is a certain number of hours, say. If one takes the real number line as a model for the time line, with each number corresponding to an instant, then in addition to the quantitative precision afforded by the use of the real numbers, the model makes some strong presuppositions about the *qualitative* nature of the temporal ordering, which may be expressed in terms of a 'precedence' relation taken as the



primitive basis for the ordering. These presuppositions are as follows:

- 1. Irreflexivity: No instant precedes itself.
- 2. *Transitivity*: If *t* precedes *t'*, which in turn precedes *t''*, then *t* also precedes *t''*.
- 3. *Linearity*: Of any two distinct instants, one precedes the other.
- 4. *Past unboundedness*: For every instant, there is an instant which precedes it.
- 5. *Future unboundedness*: For every instant, there is an instant which it precedes.
- 6. *Density*: Between any two distinct instants there is a third instant which precedes one and is preceded by the other.
- 7. Continuity: If a set S of instants is such that (a) no instant in S is preceded by an instant not in S, and (b) there is at least one instant not in S, then either (c) there is a last instant in S or (d) there is a first instant not in S.

The continuity condition expresses the Dedekind property of the real numbers, which distinguishes their order type from that of the rational numbers, which also possess the first six properties above. In logical terms, it is notable that the first six properties can be expressed in first-order logic, whereas the seventh property cannot.³ Moreover, these six first-order properties are complete in the sense that *any* first-order-expressible property of the ordering on the real numbers necessarily follows, by pure logic, from the six given properties; as such, the six properties can be taken as providing an *axiomatisation* of the first-order logic of the qualitative ordering of real-valued time (Van Benthem 1983), and hence a suitable basis for qualitative reasoning.

For the purpose of qualitative temporal reasoning over particular domains, one may wish to omit or replace one or more of the axioms. For example, the density axiom might be replaced by the following

6'. *Discreteness*: If an instant has a predecessor (successor), then it has an immediate predecessor (successor),

where a predecessor is an instant which precedes the given instant, an immediate predecessor is a latest predecessor, i.e., a predecessor of the given instant which is preceded by all the other predecessors of that instant; and similarly for successor, with 'precedes' replaced by 'follows' (the inverse of 'precedes'). The axioms 1–5 and 6' characterise the first-order properties of the ordering of the integers, but again not exhaustively—to

Discrete (integer-like) time can be a useful model for situations where the system under study evolves in a sequence of discrete steps, as for example the execution of a computer program, or the moves in a board game. If one is studying a phenomenon that only displays significant structure on a time-scale of more than a day, say, then it is reasonable to use a discrete sequence of days as an idealised representation of time for the purposes of modelling that phenomenon, and if one does this then the implicit logic is that of axioms 1–5, 6' rather 1–6.

Another interesting possibility is to drop the linearity axiom (3). By doing so, one can build into the temporal model an asymmetry between the past and the future by which, at any point of time, there is a unique past but many possible futures. The possibility here may represent genuine indeterminacy, or simply uncertainty, so one can adopt this model without having to commit oneself to any particular philosophical stance on determinism and free will. This 'branching time' model can be obtained by replacing axiom 3 by

3'. *Past-linearity*: Of any two distinct instants which both precede some third instant, one precedes the other.

In Fig. 1 is shown a simple branching-time model. Here any two instants which both precede the time labelled t must lie along the line of times stretching into the past from t; these form a linear series, and hence one of the instants must precede the other, as required by axiom

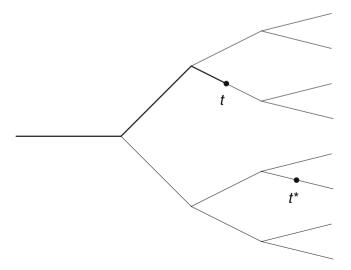


Fig. 1 A simple branching-time model



capture the integers uniquely one again needs to supplement these axioms with a property not expressible in first-order logic (e.g., that between any two instants there are at most finitely many others).

³It requires second-order logic to express it.

3'. But of the distinct times t and t^* , which do not both precede some common third time, neither precedes the other.

Thus far, time is modelled as a set of instants with some kind of ordering relation. Within AI knowledge representation, it was soon noticed that for many purposes we are more interested in intervals than instants. It is the difference between a clock and a calendar: where the clock provides us with an endless supply of time instants (e.g., 4.26 p.m., on the stroke of midnight), a calendar supplies us with intervals: days, weeks, months, and years. Moreover, if we are talking about an event or a state of affairs in the world, and ask when the event happened or the state of affairs obtained, the answer is almost always an interval: e.g., John travelled to London on August 26th; Mary was in the house from 2 pm to 3.45 p.m.

An early, and highly influential, interval-based theory of time was that of James Allen (Allen 1984; Allen and Hayes 1985). Allen showed that the qualitative relationships amongst time intervals can all be expressed in terms of a single primitive relation 'meets', where ' i_1 meets i_2 ' means that the interval i_1 ends exactly when i_2 begins. The systematic treatment of intervals, and the 13 distinct qualitative relations between them, has become known as the Interval Calculus. The relations are:

```
is before
                                      is after
<
                                      is met by
m
    meets
                                mi
                                      is overlapped by
0
    overlaps
                                oi
    starts
                                      is started by
                                si
S
                                      is finished by
f
    finishes
                                fi
    is during
                                di
                                      contains
                      equals
```

It will be seen that the relations come in pairs which are mutually inverse (e.g., if i_1 meets i_2 then i_2 is met by i_1). The 'equals' relation is unpaired since it is its own inverse. This system of 13 relations is said to be JEPD, meaning 'Jointly Exhaustive and Pairwise Disjoint'—meaning that, for any two intervals, exactly one of the relations must hold.

Freksa (1992a) neatly characterised the 13 relations of the Interval Calculus in terms of relations between the endpoints of the intervals concerned. Figure 2, adapted from Freksa's paper, illustrates this; here α and A are the beginnings of the two intervals, and ω and Ω are the corresponding endings. Clearly we must have $\alpha < \omega$ and $A < \Omega$; the 13 interval relations correspond to the possible ways of consistently assigning a relative ordering to each of remaining pairs (α, A) , (α, Ω) , (A, ω) , (ω, Ω) , as shown in Freksa's diagram.

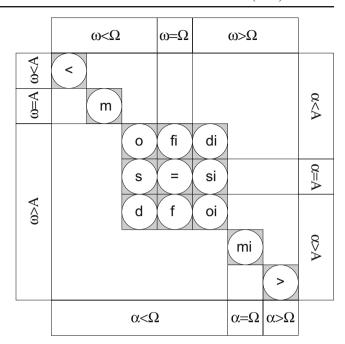


Fig. 2 The Interval Calculus relations derived from the relative ordering of the endpoints (after Freksa 1992a)

Freksa introduced the concept of *conceptual neigh-bourhood*, which he defined as follows:⁴

Definition 1 (Freksa 1992a, b): Two relations between pairs of events are (*conceptual*) *neighbours*, if they can be directly transformed into one another by continuously deforming (i.e., shortening, lengthening, moving) the events (in a topological sense).

Thus, for example, the relation o is a conceptual neighbour of s since if i_1 overlaps i_2 , then by delaying the beginning of i_1 until it coincides with the beginning of i_2 , it then holds that i_1 starts i_2 . This can be read off from Fig. 2 from the fact that the relations o and s occupy neighbouring cells in the diagram. Freksa distinguished three different neighbourhood relations corresponding to different allowed deformations, and conjectured that 'if a cognitive system is uncertain as to which relation between two events holds, uncertainty can be expected particularly between neighboring concepts'.

The Interval Calculus is governed by a comprehensive set of *composition rules*, of which a typical example is:

If i_1 overlaps i_2 and i_2 overlaps i_3 then i_1 either is before, meets, or overlaps i_3 .



⁴Note that Freksa here refers to events rather than intervals; this fully accords with Allen's declaration that intervals are of interest only insofar as they are the times at which events occur.

This is illustrated in Fig. 3, where each of the three possibilities for the relationship between i_1 and i_3 is illustrated. It should be noted that before, meets, and overlaps form a connected set ($\{<,m,o\}$) under conceptual neighbourhood—and indeed all the entries in the composition table have this property, corroborating Freksa's conjecture.

The complete set of composition rules is presented in a table (variously known as a 'transitivity table' or a 'composition table'). The idea is that a reasoner equipped with this table and an appropriate algorithm for deductive reasoning (e.g., using constraint propagation) is able to make inferences such as the following: Given that the time of the earthquake overlaps the time of the landslide, and the time of the landslide overlaps the collapse of the dam, it follows that the time of the earthquake overlaps, meets, or precedes the collapse of the dam.

Unfortunately, this kind of reasoning is rather limited, because many of the entries of the composition table contain too many alternative possibilities. The example in Fig. 3 has three; but if all we know is that i_1 precedes i_2 and i_2 follows i_3 , then the relationship between i_1 and i_3 could be any of the 13 interval calculus relations. In such a case we require information which goes beyond the purely qualitative, e.g., about the duration of the intervals, or the extent of their overlap.

More generally, for the purposes of reasoning, we need to consider not just the JEPD set of 'base' relations listed above, but also disjunctions of these. For example, if all we know about two intervals is that i starts before j, then the relation between them could be any of <, m, o, fi, and di (as can be read off from Fig. 2, expressing the constraint as $\alpha < A$), and hence may be represented as a *disjunctive* relation which we may denote $\{<,m,o,fi,di\}$. The total number of such relations is $2^{13} = 8192$, and these constitute the full *interval algebra*, often denoted A. The composition table for A thus has $8192^2 = 67$, 108, 864 entries—these do not need to be computed individually, however, as any individual entry is readily derived from the composition table for the base relations.

There is a considerable body of research into the mathematical and computational properties of the Interval Calculus and related systems. Here we shall merely give the briefest indication of the main results of this research. Of particular interest, as regards the applicability of any of these formalisms, is whether or not it is *decidable*, and if so, what is the *computational complexity* of its decision problem. A formalism is decidable if there is an algorithm, specifiable in advance, which will determine, for any collection of statements in the theory, whether or not they have a model, that is an actual example of a situation which is correctly described by those statements.

The most basic problem for reasoning over \mathcal{A} is the *constraint satisfiability problem*. An instance of this problem consists of a set of constraints of the form 'i stands in relation R to j', where 'i' and 'j' are variables standing for intervals and 'R' is any one of the 8192 relations in \mathcal{A} . Given such a set, the problem is to determine whether it is possible to associate actual intervals (represented using real-number pairs, e.g., (1.5,2.7)) with the variables appearing in the set in such a way as to satisfy all the constraints. As a simple illustration, consider the following set of constraints:

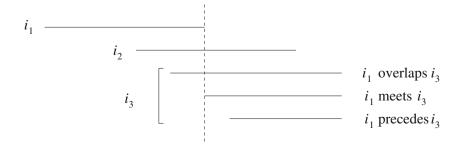
$$i_1 \text{ {m, o} } i_2$$

 $i_2 \text{ {f, =, fi} } i_3$
 $i_3 \text{ mi } i_4$
 $i_4 < i_1$

It is easy enough, in this case, to find an allocation of intervals to the variables so that these constraints are satisfied, e.g., $i_1 = (3, 4)$, $i_2 = (4, 5)$, $i_3 = (2, 5)$, $i_4 = (1, 2)$. Equally, if the third constraint is replaced by $i_3 \text{ m } i_4$, it is easy to see that there is *no* such allocation. A general solution of the constraint satisfiability problem for \mathcal{A} would be an algorithm which will correctly determine for an *arbitrary* set of constraints whether or not it is satisfiable.

An approach which provides partial solutions to this makes use of the idea of *path-consistency*. Suppose

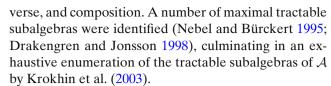
Fig. 3 Composition of qualitative relations on temporal intervals





we have a set of interval variables, say i_1, i_2, \ldots, i_n , together with constraints (i.e., A relations) for at least some pairs of intervals—a pair of intervals for which no explicit constraint is given will be allocated the 'nonconstraint' consisting of all 13 Allen relations. Such a system of relations is said to be path-consistent so long as every solution for any two variables can be extended to a solution for those two variables together any other variable. For A (though not in general), this property is equivalent to algebraic closure (Ligozat and Renz 2004), which means that for any three intervals i, j, k, if the A relation assigned to i and j is R_1 and that assigned to j and k is R_2 , then the relation assigned to i and k is contained in the composition of R_1 and R_2 as given in the composition table. Given such a set of constraints, one seeks to enforce path-consistency by systematically removing from the stated constraints any that conflict with the algebraic closure criterion, and iterating this process until either the relation assigned to some pair of intervals becomes empty, or no further changes to the relations occur. In the former case, we can infer that the original set of constraints cannot be satisfied, and we are finished. In the latter case, however, although the constraints have now been made path-consistent, we are still not guaranteed the existence of a model, since unfortunately path-consistency turns out not to be equivalent to satisfiability. The upshot of this is that in order to solve the constraint satisfiability problem in cases where path-consistency is established, further work needs to be done.

How much more work is needed? The pathconsistency algorithm of van Beek (1992) runs in cubic time—that is, the number of computational steps required is approximately proportional to the cube of the number of variables in the instance being tested—and is therefore tractable. Vilain and Kautz (1986) showed that the constraint satisfiability problem for A is NPcomplete. No tractable algorithms for NP-complete problems are currently known, and it is widely believed that none can exist; if this proves to be correct,⁵ then the amount of computational effort needed to resolve the hardest cases of the problem will always increase as an exponential function of the number of intervals. For this reason, it is important to identify restrictions to the general decision problem for A which make it tractable. A number of researchers have investigated restrictions taking the form of *subalgebras* of A, that is, subsets of the full set of interval algebra relations that are closed under the operations of intersection, con-



It is generally agreed nowadays that a model of time based on intervals alone is not wholly satisfactory: it seems that we will always want to talk about both intervals and instants. See Galton (1990) for an early critique of the Interval Calculus in this respect.

In almost any application context, we are not so much interested in time itself as in what goes on in time. Modelling time itself is of importance only insofar as it facilitates representing and reasoning about all the states, processes, and events which take place in time. An important strand in knowledge representation has therefore been concerned with establishing an appropriate formalism for handling states, processes, and events, and how they are related to each other and to the times at which they occur. Such matters have also been of concern in Philosophy and Linguistics, particularly in relation to the various forms of linguistic expression used for talking about states, processes and events, in particular the classification of different kinds of verb whose behaviour reflects the underlying logic of these categories.

The three categories of state, process, and event can be roughly characterised as follows:

- We are dealing with a state when we abstract away from any changes that might be taking place and focus on the unchanging aspects of any situation. For example, we might distinguish various states of a river: normal flow, in full spate, or reduced to a trickle. In all these states there is, of course, change happening, since the water is flowing, but in focusing on the state we are thinking about the constant properties of the flow rather than the fact that it involves many samples of water undergoing changes in position. We may say that a state holds or obtains at a particular moment or throughout a particular interval, e.g., the river was reduced to a trickle for the whole of August-and therefore in particular it was in this state at midday on August 11th which was, say, when a certain observation was made.
- In the case of a process we are concerned with change, considered as something ongoing, i.e., as it proceeds from moment to moment, rather than as a completed whole. Thus the flow of the river can be regarded as a process, involving as it does the continuous and systematic movement of water downstream. It is ongoing in the sense that we are



⁵That is, $P \neq NP$. See Garey and Johnson (1979) for details on NP-completeness and the 'P=NP' problem

not thinking in terms of something that begins and then ends but rather with something that continues throughout the period of interest; it is present at each moment during that period, and may change its character from one moment to the next, e.g., it may become faster or more turbulent, or, indeed, dwindle to a trickle. We do not say that a process holds or obtains but that it goes on, proceeds, or is in operation; and we can say that it is in operation throughout some interval or at any particular moment of that interval.

• An event also involves change, but now considered as a completed individual whole from beginning to end, e.g., a particular episode of flooding starting from when the river first overflows its banks and ending when all the floodwater has drained away. Events may last a very short time (e.g., a flash of lightning) or a very long time (e.g., the extinction of the dinosaurs); but in either case there is an interval, correspondingly short or long, over which we say that the event *happens* or *occurs*. More normally, in fact, we do not pinpoint the exact interval of its occurrence but locate it within some wider interval, as when we say that the thunderstorm occurred during the afternoon without implying that it lasted the whole afternoon.

Although these three categories are conceptually quite distinct, there are important and far-reaching relationships amongst them. Focussing in particular on processes and events, we may note that there is a mutual interdependence between them. Many processes are as it were "made of" events, and conversely many events are "made of" processes. To illustrate, consider the phenomena of coastal erosion. We may say that a particular stretch of coastal cliffs have been being eroded over many centuries, and here we are clearly referring to a process which may persist in a more or less steady fashion over a long period. Looked at more closely, however, we see that this process involves a succession of individual cliff falls, each of which is clearly an event, a bounded episode with a definite location in time. An individual cliff fall involves a mass of rock and earth sliding or tumbling down to the beach below, and this sliding or tumbling may be regarded as a process which is in operation for the duration of the fall.

Speaking in general terms, we may say that if a particular process comes into operation at a certain time t_1 and remains in operation until a later time t_2 , at which it ceases, then we have here an occurrence of an event on the interval $[t_1, t_2]$, and we could say that the event is "made of" the process. The event

and the process are not the same thing, however; for example, the event occurred over the interval $[t_1, t_2]$ but it would not be correct to say that it occurred over any proper subinterval of that interval—instead, we would say that the first half of the event occurred over the first half of the interval, and so on. On the other hand, the process was in operation over each subinterval of $[t_1, t_2]$ and indeed at each instant within it. If now we have a situation in which some event type is repeated many times over a period, we can regard this repetition as constituting an ongoing process which in principle may be continued indefinitely; thus here we have a process which is "made of" events. A simple example is the process of raining, which consists of many individual events each of which is the falling of a single drop. Each such falling happens at a particular time, but the raining process is in operation over a period which spans all those particular times; whether we see the process or many individual events is essentially a matter of the granularity of the view we are taking. Thus we see that processes and events stand in a different relation to the times at which they are located, and this has important implications for any systematic logical treatment of

Although the account just given may seem straightforward enough, it would be a mistake to give the impression that there is a consensus within the concerned academic communities as to how states, processes, events and the like should be handled, and indeed one finds a wide variety of different proposals in the literature (see Galton (2007) for a brief survey of such proposals in linguistics, philosophy, and AI). When reading the GIScience literature, one can find varied (and sometimes idiosyncratic) usage conventions regarding the key terms 'state', 'process', and 'event', but it is always a good policy here to adopt the 'principle of charity' (Wilson 1959) and seek to understand an author's usage on the assumption that what they are saying is sensible.

As noted above, there are constitution relationships running both ways between processes and events: a process can often be seen as composed of events, and equally, an event can often be seen as composed of processes. Which of these relationships you see can depend on granularity, context, and the particular focus of your interest. For example, Yuan (2001) focuses on the latter aspect, defining an event as a 'spatial and temporal aggregate of its associated processes'. This makes sense in the context of the application area under discussion—the analysis of precipitation data—where the events of interest are individual storms or showers which can be further analysed into processes representing the evolving sequence of precipitation



states at different places and times within the storm. The fact that the precipitation process consists of myriad individual raindrop falling events lies at a level of granularity far too fine to be considered in this analysis; and the long-term pattern of storm events may constitute a process when viewed at a coarse granularity, but this view likewise would seem to lie outside the purview of Yuan's study.

Spatial knowledge representation

Within KR, it is noteworthy that, for the most part, the treatment of space has lagged behind the treatment of time by approximately a decade. A number of reasons for this may be advanced, but two in particular seem paramount. First, because space has three dimensions, whereas time has only one, it affords a far greater variety of possible structures which a representational system has to handle. One has only to think here of the concept of shape, which is of importance in many spatial reasoning contexts; the variety of possible shapes in two dimensions, let alone three, already presents a formidable array of problems to anyone seeking to systematise in a tractable way the processes of representing and reasoning with spatial knowledge. In time, by contrast, the analogous concept to shape (of an interval or event) is virtually empty. The second reason for the delayed development of spatial KR has a rather specific origin in the 'poverty conjecture' of Forbus et al. (1987), that 'there is no problem-independent, purely qualitative representation of space or shape' (Forbus 1995). Forbus adduces this as an explanation for the fact that we humans are so reliant on diagrams and other perceptual representations for our spatial reasoning, since these can capture metric properties absent from a purely qualitative representation, and it is clear that many spatial reasoning tasks cannot be accomplished without access to at least some metric information. But Forbus does not on that account dismiss qualitative spatial representations as useless: on the contrary, he proposes (in the MD/PV—that is, 'metric diagram' plus 'place vocabulary'-framework of Forbus et al. (1987)) a marriage of qualitative and quantitative representation that perhaps better reflects at least some of our human spatial reasoning practices.

That the poverty conjecture was advanced for space but not for time reflects the dimensionality difference already alluded to. Apart from the difference in dimensionality, a second key difference between time and space is that time, unlike space, has an intrinsic *directedness*, which we perceive as the asymmetry between the past and the future. Time flows, we say, and always in the same direction; there is no 'going backwards' in time (whatever that could mean), whereas in space movement is possible in any direction.

Despite these two major differences between space and time, there are many formal analogies between the structures they present. Some examples are:

instant — point
interval — region
endpoints of interval — boundary of region
event occurs in interval — object located in region
process goes on during — stuff fills region
interval

as well as common relations such as overlap, adjacency and containment, and common problems such as discrete vs continuous (Galton 1999), vagueness and indeterminacy, granularity (Euzenat 1995).

All this should suggest the possibility of a formal treatment of space and spatial relations that is analogous to the treatments of time exemplified by the Interval Calculus. Within AI, the breakthrough occurred in a landmark paper by Randell et al. (1992) which introduced an exhaustive set of qualitative spatial relations that is analogous to the thirteen temporal relations of the Interval Calculus; this system has come to be known as the Region Connection Calculus (RCC) and has been the subject of intensive research for over a decade. At the same time, within the Geographical Information Science community, Egenhofer and collaborators were independently pursuing work (Egenhofer 1989, 1991; Egenhofer and Franzosa 1991) that exhibits some remarkable parallels with the RCC work, despite a number of significant differences. Egenhofer's work is the more familiar of the two within the GIScience community, but because RCC exhibits greater continuity with the temporal work discussed in the preceding section, we shall discuss this first.

The key notions of RCC, as suggested by the title, are *region* and *connection*. A region is an extended portion of space; the notion can be applied in any number of dimensions, the only caveats being that regions must have the same dimensionality as the space in which they are regarded as being embedded,⁶ (e.g., two-dimensional regions on the *surface* of the earth, three-dimensional regions in the *interior* of the earth), and that all the regions under consideration should have the *same* dimensionality—thus RCC, at least in its original form, cannot speak of the relation between,



⁶In other words, the regions should have *co-dimension zero*.

say, the water in a lake (occupying a three-dimensional region) and a stretch of the lake shore (occupying a one-dimensional region).

Connection is to be understood as a relation between regions, taken as a primitive in the logical theory; there are a number of different ways in which it can be interpreted (Cohn and Varzi 1998, 1999), but the essential idea is that two regions are to be regarded as connected⁷ so long as they are not separated—for example, they may touch, overlap, or coincide, or one may be contained in the other, and all these count as cases of connection. Beginning with connection (C) as the primitive relation, other relations between one region (R_1) and a second (R_2) can be defined as follows:

Relation	Symbol	Meaning
R_1 is disconnected from R_2	DC	R_1 and R_2 are not connected.
R_1 is part of R_2	P	Every region connected to R_1 is connected to R_2 .
R_1 overlaps R_2	O	Some region is part of both R_1 and R_2
R_1 is discrete from R_2	DR	R_1 does not overlap R_2
R_1 is externally connected to R_2	EC	R_1 and R_2 are connected but do not overlap
R_1 partially overlaps R_2	PO	R_1 overlaps R_2 but neither is part of the other
R_1 is equal to R_2	EQ	Each of R_1 and R_2 is part of the other
R_1 is a <i>proper</i> part of R_2	PP	R_1 is part of R_2 but not equal to it
R_1 is a tangential proper part of R_2	TPP	R_1 is a proper part of R_2 and some region is EC to both.
R_1 is a non- tangential proper part of R_2	NTPP	R_1 is a proper part of R_2 but not a TPP

Some of these relations are *symmetric*, i.e., if the relation holds between R_1 and R_2 then it automatically holds between R_2 and R_1 as well (in other words, the relation holds of the two regions without reference to the order in which they are considered): these are C, DC, DR, O, PO, EC, and EQ. The remaining relations, namely P, PP, TPP, and NTPP, are not symmetric, so they can hold between two regions taken in one order without holding between the same two regions taken in the opposite order;8 each of these non-symmetric relations has an inverse, represented as PI, PPI, TPPI, NTPPI, such that if R_1 and R_2 are PP, then R_2 and R_1 are PPI, and so on. Amongst these 15 relations, eight are singled out as forming a JEPD set analogous to the 13 relations of the Interval Calculus, namely DC, EC, PO, EQ, RPP, NTPP, TPPI, and NTPPI. These form the system RCC8, and Fig. 4 shows the wellknown Conceptual Neighbourhood Diagram for these relations.

As with the Interval Calculus, a composition table can be made for the RCC8 relations, which can be used as a basis for qualitative reasoning about location. An example is that if R_1 is externally connected to a non-tangential proper part of R_2 then the relation between the two regions must be either partial overlap or proper part (either tangential or non-tangential). This is illustrated in Fig. 5.

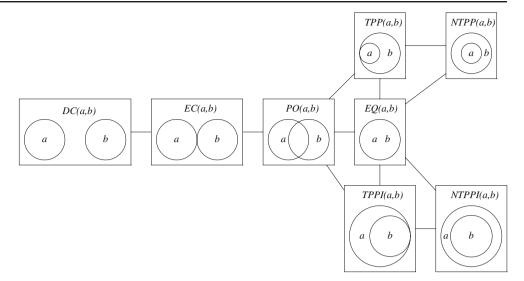
As with the Interval Calculus, considerable efforts have been expended in investigation of the mathematical and computational properties of RCC. Bennett (1994) showed that, like the Interval Calculus, RCC is decidable, but again, as with the Interval Calculus, the decision problem is known to be NP-complete (Renz and Nebel 1999). Renz and Nebel (1999) identified a maximal tractable fragment of RCC-8, and this was extended to a complete analysis of tractability in Renz (2002). For spatio-temporal reasoning (Wolter and Zakharyaschev 2000) established some partial complexity results for a combination of RCC-8 with a certain discrete point-based temporal logic.

The qualitative spatial properties expressed by RCC combine *mereology* (the study of the part-whole relation) with *topology* (the study of the connection relation). Accordingly, this approach to spatial representation and reasoning goes under the name of *mereotopology*. Inspired by RCC, numerous other systems of qualitative spatial relations have been introduced, handling more specialised aspects such as direction and orientation (Freksa 1992b; Frank 1992,

⁷Beware of a confusion here with the normal mathematical use of the term 'connected' to refer to a region which cannot be expressed as the union of two disjoint non-empty sets that are open in the region's subspace topology, and the related (but not equivalent) definition of a 'path-connected' region as one within which any two points can be joined by a continuous path lying wholly within it. To avoid overloading the term in this way, some authors prefer to use the term 'contact' to refer to the binary relation discussed here. This is sensible, but has not been widely adopted.

⁸In fact PP, TPP, and NTPP are *asymmetric*, meaning that if any of them holds between R_1 and R_2 , say, then it *cannot* hold between R_2 and R_1 .

Fig. 4 The RCC8 system of qualitative spatial relations, arranged in the form of a Conceptual Neighbourhood Diagram



1996), distance (Hernández et al. 1995), shape (Cohn 1995; Schlieder 1996), modes of overlap (Galton 1998), occlusion (Galton 1994; Randell et al. 2001), and dimensionality (Galton 1996). It is probably fair to say that none of these systems has achieved anything like the definitive status of RCC within the KR community.

As already noted, within GIScience, the work of Egenhofer and his collaborators (Egenhofer 1989, 1991; Egenhofer and Franzosa 1991) is more familiar than RCC. The parallels with RCC extend to a duplication, in Egenhofer's own formalism, of the conceptual neighbourhood graph for RCC-8 (Egenhofer and Al-Taha 1992). Egenhofer's approach is entirely different from RCC, although the goal, to characterise the possible mereotopological relations amongst spatial regions, is essentially the same.

In his first formulation (Egenhofer 1989), Egenhofer considered each region as characterised by its interior and boundary, and then described the topological rela-

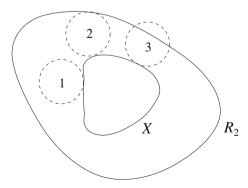


Fig. 5 Composition of RCC relations: If R_1 is EC to X and X is NTPP to R_2 , then R_1 (the dashed circle) must be one of (1) NTPP, (2) TPP, or (3) PO to R_2

tion between two regions by means of a 2×2 matrix, called the *4-intersection*, indicating the emptiness or otherwise of the intersection between the interior (*int*) and boundary (bdy) of one with both the interior and boundary of the other:

$$bdy(R_1)\begin{pmatrix} a & b \\ int(R_1) & c & d \end{pmatrix}$$

Here a, b, c, d are 0 or 1 depending on whether or not the corresponding intersection is empty. For regions of co-dimension zero, eight of the 16 possible matrices correspond exactly to the RCC-8 relations, so long as both regions are assumed to consist of a single connected component (see Fig. 6), and the other eight cannot occur; but the scheme can be applied equally to regions of higher co-dimension, so that, for example, for relationships between regions of co-dimension 1 (e.g., lines in the plane), all 16 matrices correspond to possible relationships.

In a subsequent formulation (Egenhofer 1991), Egenhofer considered the exterior (*ext*) of each region in addition to its interior and boundary, leading to the *9-intersection* matrix, which allows finer discriminations amongst the relations involving regions with higher codimension. 'Exterior' here simply means all of the space

Fig. 6 Correspondence between RCC-8 and 4-intersection



under consideration that is not covered by the interior and boundary; note that for regions of positive codimension, the boundary and interior, in Egenhofer's sense, can only be understood as the topological boundary and interior with respect to some subspace topology having the same dimensionality as the region, whereas the exterior is the topological exterior with respect to the space as a whole. In later work, Egenhofer has gone on to consider relations between, amongst others, regions in discrete space (Egenhofer and Sharma 1993), regions with holes (Egenhofer et al. 1994), regions on the surface of a sphere (Egenhofer 2005), directed line segments (Kurata and Egenhofer 2006), and also more discriminating versions of 4-intersection obtained by including more information about the intersections (Egenhofer and Franzosa 1995).

Perhaps the most significant difference between RCC and the Egenhofer systems is that the latter, but not the former, readily accommodates regions of positive co-dimension, allowing the expression of relations between solids, surfaces, edges, and points, which cannot be handled simultaneously in RCC. This is important in geographical applications, for example to express the relationship between a linear region L and an area A. To see how this works, consider the three cases shown in Fig. 7. The boundary of the linear region is considered to consist of its two endpoints. The 9-intersection matrices for these three cases are

showing that this scheme is able to discriminate these qualitatively distinct configurations.

Another important difference is that Egenhofer's methods can be applied with equal ease to discrete spaces (e.g., the pixels of a computer display), whereas in RCC the definition of 'part' (namely, that X is part of Y if and only if everything connected to X is also



Fig. 7 Three relationships between a line and an area

connected to Y) leads to bizarre and counterintuitive consequences when applied in discrete space.⁹

Moving beyond RCC and the Egenhofer schemes, perhaps the most important of the additional qualitative spatial relations mentioned above, in the geographical context, are distance and direction. Qualitative distance is expressed using terms such as 'near' and 'far'; qualitative direction using 'left', 'right', 'ahead', and 'behind', and 'north', 'south', 'east', and 'west'. A number of different formal schemes for handling relations of this sort have been proposed. We illustrate three such schemes for direction in Fig. 8. The conical and projective schemes are discussed by Frank (1992, 1996), while the double-cross scheme is due to Freksa (1992b). Both the conical and projective schemes have eight cardinal directions, or nine if one counts the centre. The central shaded square in the projective scheme, which Frank describes as a 'neutral' area ("here"), could be used to represent points sufficiently near to the reference point not to be assigned a cardinal direction (in keeping with Frank's observation that cardinal directions are 'almost exclusively used in large-scale space'). The double-cross scheme relates the position of a point to a directed line segment which may be thought of as joining a point of observation to a reference point. The labels attached to them in Fig. 8c correspond to those used by Scivos and Nebel (2001), who investigate the computational properties of reasoning in the double-cross calculus.

For any of these schemes, unrestricted composition quickly leads to loss of information. For example, if A is north of B and B is south-east of C, then in the conical scheme A could be north, north-east, east, or south-east of C, while in the projective scheme it could be south of C as well. Frank (1996) presents instead composition tables derived using a somewhat mysterious 'averaging rule' which results in each entry of the table consisting of a single relation, though this is in many cases qualified as 'approximate'.

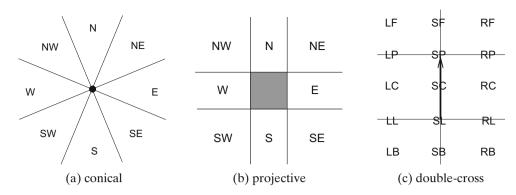
Combining space and time

Although, as noted, KR came to grips with the formal treatment of time before that of space, in GIScience, the development has, for obvious enough reasons, been the other way round. Space being the primary domain here—since a geographical phenomenon is necessarily spatial—the early development of GIS naturally

⁹Specifically: any region all of whose pixels are adjacent to pixels outside the region would be part of its complement.



Fig. 8 Schemes for qualitative direction and orientation (**a-c**)



concentrated on ways of handling space and spatial relationshiops without reference to the temporal dimension. As a result, GIS evolved into a highly sophisticated medium for representing and, to some extent, reasoning about the spatial relationships amongst geographical entities existing at one moment in history. This is, in effect, a snapshot of the world, although it is natural to assume that the spatial properties under consideration are sufficiently durable that the 'snapshot' can be considered to remain valid over a long enough period to be useful without having to be continuously updated. Many geographical facts do indeed have the required durability: on the whole, the boundaries of countries and other political and administrative units remain stable over periods of years (many years in many cases), as do most geographical features both natural (e.g., lakes, rivers, mountains, coasts) and man-made (e.g., roads, railways, buildings, towns): the admittedly conspicuous exceptions (e.g., the dramatic changes in political boundaries in eastern Europe after the end of the Cold War, areas of relatively rapid coastal erosion) only serve to highlight the relative stability of the rest. Without such stability, a 'snapshot' of the world at one time would be of limited use beyond that time—as is easily seen by considering, as a bizarre thought experiment, the idea of a map which purported to give the location of every human and every car.

In reality, one is almost never presented with a truly instantaneous snapshot, since the collection of data must happen over an extended temporal period. Still, it is not unreasonable to publish a map purporting to give, say, the layout of streets and important buildings in a particular town as they were at the time of publication, even though we know there are likely to be some inaccuracies arising, not from inaccurate surveying, but from changes that occur between the time that a particular data item was recorded and the time that the map was published. As an idealisation, then, we may take the idea of a snapshot view as being

the paradigm of *atemporal GIS*, that is, GIS in which the time-dimension is entirely suppressed.

From here the natural first step in the incorporation of time into GIS is to consider a time-indexed *sequence* of snapshots. This is identified as the Stage I in the 'brief history of time in GIS' presented by Worboys (2005). But as Worboys pointed out, a sequence of snapshots does not hold any explicit information about change in objects or about processes and events. Worboys designates the introduction of explicit reference to change as Stage II in his brief history, while Stage III brings in a 'full-blooded treatment of change, in terms of events and actions'.

The need for GIScience to handle processes, events and other temporal phenomena has been highlighted by many authors in recent years (Peuquet 1994; Claramunt and Thériault 1995; Yuan 2001; Worboys 2005), but despite these repeated calls to action, and many individual successes, it remains true that GI-Science has not yet evolved a stable, agreed set of conceptual and formal tools for this comparable to what has been available for many years as regards the purely spatial part of the enterprise. Within the KR community, as we have seen, there has been much work devoted to the development of good formalisms for space, time, motion, change, etc., and it is natural to ask whether it is from this work that GIScience will eventually acquire the much-needed formal and computational tools to enable it to handle temporal phenomena with the ease and fluency already manifested in its handling of the purely spatial. That this may be so is



¹⁰It may of course contain *implicit* information, simply because if a particular spatial property holds in one snapshot but not in the next later one, then it follows that a change with respect to that property must have occurred in the interval between the two snapshots—but regarding the nature of that change, e.g., whether it was a sudden event occurring effectively instantaneously, or a more drawn-out gradual process going on throughout the interval, the succession of snapshots remains silent.

suggested by the increasing interest taken in GIScience by practitioners of KR and related disciplines over the past 20 years or so (as witness, for example, the attendance of many such practitioners at conference series such as COSIT and GIScience).

The remainder of this section will concentrate on a number of specific themes which involve both time and space and are of importance to GIScience: continuity, causality, and identity.

Continuity is a fundamental organising principle for spatial and temporal data. Because most effects diminish with distance (in space or time), there is a general tendency for nearby locations to be more similar to one another than more distant locations. This is the phenomenon of autocorrelation, which forms the basis of most techniques for spatial and temporal interpolation.¹¹ For real-valued variables, this leads to a presumption of continuity in the strict mathematical sense, and this presumption underlies most widelyused methods of interpolation. Interpolation is a form of non-monotonic reasoning: that is, reasoning that leads to conclusions regarded as plausible but not logically guaranteed, and hence vulnerable to falsification after further data acquisition (Antoniou 1997). Nonmonotonic reasoning has been the subject of intensive study in the KR community, with the aim of devising extensions to or replacements for standard logics to emulate the plausible reasoning of everyday life. The case of temporal non-monotonic reasoning (Sandewall and Shoham 1995; Shanahan 1997) has received particular emphasis, key notions here being various forms of change minimisation: e.g., persistence, the principle that, in the absence of any reason to infer that a change has taken place, it can be assumed by default that nothing has changed; and more generally, the selection of that inference which minimises the amount of unexplained change, in effect maximising temporal autocorrelation. Where, as in these cases, the variable to be interpolated has a discrete value range (e.g., truth values for propositions in non-monotonic logics), there can be no continuity in the mathematical sense; autocorrelation in effect takes the place of strict continuity here. Something similar is seen in the conceptual neighbourhood diagrams for systems of relations such as RCC; under continuous motion or deformation, the RCC relation

between two regions will trace a path in the conceptual neighbourhood diagram. Regarding neighbouring relations as 'close', this is simply a form of temporal autocorrelation. The autocorrelation found in a discrete variable may reflect continuity in some underlying continuous variables on which the discrete variable depends; for example, the discrete, qualitative RCC-8 relations depend on the continuous, quantitative characteristics of the regions, and the conceptual neighbourhood diagram for RCC-8 can be derived from a quantitative description of the space of possible regions by means of a precisely definable homomorphic mapping (Galton 2001). Depending on which continuous variables are chosen, different conceptual neighbourhood relations may be obtained for the discrete values (Freksa 1992a; Davis 2001).

Temporal interpolation can be used to fill in the gaps between observations at different times, but for extrapolating into the future or explaining the present in terms of the past we need something more powerful: a causal model of the processes giving rise to change. Causality has been an important theme in KR, and is intimately bound up with the problems of non-monotonic temporal reasoning. In the AI context, causality has been studied as a component of theories of action, where an agent needs to reason about the effects of its actions on the world, as well as those of other agents, both in order to predict what will happen and to formulate plans of action. An early formalism, which has had a far-reaching influence on subsequent developments is AI and KR, is the Situation Calculus (McCarthy and Hayes 1969); at the core of the formalism is the representation of an action-type as, in effect, a mapping from situations to situations, providing the basis both for determining the result of performing a given sequence of actions starting from a given situation (prediction) and for inferring an appropriate sequence of actions to obtain a desired state of affairs starting from a given situation (planning). A separate strand of development owes its origin to another highly influential paper (Kuipers 1984) which used 'qualitative differential equations' as the basis for a system for predicting the qualitative behaviour of mechanisms characterised by continuous time-varying parameters. The behaviours can be visualised as tracking a path through an 'envisionment' network which is reminiscent of the Conceptual Neighbourhood Diagrams for qualitative relation spaces; this connection is explored further in Galton (2001).

A third important theme is *identity*. Problems concerning identity naturally arise as a result of the increasing adoption of object-oriented approaches to data-modelling in GIS (Egenhofer and Frank 1992;



¹¹A famous expression of this is Tobler's First Law of Geography: "Everything is related to everything else, but near things are more related than distant things" (Tobler 1970). Note that, strictly speaking, the term 'autocorrelation' refers to a measure of the extent to which closer things tend to be more similar than more distant things; as used in this paragraph, it is to be understood to mean *positive* autocorrelation.

Worboys 1994). An object's identity is whatever it is that can be said to persist through all the changes that the object can undergo; this need not be anything material, but perhaps an enduring pattern of organisation or activity, or maybe simply a label. But some changes in the world can affect the identities of the objects it contains: objects can come into or go out of existence, and for some types of objects transformations such as merging or splitting can occur; for yet other objects, reincarnation may be possible, that is resumption of existence after temporary suspension (as when, say, a committee is disbanded and subsequently reinstated) (Medak 2001). Changes of this kind have been systematised under the rubric of 'identity-based change' (Hornsby and Egenhofer 1998, 2000). While this work has mainly concentrated on man-made geographical entities such as countries and other political or administrative units, some of the ideas can also be applied to natural objects, for which problems of identity can be particularly acute: consider for example the case of a lake which becomes two separate lakes as a result of a lowering of the water-table; if the water-table subsequently rises again and the lakes merge to become one, should we say that this is the original lake reincarnated or a new lake occupying the same position? Nature itself provides no answer to this, but it is a question which has to be answered for the purposes of recording the facts about the lake(s) in any information system which recognises a lake as a specific type of entity. All this is of course closely connected to the topic of ontology, mentioned earlier.

Conclusions

In this article I have concentrated on techniques for representing spatial and temporal knowledge within the KR subfield of AI and, to a lesser extent, within GIScience. In the space available it has not been possible to do more than give a brief overview of some of the main contributions to this developing enterprise. The emphasis has been on those systems which have addressed the most general conceptual issues relating to space and time; particular applications will always, of course, have to go beyond this, with domain-specific knowledge encoded alongside, or rather on top of, the general framework provided by the spatio-temporal theory. Being in this way general-purpose systems, such spatio-temporal theories are prime candidates for implementation within information systems in the drive to enable them to represent bodies of knowledge, and not just information.

References

- Allen J (1984) Towards a general theory of action and time. Artif Intell 23:123–154
- Allen JF, Hayes PJ (1985) A common-sense theory of time. In: Joshi A (ed) Proceedings of the ninth international joint conference on artificial intelligence (IJCAI'85). International joint conferences on artificial intelligence, inc. Morgan Kaufmann, San Francisco, pp 528–531
- Antoniou G (1997) Nonmonotonic reasoning. MIT, Cambridge Bennett B (1994) Spatial reasoning with propositional logics. In:

 Doyle J, Sandewall E, Torasso P (eds) Principles of knowledge representation and reasoning: proceedings of the 4th international conference (KR94). Morgan Kaufmann, San Francisco
- Bennett B (2001a) Application of supervaluation semantics to vaguely defined spatial concepts. In: Montello DR (ed) Spatial information theory: foundations of geographic information science. Proceedings of international conference COSIT 2001. Springer, Berlin Heidelberg New York, pp 108–123
- Bennett B (2001b) What is a forest? On the vagueness of certain geographic concepts. Topoi 20:189–201
- Bittner T (2002) Granularity in reference to spatio-temporal location and relations. In: Haller S, Simmons G (eds) Proceedings of the 15th Florida artificial intelligence research symposium (FLAIRS). AAAI, Menlo Park
- Bittner T, Stell JG (2001) Rough sets in approximate spatial reasoning. In: Ziarko W, Yao Y (eds) Proceedings of the second international conference on rough sets and current trends in computing (RSCTC2000). Springer, Berlin
- Bittner T, Smith B (2003) Vague reference and approximating judgments. Spat Cogn Comput 3(2,3):137–156
- Brachman RJ, Levesque HJ (2004) Knowledge representation and reasoning. Morgan Kaufmann, San Francisco
- Bratko I (2009) PROLOG programming for artificial intelligence, 4th edn. Addison-Wesley, Reading
- Brooks RA (1991) Intelligence without representation. Artif Intell 47:139–159
- Burrough P, Frank AU (eds) (1996) Geographic objects with indeterminate boundaries. Number 2 in GISData. Taylor and Francis. London
- Claramunt C, Thériault M (1995) Managing time in GIS: an event-oriented approach. In: Clifford J, Tuzhilin A (eds) Recent advances in temporal databases. Springer, Berlin, pp 23–42
- Cohn AG (1995) A hierarchical representation of qualitative shape based on connection and convexity. In: Frank AU, Kuhn W (eds) Spatial information theory: a theoretical basis for GIS. Lecture notes in computer science, vol 988. Proceedings of international conference COSIT'95. Springer, New York, pp 311–326
- Cohn AG, Gotts NM (1996a) The "egg-yolk" representation of regions with indeterminate boundaries. In: Burrough P, Frank AU (eds) Geographic objects with indeterminate boundaries, vol 2 of GISData. Taylor and Francis, London, pp 171–187
- Cohn AG, Gotts NM (1996b) Representing spatial vagueness: a mereological approach. In: Aiello LC, Doyle J, Shapiro S (eds) Principles of knowledge representation and reasoning: proceedings of the fifth international conference (KR'96). Morgan Kaufmann, San Mateo, pp 230–241
- Cohn AG, Varzi AC (1998) Connection relations in mereotopology. In: Prade H (ed) Proceedings of the thirteenth European conference on artificial intelligence (ECAI'98). Wiley, Chichester, pp 150–154



- Cohn AG, Varzi AC (1999) Modes of connection. In: Freksa C, Mark DM (eds) Spatial information theory: cognitive and computational foundations of geographic science. Proceedings of international conference COSIT'99. Springer, Berlin
- Colomb RM (2007) Ontology and the semantic web. IOS, Amsterdam
- Davis E (2001) Continuous shape transformations and metrics of regions. Fundam Inform 46:31–54
- Davis E (2008) Pouring liquids: a study in commonsense physical reasoning. Artif Intell 172:1540–1578
- Drakengren T, Jonsson P (1998) A complete clasification of tractability in Allen's algebra relative to subsets of basic relations. Artif Intell 106:205–219
- Egenhofer MJ (1989) A formal definition of binary topological relationships. In: Litwin W, Schek H-J (eds) Third international conference on foundations of data organisation and algorithms (FODO). Springer, New York, pp 457–472
- Egenhofer MJ (1991) Reasoning about binary topological relations. In: O Günther, Schek H-J (eds) Advances in spatial databases. Springer, New York, pp 143–160
- Egenhofer M (2005) Spherical topological relations. J Data Semant III:25–49
- Egenhofer MJ, Al-Taha KK (1992) Reasoning about gradual changes of topological relationships. In: Frank AU, Campari I, Formentini U (eds) Theories and methods of spatio-temporal reasoning in geographic space. Lecture notes in computer science, vol 639. Springer, Berlin, pp 196– 219
- Egenhofer MJ, Frank AU (1992) Object-oriented modeling for GIS. J Urban Reg Inf Syst Assoc 4:3–19
- Egenhofer MJ, Sharma J (1993) Topological relations between regions in \mathbb{R}^2 and \mathbb{Z}^2 . In: Abel D, Ooi BC (eds) Advances in spatial databases—third international symposium on large spatial databases SSD'93. Lecture notes in computer science, vol 692. Springer, New York, pp 316–336
- Egenhofer MJ, Clementini E, Di Felice P (1994) Topological relations between regions with holes. Int J Geogr Inf Syst 8(2):129–144
- Egenhofer MJ, Franzosa RD (1991) Point-set topological spatial relations. Int J Geogr Inf Syst 5:161–174
- Egenhofer MJ, Franzosa RD (1995) On the equivalence of topological relations. Int J Geogr Inf Syst 9(2):133–152
- Euzenat J (1995) An algebraic approach to granularity in qualitative time and space representation. In: Mellish CS (ed) Proceedings of the fourteenth international joint conference on artificial intelligence (IJCAI'95). Morgan Kaufmann, San Francisco, pp 894–900
- Euzenat J, Montanari A (2005) Time granularity. In: Fisher M, Gabbay D, Vila L (eds) Handbook of temporal reasoning in artificial intelligence. Elsevier, Amsterdam, pp 59– 118
- Fisher P, Wood J, Cheng T (2004) Where is Helvellyn? Fuzziness of multi-scale landscape morphometry. Trans Inst Br Geogr, NS 29:106–128
- Forbus KD (1995) Qualitative spatial reasoning: framework and frontiers. In: Chandrasekaran D, Glasgow J, Narayanan NH (eds) Diagrammatic reasoning: cognitive and computational perspectives. AAAI/MIT, Menlo Park, pp 183–202
- Forbus KD, Nielsen P, Faltings B (1987) Qualitative kinematics: a framework. In: McDermott J (ed) Proceedings of the 10th international joint conference on artificial intelligence (IJCAI'87). Morgan Kaufmann, Los Altos, pp 430–435
- Frank AU (1992) Qualitative spatial reasoning about distances and directions in geographic space. J Vis Lang Comput 3: 343–371

- Frank AU (1996) Qualitative spatial reasoning: cardinal directions as an example. Int J Geogr Inf Sci 10:269–290
- Freksa C (1992a) Temporal reasoning based on semi-intervals. Artif Intell 54:199–227
- Freksa C (1992b) Using orientation information for qualitative spatial reasoning. In: Frank AU, Campari I, Formentini U (eds) Theories and methods of spatio-temporal reasoning in geographic space. Lecture notes in computer science, vol 639. Springer, Berlin, pp 162–178
- Galton AP (1990) A critical examination of Allen's theory of action and time. Artif Intell 42:159–188
- Galton AP (1994) Lines of sight. In: Keane M, Cunningham P, Brady M, Byrne R (eds) AI and Cognitive science '94. Dublin University Press, Dublin, pp 103–113
- Galton AP (1996) Taking dimension seriously in qualitative spatial reasoning. In: Wahlster W (ed) Proceedings of the twelfth European conference on artificial intelligence (ECAI'96). Wiley, Chichester, pp 501–505
- Galton AP (1998) Modes of overlap. J Vis Lang Comput 9:61–79 Galton AP (1999) The mereotopology of discrete space. In: Freksa C, Mark DM (eds) Spatial information theory: cognitive and computational foundations of geographic science. Springer, New York, pp 251–266
- Galton AP (2001) Dominance diagrams: a tool for qualitative reasoning about continuous systems. Fundam Inform 46:55-70
- Galton AP (2007). Experience and history: processes and their relation to events. J Logic Comput 18(3):323–340
- Galton AP, Hood JM (2005) Anchoring: a new approach to handling indeterminate location in GIS. In: Mark DM, Cohn AG (eds) Proceedings of COSIT'05. Springer, Berlin
- Garey MR, Johnson DS (1979) Computers and intractability: a guide to the theory of NP-completeness. W. H. Freeman, New York
- Ginsberg M (1993) Essentials of artificial intelligence. Morgan Kaufmann, San Mateo
- Guarino N (1998) Formal ontology and information systems. In: Guarino N (ed) Formal ontology in information Syst. IOS, Amsterdam, pp 3–15
- Hayes PJ (1979) The naïve physics manifesto. In: Michie D (ed) Expert systems in the micro-electronic age. Edinburgh University Press, Edinburgh, pp 242–270
- Hayes PJ (1985) Naive physics I: Ontology for liquids. In: Hobbs JR, Moore RC (eds) Formal theories of the commonsense world. Ablex, Norwood, pp 71–107
- Hernández D, Clementini E, Di Felice P (1995) Qualitative distances. In: Frank AU, Kuhn W (eds) Spatial information theory: a theoretical basis for GIS. Lecture notes in computer science, vol 988. Proceedings of international conference COSIT'95. Springer, New York, pp 45–57
- Hobbs JR (1985) Granularity. In: Joshi A (ed) Proceedings of the ninth international joint conference on artificial intelligence (IJCAI'85). International joint conferences on artificial intelligence, inc. Morgan Kaufmann, San Francisco, pp 432–435
- Hobbs JR, Moore RC (eds) (1985) Formal theories of the commonsense world. Ablex, Norwood
- Hodges W (2001) Logic, 2nd ed. Penguin, Toronto
- Hood JM, Galton AP (2006) Implementing anchoring. In: Raubal M, Miller HJ, Frank AU, Goodchild MF (eds) Geographic information science: proceedings of the 4th international conference, GIScience 2006. Springer, Berlin, pp 168–185
- Hornsby K, Egenhofer MJ (1998) Identity-based change operations for composite objects. In: Poiker T, Chrisman N (eds)



- Proceedings of the 8th international symposium on spatial data handling. International Geographical Union, Freiburg, pp 202–213
- Hornsby K, Egenhofer MJ (2000) Identity-based change: a foundation for spatio-temporal knowledge representation. Int J Geogr Inf Sci 14(3):207–224
- João EM (1998) Causes and consequences of map generalisation. CRC, Boca Raton
- Krokhin A, Jeavons P, Jonsson P (2003) The tractable subalgebras of Allen's interval algebra. J ACM 50(5):591–640
- Kuipers BJ (1984) Commonsense reasoning about causality: deriving function from structure. Artif Intell 24:169–203
- Kulik L (2000) Vague spatial reasoning based on supervaluation.
 In: Winter S (ed) Geographical domain and geographical information systems—GeoInfo series, vol 19. Institute for Geoinformation, Vienna University of Technology, Vienna, pp 73–80
- Kulik L (2001) A geometric theory of vague boundaries based on supervaluation. In: Montello DR (ed) Spatial information theory: foundations of geographic information science. Proceedings of international conference COSIT'01. Springer, Berlin, pp 44–59
- Kurata Y, Egenhofer M (2006) Topological relations of arrow symbols in complex diagrams. In: Barker-Plummer D, Cox R, Swoboda N (eds) Diagrams 2006: fourth international conference on the theory and application of diagrams. Lecture notes in computer science, vol 4045. Springer, Berlin, pp 112–126
- Lehmann F, Cohn AG (1994) The egg/yolk reliability hierarchy: semantic data integration using sorts with prototypes. In: Proceedings of the conference on information and knowledge management, CIKM-94. ACM, New York, pp 272–279
- Ligozat G, Renz J (2004) What is a qualitative calculus? A general framework. In: Zhang C, Guesgen HW, Yeap W-K (eds) PRICAI 2004: trends in artificial intelligence, 8th pacific rim international conference on artificial intelligence, Auckland, New Zealand, August 9–13, 2004, Proceedings. Lecture notes in computer science, vol 3157. Springer, Berlin, pp 53–64
- Luger GF (2002) Artificial intelligence: structures and strategies for complex problem solving. Pearson, Harlow
- McCarthy J, Hayes PJ (1969) Some philosophical problems from the standpoint of artificial intelligence. In: Melzer B, Michie D (eds) Machine intelligence, vol 4. Edinburgh University Press, Edinburgh
- McDermott D (1987) Critique of pure reason. Comput Intell 3(3):151–160
- Medak D (2001) Lifestyles. In: Frank A, Raper J, Cheylan J-P (eds) Life and motion of socio-economic units. Taylor and Francis, London, pp 139–153
- Montello DR, Goodchild MF, Gottsegen J, Fohl P (2003) Where's downtown? Behavioral methods for determining referents of vague spatial queries. Spat Cogn Comput 3(2,3): 185–204
- Naur P (1989) Review of Shoham (1988). Comput Rev 30
- Nebel B, Bürckert H-J (1995) Reasoning about temporal relations: a maximal tractable subclass of Allen's interval algebra. J ACM 42(1):43–66
- Pearl J (1988) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann, San Francisco
- Peuquet DJ (1994) It's about time: a conceptual framework for the representation of temporal dynamics in geographic information systems. Ann Assoc Am Geogr 84(3): 441–461

- Randell DA, Cui Z, Cohn AG (1992) A spatial logic based on regions and connection. In: Nebel B, Rich C, Swartout W (eds) Principles of knowledge representation and reasoning: proceedings of the third international conference (KR'92). Morgan Kaufmann, San Mateo, pp 165–176
- Randell D, Witkowski M, Shanahan M (2001) From images to bodies: modelling and exploiting spatial occlusion and motion parallax. In: Nebel B (ed) Proceedings of the seventeenth international joint conference on artificial intelligence (IJCAI'01). Morgan Kaufmann, San Francisco, pp 57–63
- Renz J (2002) Qualitative spatial reasoning with topological information. Number 2293 in lecture notes in artificial intelligence. Springer, Berlin Heidelberg New York
- Renz J, Nebel B (1999) On the complexity of qualitative spatial reasoning: a maximal tractable fragment of the region connection calculus. Artif Intell 108:69–123
- Russell S, Norvig P (2003) Artificial intelligence: a modern approach, 2nd ed. Pearson Education International, Upper Saddle River
- Sandewall E, Shoham Y (1995) Non-monotonic temporal reasoning. In: Gabbay DM, Hogger CJ, Robinson JA (eds) Handbook of logic in artificial intelligence and logic programming, vol 4. Oxford University Press, Oxford, pp 439–498
- Schlieder C (1996) Qualitative shape representations. In: Burrough PA, Frank AU (eds) Geographic objects with indeterminate boundaries, vol 2 of GISdata. Taylor and Francis, London, pp 123–140
- Schmidtke HR, Woo W (2007) A size-based quantitative approach to the representation of spatial granularity. In: Veloso M (ed) Proceedings of the 20th international joint conference on artificial intelligence, pp 563–568
- Scivos A, Nebel B (2001) Double-crossing: decidability and computational complexity of a qualitative calculus for navigation. In: Montello DR (ed) Spatial information theory: foundations of geographic information science (Proceedings of COSIT 2001). Springer, Berlin, pp 421–446
- Shanahan M (1997) Solving the frame problem: a mathematical investigation of the common sense law of inertia. MIT, Cambridge
- Shoham Y (1988) Reasoning about change. MIT, Cambridge
- Smith B (2004) Ontology. In: Floridi L (ed) The blackwell guide to the philosophy of computing and information. Blackwell, Oxford, pp 155–166
- Smith B, Mark DM (2003) Do mountains exist? Towards an ontology of landforms. Environ Plann Law J Plann Des 30:411–427
- Sowa JF (2000) Knowledge representation: logical, philosophical, and computational foundations. Brooks/Cole, Pacific Grove
- Stell JG (2003) Qualitative extents for spatio-temporal granularity. Spat Cogn Comput 3:119–136
- Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. Econ Geogr 46:234–240
- van Beek P (1992) Reasoning about qualitative temporal information. Artif Intell 58:297–326
- Van Benthem JFAK (1983) The logic of time. Kluwer, Dordrecht [Second Edition (1991)]
- Vilain M, Kautz H (1986) Constraint propagation algorithms for temporal reasoning. In: Proceedings of AAAI-86. Morgan Kaufmann, San Mateo, pp 377–382
- Vögele T, Schlieder C, Visser U (2003) Intuitive modelling of place name regions for spatial information retrieval. In: Kuhn W, Worboys M, Timpf S (eds) Spatial information theory: foundations of geographic information science. Proceedings of international conference COSIT 2003. Springer, Berlin, pp 239–252



- Wang F, Hall GB (1996) Fuzzy representation of geographical boundaries in GIS. Int J Geogr Inf Syst 10(5):573–590
- Wilson NL (1959) Substances without substrata. Rev Metaphys 12:521–539
- Wolter F, Zakharyaschev M (2000) Spatio-temporal representation and reasoning based on RCC-8. In: Cohn AG, Giunchiglia F, Selman B (eds) Principles of knowledge representation and reasoning: proceedings of the seventh international conference (KR2000). Morgan Kaufmann, San Francisco, pp 3–14
- Worboys MF (1994) Object-oriented approaches to georeferenced information. Int J Geogr Inf Syst 8(4):385–399
- Worboys MF (1998) Spatio-bitemporal geographic information. In: Egenhofer MJ, Golledge RG (eds) Spatial and temporal reasoning in geographic information systems. Oxford University Press, Oxford, pp 25–39
- Worboys MF (2005) Event-oriented approaches to geographic phenomena. Int J Geogr Inf Sci 19(1):1–28
- Yuan M (2001) Representing complex geographic phenomena in GIS. Cartogr Geogr Inf Sci 28(2):83–96

