Diagrammatic transformation processes on relational maps

Christoph Schlieder

University of Freiburg, Institute of Computer Science and Social Research,
Center for Cognitive Science, Friedrichstrasse 50, 79098 Freiburg, Germany
E-mail: cs@cognition.iig.uni-freiburg.de

Abstract

Psychological evidence indicates that human reasoners solve spatial relational inferences by constructing and inspecting mental models in visuospatial working memory. From a computational point of view, this reasoning strategy seems to combine relational representations comparable to those described in the AI literature on qualitative spatial reasoning with the type of local spatial transformations that are characteristic of diagrammatic reasoning. However, applying local transformations to arbitrary relational representations involves solving the computationally intractable subgraph isomorphism problem. This paper describes a class of representations, relational maps, for which the problem becomes tractable and, as a consequence, for which diagrammatic inference is implementable by efficient local transformations.

1 Introduction

Two rather different AI research approaches on diagrammatic reasoning can be taken. For the purpose of discussion, we call them the *inferentialist* and the *computationalist perspective*. The primary interest of the inferentialists consists in giving an adequate logical description of the use of diagrammatic representations in valid reasoning. Barwise and Etchemendy (1990), the originators of this line of research, stress that "visual forms of representation can be important, not just as heuristics and pedagogical tools, but as legitimate elements of mathematical proofs." Inferentialists look for a notion of valid inference which is not tied to a specific (sentential) representational format. However, they are typically not concerned with finding valid inferences, i.e., with automating theorem proving in a hybrid reasoning system. In contrast, the computationalist approach focuses on the procedural aspects of diagrammatic inference. For computationalists, valid inference is not the only concern, since they consider the heuristic use of diagrams valuable. As a consequence, the issues of representational format and computational costs become important. An example for an analysis from this perspective is the seminal work of Larkin and Simon (1987).

This paper adopts the computationalist point of view to describe diagrammatic transformation processes on relational maps — a class of representations yet to be defined. There are two reasons why this perspective is more adequate for the task at hand. First, the paper explores an algorithmic idea that originated from our empirical investigations in the psychology of reasoning. As described in section 2, which summarizes the relevant findings, procedural aspects (premise order effects) were found to have an important influence on the reasoning result, whereas there is no evidence that human reasoners are committed to valid inference (verification bias). Second, the objective of this paper is to provide a generic and efficient computational framework to implement diagrammatic transformations. While the aspect of generality could have been handled within the inferentialist approach, efficiency is an issue only the computationalist approach deals with. Section 3 shows how to generalize the notion of local operator on a digital image to arbitrary relational data structures. However, applying the resulting abstract diagrammatic operators involves solving the computationally intractable subgraph isomorphism problem (section 4). A solution is provided by defining a sufficiently general class of representations, namely relational maps, for which efficient diagrammatic transformations exist.

2 The cognitive approach to spatial relational infernece

Spatial relational inference has been studied by cognitive psychologists for almost twenty years now. A typical spatial reasoning task is the three-term series which consists of two premises, $X\ r_1\ Y$ and $Y\ r_2\ Z$, and a conclusion $X\ r_3\ Z$ (X, Y, Z denote spatial objects, r_1, r_2, r_3 denote binary spatial relations; the conclusion has to be generated or verified). The rich body of empirical evidence currently available asks for a unifying explanatory framework. The theory of mental model reasoning proposed by Johnson-Laird and Byrne (1991) has proven to be very successful in this respect.

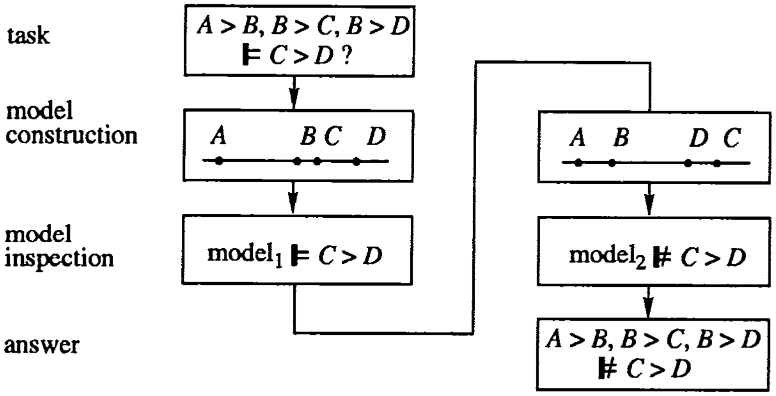


Fig. 1: Variation of mental models during spatial relational inference

Reasoning with mental models

Underlying the mental model account of inference is the general assumption of premise integration. Mental model theory assumes that the information conveyed by the premises of a reasoning task is integrated into a unified representation in working memory, the mental model. Formally, mental models play a role in reasoning that is comparable to models in logic: they represent structures in which the premises are valid (under an appropriate interpretation). According to Johnson-Laird and Byrne (1991) spatial relational inference is a three stage process consisting of a *model construction* phase which integrates the premises into a model, a *model inspection* phase which produces a conclusion valid in that model, and, finally, an optional search for alternative models (*model variation*). Fig. 1 illustrates this process for a simple inference about a linear ordering. Two mental models are constructed before it is found that the premises do not entail the conclusion. Unfortunately, there is very little empirical evidence about model variation. It remains a major research topic to find how many and which models reasoners construct.

Evidence for model preference

Experiments with reasoning tasks constructed from the system of interval relations introduced by Allen (1983) produced two essential findings (Knauff, Rauh & Schlieder, 1995; Rauh & Schlieder, in print). First, the *existence of preferred mental models*, and second, the *presence of order effects*. We will discuss model preference using the task illustrated in Fig. 2. Subjects are asked to generate a conclusion which is consistent with the premises. For this specific task there are three correct answers, each corresponding to a different model. From a logical point of view, none of the models of an Allen three-term series is more preferable than the others. However, it was found that subjects agree considerably in their choice of a model, which we therefore call preferred model.

We can learn about model construction by looking at the general dependency between three-term series tasks due to symmetry. The relevant group of symmetry transformations is generated by two transformations, reorientation and transposition. It turns out that the number of violations of transposition symmetry is much higher than the number of violations of reorientation symmetry. This is called the order effect since violation of transposition symmetry means that the result of processing a premise depends upon which other premise has previously been processed.

A overlaps *B* from the left
B overlaps *C* from the left
How could *A* lie with respect to *C*?

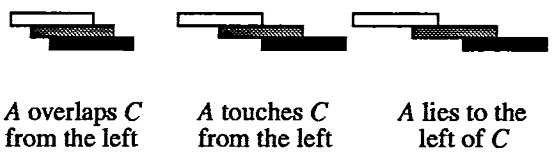


Fig. 2: Mental models of the premises of an Allen three-term series

A computer simulation has been described by Schlieder (1996) that accounts for both, model preference and order effect. Although the computer simulation represents the mental model by a relational data structure rather than by a digital image, it behaves diagrammatically in several respects. The interval relations are represented implicitly in terms of the linear ordering of the interval's startpoint and endpoint, which means that the mental model has to be inspected to retrieve the relations. It is possible to implement both model construction and model inspection by means of local computation steps (shifts of a spatial focus). The same would hold for the transformation that generate alternative models during the model variation phase if empirical evidence indicated that those transformations are restricted to neighboring elements in the point ordering. Schlieder (1996) introduced the term *abstract imagery* form this type of spatially local processing of relational representations.

3 Implicit representation of spatial relations

We will now study abstract imagery, i.e., diagrammatic reasoning with relational representations, from an algorithmic perspective. It turns out that outside the one-dimensional world of interval relations, the local transformations used in diagrammatic inferences are computationally very expensive. Finding a way to resolve this complexity issue is a prerequisite to applying abstract imagery to engineering problems. However, it does not imply that the solution also reflects cognitive processes involved in spatial relational inference. Note that our primary objective is to use an intuition about cognition to provide a generic and efficient computational framework for abstract imagery. cognitive adequacy is not claimed (nor is it denied).

Relational description

A syntactically simple type of relational description, which is nevertheless sufficiently expressive for many applications of qualitative spatial reasoning (QSR), represents the position of points in the plane by means of binary spatial relations. We restrict our analysis to this type of description. Two kinds of symbols are needed to formulate a description, point symbols $p_1, \dots, p_n$ and binary relations symbols $r_1, \dots, r_m$. Using infix notation we write a relational expression as $p_i\ r_j\ p_k$. A relational description is simply a finite set of relational expressions.

Many examples of binary spatial relations between points can be found in the QSR literature. Probably the most common among the relations encoding ordinal information are the cardinal directions (e.g. p_i *north-of* p_k, see Frank, 1991, and others). Their geometrical meaning is determined by two lines through p_i delimiting a sector within which p_k has to lie such that the relation holds. An example for relations encoding metrical information is the system of qualitative distances (e.g. p_i *close-to* p_k) described by Hernández, Clementini, and Di Felice (1995). It uses concentric circles around p_i to delimit the region where p_k must be located for a certain qualitative distance relation.

The examples illustrate a standard way of specifying the geometrical meaning of spatial relations that is adopted by many QSR approaches. Rather than axiomatizing the relations, one describes their interpretation in some intended structure. In the examples, the intended structures are arbitrary configurations of points in the Euclidean plane. Sometimes it is useful to consider more restricted point sets, e.g., only points on a triangular, rectangular, or hexagonal grid. Even more complex structures may be needed. To define a relation such as p_i *visible-from* p_k some additional point sets are distinguished in the plane and serve as opaque obstacles. Similarly complex structures are needed to interpret the incidence relations which we will use as our running example. Intuitively, these relations describe the order in which the points are incident with a line.

Let us consider a concrete case, relational descriptions with point symbols $p_1 \dots, p_{14}$ and incidence relation symbols $r_1, \dots, r_4$ (see Fig. 3). The structure for interpreting these relational descriptions is constructed as follows. Four lines $l_1, \dots, l_4$ are arranged in such a manner that no three lines have a point in common. The 6 intersection points partition the lines into segments. On each of the 8 unbounded segments an arbitrary point is chosen. The resulting configuration of 14 points has 6 + 8 = 14 points. Every configuration of 14 points that is related to an arrangement of 4 lines in the way described above is considered an intended structure.

A description is given an interpretation by specifying a one-to-one mapping $\mathfrak{I}$ of the point symbols onto points of an intended structure. The relational expression $p_i\ r_j\ p_k$ is true under the interpretation iff $\mathfrak{I}(p_i)$ and $\mathfrak{I}(p_k)$ are neighboring points on the line, such as D and F on l_1. Once intended structures are specified, a notion of entailment restricted to these structures may be defined. Although this way of proceeding is close to how designers of relational representations think of their semantics, it should be mentioned that the use of intended structures makes an analysis of logical properties of relational descriptions rather difficult.

Diagrammatic inference procedures

A fundamental idea of diagrammatic reasoning consists in using the topology of the spatial representation (diagram) to guide the flow of control during the reasoning process and thus to avoid the combinatorial

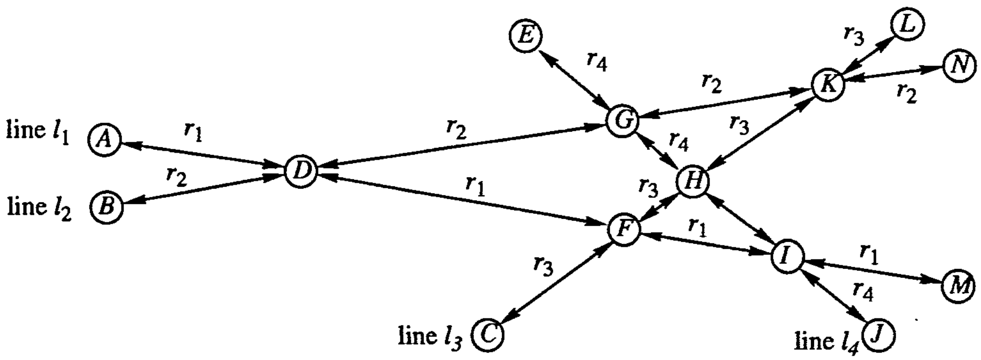


Fig. 3: Incidence relations induced by an arrangement of four lines

explosion from which classical inference procedures suffer. This intuition is already present in Sloman (1975) — one of the first papers published on the subject — where it is claimed that a spatial "indexing scheme" provides "ways of controlling the order in which assertions or inference steps are tried." In the meantime, different models of local computation have been proposed to formalize this intuition, for example the graphical search and replace operators of Furnas (1991) and the activation spreading in depictions of Habel (1987), Khekhar (1990) and Pribbenow (1992). A common feature of all these diagrammatic inference procedures is that in each computation step they process only information from a spatial neighborhood (see Fig. 4).

The diagrammatic transformations illustrated in Fig. 4 are intended to simulate very simple physical behavior: the movement of a set of objects falling from and gliding along obstacles. Furnas (1991) showed that such transformations of digital images can be implemented by graphical search and replace operations. Computational costs are generally not an issue with such transformations. Of course, there is the question of whether the operator allows for parallel application; but the step of applying the operator once, i.e., the graphical search and replace process, does not constitute a computationally hard problem. The search part dominates the replace part. For a sequential operator (which is the worst case) defined by a mask of size m and an image of size $n > m$ it takes a maximum of $m \times n = \theta(n^2)$ pixel comparisons to decide where the operator is applicable in the image.

Local operators on relational descriptions

As a first step towards defining local transformations of relational descriptions, let us show how graphical search and replace operations on a digital image can be interpreted as transformations of a specific type of relational description. The positional information of the pixels of the image in Fig. 5 is easily encoded by a system of four relations $r_1, \dots, r_4$, with the following

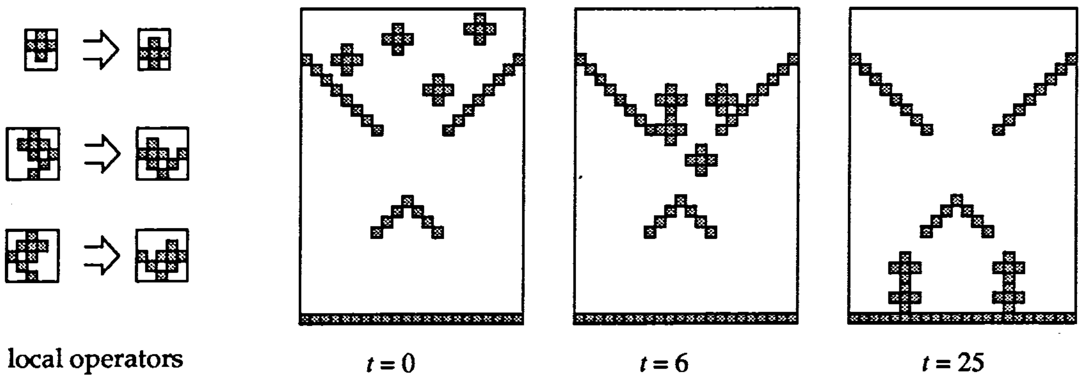


Fig. 4: Local operators acting on a digital image (adapted from Furnas, 1992)

semantics: r_1 holds between 4-adjacent pixels of the figure (grey pixels); r_4 holds between 4-adjacent pixels of the background (white pixels); r_2 holds between a background pixel and a 4-adjacent figure pixel; r_3 holds between a figure pixel and a 4-adjacent background pixel.

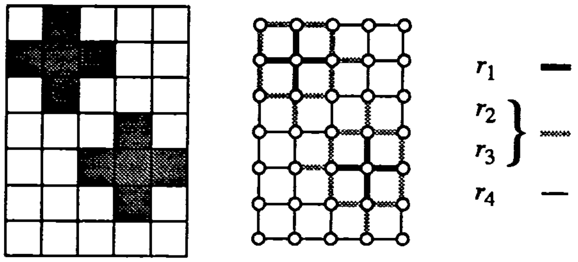


Fig. 5: Relational description of a digital image

Note that any relational description can be seen as a directed edge-colored graph (one color for each relation). Therefore, digital images correspond to a specific class of 4-connected graphs. The graphical search and replace operators which are specified by a left-hand side image and a right-hand side image correspond to *subgraph substitution rules*.

4 Relational description and their local transformations

The graph theoretical formulation of the diagrammatic operators on digital images lends itself to generalization. On arbitrary relational descriptions one could consider implementing local operators by subgraph substitution rules. Whereas the graphical search involved in the application of a local image operator can be efficiently implemented by shifting the corresponding left-hand side mask once over the image, the situation changes considerably if we consider operators on arbitrary relational descriptions. No efficient algorithm is known to exist for the search part, that is, for finding a subgraph of the relational description which is isomorphic to the left-hand side of the substitution rule. Furthermore, it is very improbable that an efficient algorithm will be found since the subgraph isomorphism problem is known to be NP-complete (e.g. Garey & Johnson, 1979).

Subgraph isomorphism for embedded graphs

Usually, computationally intractable problems are approached either by approximating the solution or by modifying the problem slightly in order to obtain a tractable problem. We will follow the latter route. But first we need to summarize some relevant algorithmic results. The graph isomorphism problem (not to be confused with the subgraph isomorphism problem) has so far resisted all attempts to prove its NP-completeness. Nevertheless, no polynomial algorithm has been found. It is believed to belong to a class of problems of intermediate complexity (Köbler, Schöning & Toran, 1993). However, it has been known for a long time that the graph isomorphism problem is solvable in polynomial time for planar graphs (Hopcraft & Wong, 1974).

A graph is embedded into the plane by mapping each of its vertices onto a point. Edges are mapped onto straight line segments joining points which are images of adjacent vertices. We call this a *geometrical embedding*. A graph is *planar* iff it has a geometrical embedding without crossing edges. Planarity is a purely combinatorial property of a graph which can be determined very efficiently in time proportional to the number of vertices (different linear time planarity testing algorithms are discussed in Nishiseki and Chiba, 1988)

Planar graphs possess more structure than general graphs, a fact that can be exploited to devise efficient algorithms. However, the subgraph isomorphism problem does not become tractable when restricted to planar graphs — it is even NP-complete if the graph is a tree and the subgraph a forest! An algorithmic intuition related to planarity proves nevertheless valuable: many of the efficient algorithms rely on the fact that an important subclass, namely the triply connected planar graphs, possess an unique combinatorial embedding in the sphere. A combinatorial embedding of a graph is specified by the circular ordering in which the edges appear around vertices in a geometrical embedding. The combinatorial embedding abstracts from the metrical information contained in the geometrical embedding retaining only ordinal information.

If graph and subgraph are given together with a combinatorial embedding, then the subgraph isomorphism problem (relative to the embedding) becomes tractable. Fig. 6 illustrates why this is the case. Note that for both, the graph Γ and the (sub)graph Σ, not only the incidences between vertices and edges, but also the counterclockwise ordering of the edges around the vertices is known. The combinatorial embeddings put strong constraints on the mapping of edges of Σ onto edges of Γ. If for example the edge xu of Σ is tentatively mapped onto the edge HK of Γ, then each edge incident with x has to be mapped onto the edge incident with H that has the corresponding position in the circular edge ordering: xy onto HI, xz onto HF, xr onto HG. The process can be iterated yielding a mapping of all edges of Σ if a structure preserving mapping exists (see appendix).

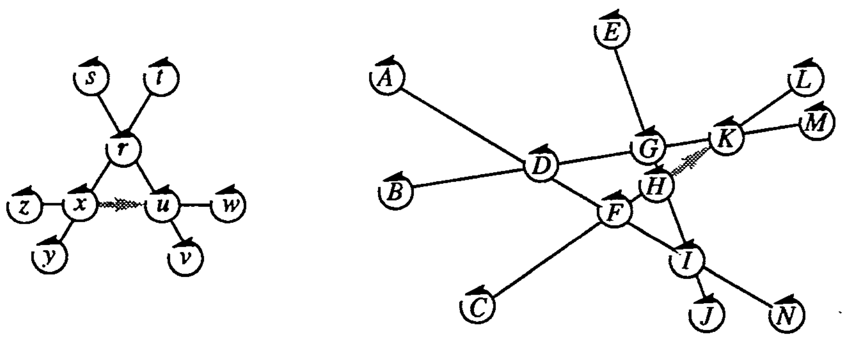


Fig. 6: Using the combinatorial embedding to find a subgraph isomorphic to Σ in Γ

In other words, any mapping of a single edge of Σ onto an edge of Γ can be extended to a complete mapping of all edges without backtracking over edge assignments — a simple implementation of this procedure by a depth-first traversal of the subgraph is described in the appendix. The NP-completeness of the subgraph isomorphism problem suggests that backtracking cannot be avoided if Σ and Γ are arbitrary graphs.

Reasoning with relational maps

We are now in a position to define efficient diagrammatic reasoning operators on a specific type of relational descriptions. A *relational map* is a relational description together with a combinatorial embedding. Subgraph substitution rules whose left-hand side and right-hand side are specified by relational maps are called *diagrammatic map transformations*. Looking for examples of diagrammatic map transformations, we first note that all graphical search and replace operators on digital images satisfy the definition, because the grid has a canonical combinatorial embedding in the plane. An example for a diagrammatic map transformation of a more relational character is found among the algorithmic paradigms studied in computational geometry: the movement of a sweep line over an arrangement of lines (see Fig. 7). Sweeping a simple (= no three lines have a common point) arrangement of four lines can be implemented by the diagrammatic map transformation shown. The geometrical reason for this is that the sweep line will always form a triangle with the two lines whose intersection point it is going to pass next. At each step of the process the relational map is searched for subgraphs to which the map transformation is applicable. Judging from the example, it seems that diagrammatic map transformations are sufficiently expressive to describe quite complex geometrical operations. Since they encompass local operators on digital images, we may consider them interesting candidates form formulating diagrammatic inferences.

Appendix: Extending a vertex mapping

Σ and Γ are relatinal maps, e_Σ is an edge of Σ and e_Γ an edge of Γ. The procedure *match* extends a given vertex mapping I from the vertices of Σ to the vertices of Γ in such a way that the induced edge mapping maps e_Σ onto e_Γ. Initially, the procedure is called with empty I. For convenience, directed edges are used in the description of the procedure. A directed edge e is incident with two vertices, *origin*(e) and *destination*(e). Redirecting the edge e, i.e., exchanging origin and destination, yields the edge *converse*(e).

Before the procedure starts, the statuts of all vertices and edges of Σ is *unvisited*. The procedure returns a partial mapping of the vertices of Σ onto the vertices of Γ. If this vertex mapping is not complete (i.e. not defined for all vertices of Σ) then no isomorphism mapping e_Σ onto e_Γ exists. Otherwise, the the isomorphism is specified by the vertex mapping returned.

match(e_Σ, e_Γ, I)

v_Σ := destination(e_Σ);
v_Γ := destination(e_Γ);

Distinguish three cases according to the status of v_Σ and the status of converse(e_Σ):

v_Σ unvisited:
    e_Σ' := successor(e_Σ,v_Σ);
    e_Γ' := successor(e_Γ,v_Γ)
    add (v_Σ,v_Γ) to the vertex mapping I;

v_Σ visited and converse(e_Σ) unvisited:
    e_Σ' := converse(e_Σ);
    e_Γ' := converse(e_Γ);

=
v_Σ visited and converse(e_Σ) visited:
    if successor(e_Σ,v_Σ) visited
    then return vertex mapping *I*;
    e_Σ' := successor(e_Σ,v_Σ);
    e_Γ' := successor(e_Γ,v_Γ)

mark e_Σ and, if necessary v_Σ as visited;
return match(e_Σ', e_Γ', *I*)

References

Allen, J. (1983). Maintaining knowledge about temporal intervals. *Comm. of the ACM, 26,* 832-843.

Barwise, J., & Etchementy, J. (1990). Visual information and valid reasoning. In W. Zimmerman (ed.) *Visualization in Mathematics* (pp. 9-24). Washington, DC: Mathematical Association of America.

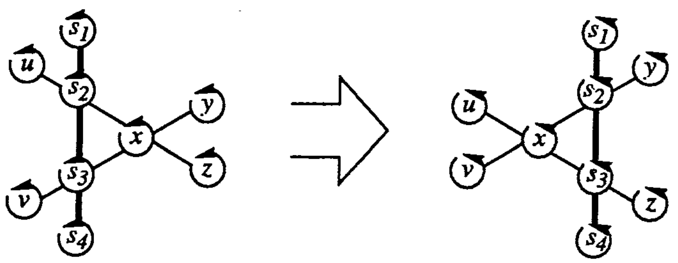
Frank, A. (1991). Qualitative reasoning with cardinal directions. In *Proc. Austrian Conference on Artificial Intelligence* (pp. 157-167). Berlin: Springer.

Furnas, G. (1991). New graphical reasoning models for understanding graphical interfaces. In *Human Factors in computing systems, Proc. CHI'91* (pp. 71-78).

Garey, M., & Johnson, D. (1979). *Computers and intractability.* New York: Freeman.

Habel, C. (1987). Cognitive linguistics: The processing of spatial concepts. *T.A. Informations - ATALA, 28,* 21-56.

Hernández, D., Clementini, E., & Di Felice, P. (1995). Qualitative distances. In Proc. Conference on Spatial information theory (pp. 45-57). Berlin: Springer.



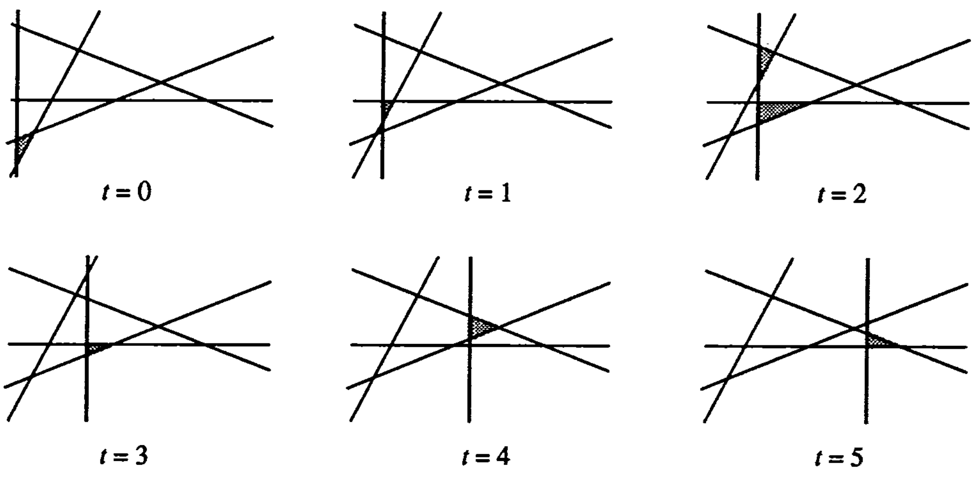


Fig. 6: Moving a sweep line ober an arrangement of lines

Hopcraft, J. & Wong, J. (1974). A linear time algorithm for the isomorphism of planar graphs. *Proc. STOC '74* (pp. 172-184).

Johnson-Laird, P. & Byrne, R. (1991). *Deduction*. Hillsdale, NJ: Lawrence Erlbaum.

Khenkhar, M. (1990). Object-oriented representation of depictions on the basis of cell matrices. In O. Heryog & C. Rollinger (eds.), *Text understanding in LILOG*. Berlin: Springer.

Knauff, M., Rauh, R. & Schlieder, C. (1995). Preferred mental models in qualitative spatial reasoning: A cognitive assessment of Allen's calculus. In J. Moore & J. Lehman (Eds.) *Proc. COGSCI-95* (pp. 200-205), Mahwah, NJ: Lawrence Erlbaum.

Köbler, J., Schöning, U., & Toran, J. (1993). *The graph isomorphism problem: Its structural complexity*. Basel: Birkhäuser.

Larkin, J., & Simon, H. (1987). *Why a diagram is (sometimes) worth ten thousend words*. Cognitive Science, 11, 65-99.

Nishizeki, T., & Chiba, N. (1988). *Planar graphs: Theory and algorithms*. Amsterdam: North-Holland.

Rauh, R., & Schlieder, C. (in print). Symmetries of model construction in spatial relational inference. To appear in *Proc. COGSCI-97*.

Pribbenow, S. (1992). Computing the meaning of localizing expressions involving prepositions. C. Zelinsky-Wibbelt (ed.) *The semantics of prepositions in natural language processing*. Berlin: De Gruyter.

Schlieder, C. (1996). Diagrammatic reasoning about Allen's interval relations. In P. Olivier & al. (eds.) Working notes of the *AAAI Spring Symposium on cognitive and computational models of spatial representation* (pp. 83-91). Stanford, CA: AAAI.

Sloman, A. (1975). Afterthoughts on analogical representations. In *Proc. Theoretical issues in natural language processing* (pp. 164-168). Cambridge, MA.