To appear in:
Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics.
Chicago, 18-21 October 1992.

On the Utilization of Spatial Structures for Cognitively Plausible and Efficient Reasoning

Christian Freksa and Kai Zimmermann

Fachbereich Informatik, Universität Hamburg
Bodenstedtstr. 16, 2000 Hamburg 50, Germany
freksa@informatik.uni-hamburg.de
zimmerma@informatik.uni-hamburg.de

Abstract **We present an approach to qualitative spatial reasoning which is based on directional orientation information as available through perceptual processes. Qualitative orientations in 2-dimensional space are given by the relation between a point and a vector. The paper presents our basic iconic notation for spatial orientation relations which exploits the spatial structure of the domain and explores a variety of ways in which these relations can be manipulated and combined for spatial reasoning.**

I. INTRODUCTION

Our knowledge about physical space differs from all other knowledge in a very significant way: we can perceive space directly through various channels conveying distinct modalities. Unlike in the case of other perceivable domains, spatial knowledge obtained through one channel can be verified or refuted through the other channels. As a consequence, we are disproportionally confident about what we know about space: we take for granted that it is as we perceive it.

Our research on spatial representations and reasoning is motivated by the intuition that 'dealing with space' should be viewed as cognitively more fundamental than abstract reasoning. Afterall, one of the very first tasks we learn to accomplish is to orient ourselves in the environment. The use of spatial metaphors in language and problem solving tasks also indicates that there might be a specialised, maybe less expressive, but optimized, spatial inference mechanism. Why else would we translate a problem into the specialised domain of space if it could be handled by a general inference mechanism? As a consequence, we want to understand dedicated spatial reasoning before we construct general abstract reasoning engines. The goal of this research is the conception of a 'spatial inference engine' which deals with spatial knowledge in a way more similar to biological systems than systems based on abstract logic languages.

Manuscript received July 24, 1992.

Spatial orientation information, or more specifically, directional information about the environment, is directly available to animals and human beings through perception, and is crucial for establishing spatial location and for path finding. Such information typically is imprecise, partial, and subjective. In order to deal with this kind of spatial information we need methods for adequately representing and processing the knowledge involved. In this paper we present an approach to representing and processing qualitative orientation information which is motivated by cognitive considerations about the knowledge acquisition process.

Consider a simple localization task: you walk straight along a road, turn to the right, walk straight, turn left, and walk straight again. Now you would like to know where you are located with respect to the first road you walked on. Tasks like this are very fundamental for almost all animals and human beings. We mostly carry them out subconsciously – except when we fear getting lost in underground walkways. In the following we describe how we represent this knowledge for modelling spatial reasoning.

II. THE REPRESENTATION STRUCTURE

We introduce an *orientation grid* for representing qualitative orientation information. The grid is aligned to the orientation determined by two points in space, the start point and the end point of a movement, e.g. the points *a* and *b*. First we can distinguish three different possible positions of an additional point *c* with respect to (w.r.t.) the imagined line through *a* and *b*. Point *c* might be left of the line, on the line or right of the line. We call these relations *left*, *straight*, and *right*. But since we consider *a* and *b* to be the start point and the end point of a movement, respectively, we can not only refer to the line through *a* and *b*, dividing the whole plane, but to the vector ***ab*** denoting the oriented path from *a* to *b*.

With this in mind we can impose new qualitative distinctions in the relative position of point *c* w.r.t. the vector ***ab***, called *front*, *neutral* and *back*, each related to one of the lines through the end points *a* and *b* orthogonal to ***ab***, where front means the direction to which the vector points and back its inverse. Fig. 1a shows the lines that form the orientation grid and Fig. 1b and 1c show the iconic representation we use to denote the 15 possible positions and orientations of *c* that can be distinguished by means of the grid.

Although people and most animals do not have a perception system for explicit front/back or forward/backward discrimination as they do for left/right discrimination, the segmentation of the plane into a front and a back semi-plane also is meaningful from a cognitive point of view: we conceptualize people, animals, robots, houses, etc. as having an 'intrinsic front side' (compare Pribbenow [1990], Mukerjee & Joe [1990]); this results in an implicit dichotomy between a front region and a back region and a forward and backward orientation.

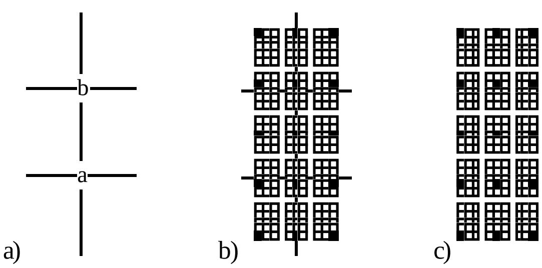


Fig.1: a) The basic position and orientation grid; b) the iconic representation of point *c* w.r.t. vector ***ab*** superimposed on the grid; c) relation of point *c* w.r.t. vector ***ab*** without the basic grid.

Comparison to Existing Approaches

A variety of approaches to qualitative spatial reasoning has been proposed. Güsgen [1989] adapted Allen's [1983] qualitative temporal reasoning approach to the spatial domain by aggregating multiple dimensions into a Cartesian framework. Güsgen's approach is straightforward but it fails to adequately capture the spatial interrelationships between the individual coordinates. The approach has a severe limitation: only rectangular objects aligned with their Cartesian reference frame can be represented in this scheme. Since we only represent the relative position and orientation information of points as abstractions of real object positions we do not make any assumptions on the shape of the objects and we are not restricted to one specific rectangular coordinate system that has to be applied to all objects.

Randell [1991] attacks the problem of representing qualitative relationships of concave objects. He introduces a 'cling film' function for generating convex hulls of concave objects; he then lists all qualitatively different relations between an object containing at most one concavity and a convex object. Egenhofer & Franzosa [1991] develop a formal approach to describing spatial relations between point sets in terms of the intersections of their boundaries and interiors. They do not use orientation information.

Hernández [1992] considers 2-dimensional projections of 3-dimensional spatial scenes. He attempts to overcome some deficiencies of Güsgen's approach by introducing 'projection' and 'orientation' relations. For the dimension of projection he adopts and extends the ideas of Egenhofer [1989], i.e. the binary topological relationship between two areas in the plane. But he combines the topological information with relative orientation information that can be defined on multiple levels of granularity. Nevertheless, he is still describing scenes within a static reference system. Freksa [1991] suggests a perception-based approach to qualitative spatial reasoning; a major goal of this approach is to find a natural and efficient way for dealing with incomplete and fuzzy knowledge.

Schlieder [1990] develops an approach which is not based on the relation between extended objects or connected point sets. Schlieder investigates the properties of projections from 2-D to 1-D and specifies the requirements for qualitatively reconstructing the 2-dimensional scene from a set of projections yielding partial arrangement information.

Frank [1991] discusses the use of orientation grids ('cardinal directions') for spatial reasoning. The investigated approaches yield approximate results, but the degree of precision is not easily controlled. Mukerjee & Joe [1990] present a truly qualitative approach to higher-dimensional spatial reasoning about oriented objects. Orientation and rectangular extension of the objects are used to define their reference frames.

Our representation also allows us to describe orientation and position qualitatively, but it does not deal with the shapes of objects and is not restricted to objects of certain shapes. Furthermore the operations on our representation do not yield approximate values but correct ranges of values. One of the major differences to previous approaches is that the relative positions of other objects are not described w.r.t. to one position but w.r.t. to a vector that describes the movement between two positions.

Our approach is motivated by cognitive considerations about the availability of spatial information through perception processes [Freksa 1991]. A major goal of this approach is to find a natural and efficient way for dealing with incomplete and fuzzy knowledge. The operations applicable on this kind of representation are described below.

III. THE OPERATIONS

So far we have only introduced the new representation scheme for position and orientation information. Let us now consider the operations that we can apply to spatial information. First we will discuss the operations working on a single relation, i.e. the relative position of one point w.r.t. to one vector, and then the operation of combining two or more vectors: composition.

Inversion

The first operation we will consider is the inversion (INV) of the vector ***ab***. This corresponds to the task of turning back and asking the question "What would the spatial orientation of *c* be if I were to walk back from *b* to *a*?" Thus from the knowledge about the relation of *c* w.r.t. the vector ***ab***, i.e. *c:**ab***, we deduce the relation of *c* w.r.t. the inverted vector ***ba***, i.e. *c:**ba***.

The inversion operation yields a precise result. All of the following operations do not; rather they sometimes yield a disjunction of possible relations. The precision of the solution is a result of the fact that when applying the inversion operation *left* simply becomes *right*, *front* becomes *back*, and vice versa. Unlike in the case of linear dimensions, incrementing quantitative orientation leads back to previous orientations, i.e. INV(INV(*z:**xy***))=*z:**xy***. In this sense, orientation is a circular dimension.

Although the above statement seems well known and trivial, existing approaches do not deal with periodicity of orientation explicitly. Periodicity is either eliminated by not admitting certain orientations, as in [Schätz 1990] or it is ignored by treating different orientations as independent entities, as in Frank [1991]. In the example in section IV we will show that the operation of inversion is essential in deducing the resulting disjunction of relations in an inference task at its maximum precision. Fig. 2 shows the resulting relations of the inversion operation.

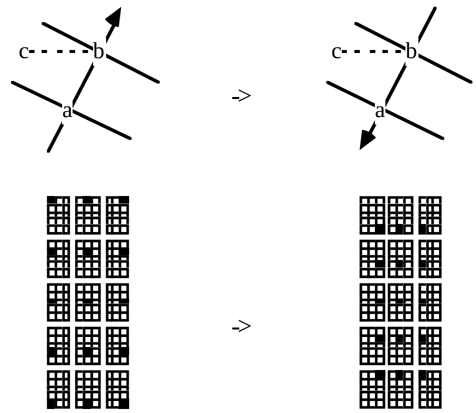
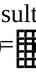


Fig. 2: The inversion operation and the resulting relations in their corresponding positions, e.g. Inv()=.

Homing

Another way to deduce new relations between a point and a vector is *homing* (HM). In this case we are interested in answering the question: "Where is the start point *a* if I continue my way by walking from *b* to *c*?" Formally this deduction step can be expressed as finding the relation of *a* w.r.t. ***bc***, i.e. *a:**bc***, from the given relation *c:**ab***. Fig. 3 shows this graphically.

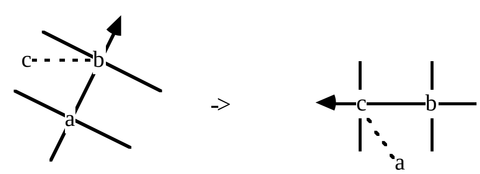


Fig. 3: Orienting the past point *a* to ***bc*** when continuing the path from *b* to *c*.

Unlike the inversion operation the operation of homing does not yield precise results in the sense that only a single possible relation always results from a single base relation. Instead for three input relations the application of HM results in a disjunction of possible relations. Fig. 4 shows the results.

Fig. 4: The left table shows the resulting relations from the HM operation. The table on the right depicts its inverse, HMI, i.e. INV(HM(*z*:***xy***))

Notice the rectangle, column two, row two. HM() = , i.e., the result is absolutely undetermined; each spatial relation is a valid candidate for a solution. This is because if you are walking from *a* to *b* and then back to *b*, maybe with some other steps in between, the reference vector for point *a* becomes ***bb***, i.e. the null vector.

But on the other hand, the undetermined relation establishes continuity of the operator w.r.t. the resulting topological relations. As we can easily see if we move from one relation to its direct neighbor, only neighboring relations are added or removed from the resulting set. This behavior of conceptually neighboring relations has been studied by Freksa [1992a] for the domain of time and for the domain of space, using the above introduced representation [Freksa 1992b]. In Hernández & Zimmermann [1992] it is shown how these neighboring structures can be applied to and combined with default reasoning methods.

Shortcut

The third possibility for recombining the points within one given reference frame is called *shortcut* (SC). This is the deduction of the relation of *b* w.r.t. ***ac***, from the given relation *c*:***ab***, see Fig. 5. This allows us to solve two very common problems in spatial reasoning. First we may deduce the relations between objects if we are really interested in finding a shortcut. On the other hand, shortcut allows us to specify the relative position of objects that are not on our path but to the side of it. Thus, we may relate observed objects on the side of our path, e.g. *b*, to the direct path from *a* to *c* and use this kind of information in the later reasoning process, using only one combination operator.

One can see the strong inner resemblance between the homing and the shortcut operations, for which only one table is necessary, because all the relations found in the HM table are found in the SC table, only in a different order. Fig. 6 shows the correspondence between the two operations.

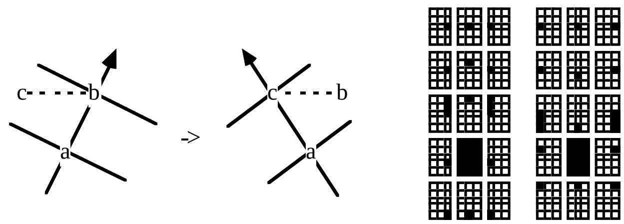


Fig. 5: The shortcut operation and the resulting sets of possible relations for SC, left table, and its inverse SCI(*z*:***xy***) = INV(SC(*z*:***xy***)), right table.

HM				SC		
1	6	11		15	10	5
2	7	12		14	9	4
3	8	13	->	13	8	3
4	9	14		12	7	2
5	10	15		11	6	1

Fig. 6: Correspondence between the operations HM and SC.

Algebraic Combination of Operations

Fig. 7 shows how the operations can be combined algebraically. This kind of combination is not commutative, but it is associative. The associativity allows us, for example, to apply a general and possibly parallel constraint propagation algorithm in which the temporal order of combination does not matter. If the combination were not associative, we would be restricted to an ordered computation, e.g., backward chaining.

o	ID	INV	SC	SCI	HM	HMI
ID	ID	INV	SC	SCI	HM	HMI
INV	INV	ID	HM	HMI	SC	SCI
SC	SC	SCI	ID	INV	HMI	HM
SCI	SCI	SC	HMI	HM	ID	INV
HM	HM	HMI	INV	ID	SCI	SC
HMI	HMI	HM	SCI	SC	INV	ID

Fig. 7: The algebraic combination of operations. The result of HM(SC(*z*:***xy***))⊇HMI(*z*:***xy***) can be found in the fourth row, column six.

Composition

If we are given two segments of a path, we can combine them. If, for example, we have *c*:***ab*** and *d*:***bc***, we can deduce *d*:***ab***, i.e., we can try to relate later stations of a path to earlier segments. This operation does not, in general, produce unique results, e.g., combining '*c* is in the right front of ***ab***' when '*d* is in the right front of ***b***c', results in a disjunction indicating that *d* can be anywhere to the right of ***ab***, see Fig. 8.

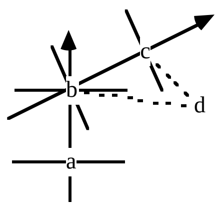


Fig. 8:
Composing *c* ***ab*** and *d* ***bc*** leads to the result *d* ***ab***.

The table showing the results of combining the 15x15=225 possible pairs can be found in Freksa [1992b] accompanied by an extensive discussion of the resulting neighborhood structure of the chosen representation. This discussion will not be repeated here. There exists a proof that with the three basic operations INV, HM, and SC, plus composition, we can build and deduce a relation for every possible combination of points w.r.t. to any vector. The proof is not presented here, because we are currently investigating whether the result achieved is at its maximum sharpness, i.e., whether the approach is complete in the mathematical sense. This has yet to be proved.

IV. EXAMPLE

Up to now only the representation and the four fundamental operations that may be performed on it have been discussed. We will now show that these four operations can be combined successfully to answer questions about the orientation of arbitrary points w.r.t. any line in 2D space.

Since it is known that landmarks play a very important role in human orientation and path finding [Lynch 1960], we consider as cognitively fundamental the task of orienting the new position not only w.r.t. places reached by the path, but also with respect to landmarks off the route. It is a very common phenomenon that if a landmark was visible from one point in the path and then becomes obscured, e.g., by a building, that we try to locate the landmark again. Most of the time we have a very clear intuition about the general direction in which we have to search for it.

We will now show that it is not necessary to define new composition tables for such tasks. The solution to this kind of question can be deduced algebraically in a goal-oriented backward chaining system or through a simple constraint propagation mechanism which would correspond to an unoriented forward chaining task.

Consider the following route description, depicted in Fig. 9: ”Walk down the road (***ab***). You will see a church (*c*) in front of you on the left. Before you reach the church turn down the path that leads forward to the right (***bd***).“ The question one might ask on his way down the path is ”Where is the church w.r.t. me?“. More formally we can state this as: Given *c*:***ab*** and *d*:***ab***, deduce *c*:***bd***.

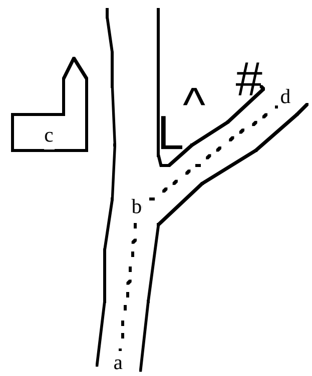


Fig. 9: A church left to the path, hidden by some trees.

In Fig. 10 we see the result obtained by directly deducing *c*:***bd*** through composition. As can easily be seen from the figure, this is not the result we expected; it is obvious that the result should be, ”The church is somewhere to my left.“ This more appropriate result may be obtained if we do not ask for the direct composition that leads to *c*:***bd***, but if we ask for the composition that gives us the inverted relation *c*:***db*** instead (see Fig. 11). Thus, in an implementation we have to ensure that each possible way to deduce a smaller disjunction of relations is explored to produce a result at its optimum.

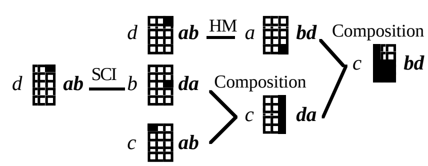


Fig. 10: Asking for the composition of *c*:***bd*** directly only yields: ”The church is somewhere to my left or behind me.“

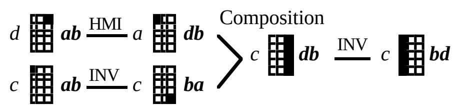


Fig. 11: Asking for the composition of the inverse of *c:**bd*** instead leads to the expected improved result: "The church is somewhere to my left."

V. OPEN QUESTIONS

Is this inference system complete? Up to now we have only proven that we can deduce a relation between an arbitrary point w.r.t. to any line. We are currently trying to prove that these deduced relations are at their maximum sharpness.

What is the overall complexity of the complete inference mechanism? We can prove that the number of possible relations of points and lines increases monotonically with $O(k^3)$, where k is the number of points. In general we would expect the runtime of a constraint propagation system that makes explicit all the implicit relations to be of exponential order. But for the domain of time interval logic there exists a proof, see Vilain, Kautz & van Beek [1989], that if only convex sets of relations are fed into the system, the algorithm is of a polynomial order. Informally speaking this means only conceptually neighboring relations without 'holes' should be specified for the orientation.

Does there exist a preferred way of carrying out the computation? We have seen that the result of a deduction depends on the way the result is derived and which relations are used. We are currently investigating whether there exists a preferred way, which would speed up the inference system, because we would have to take into account only the preferred subset of possible combinations.

ACKNOWLEDGMENT

Most of the presented ideas have been discussed in the context of the 'Space Inference Engine' project at the University of Hamburg with Marco Homann, Carsten Wiegand, Antje Wulf, and Rolf Sander. Their valuable criticism and comments helped improve the approach considerably. We thank Clayton McMillan for proofreading the manuscript.

REFERENCES

Allen, J.F., Maintaining knowledge about temporal intervals, *CACM* 26 (11) (1983) 832-843.

Chang, S.K., Jungert, E.: A spatial knowledge structure for image information systems using symbolic projections. *Proc. of the National Computer Conference*, Dallas, Texas, November 2-6, 1986, 79-86.

Egenhofer, M.: A formal definition of binary topological relationships. In W. Litwin and H.-J. Schek, editors, *Third International Conference on Foundations of Data Organization and Algorithms*, vol. 367 of Lecture Notes in Computer Science, 457-472. Springer, Berlin 1989.

Egenhofer, M., Franzosa, R.: Point-set topological spatial relations. *Intern.. Journal of Geographical Information Systems,* vol.5, n.2, 161-174, 1991.

Frank, A.U.: Qualitative spatial reasoning with cardinal directions. *Proc. Seventh Austrian Conference on Artificial Intelligence*, Wien, 157-167, Springer, Berlin 1991.

Freksa, C.: Qualitative spatial reasoning. In: D.M. Mark & A.U. Frank (eds.) *Cognitive and Linguistic Aspects of Geographic Space*, Kluwer, Dordrecht 1991.

Freksa, C.: Temporal reasoning based on semi-intervals. *Artificial Intelligence* 54 (1992a) 199-227.

Freksa, C.: Using orientation information for qualitative spatial reasoning. *Proc. of the International Conference GIS - From Space to Territory. Theories and Methods of Spatio-Temporal Reasoning*. Pisa (Italy), September 21-23. 1992. Springer, New York, 1992b. In press.

Furbach, U., Dirlich, G., Freksa, C.: Towards a theory of knowledge representation systems. In: W. Bibel & B. Petkoff (eds.) *Artificial Intelligence Methodology, Systems, Applications* (pp. 77-84). Amsterdam: North-Holland 1985.

Güsgen, H.W.: *Spatial reasoning based on Allen's temporal logic*. TR-89-049, International Computer Science Institute, Berkeley 1989.

Hernández, D.: Diagrammatical aspects of qualitative representations of space. *Proceedings of the AAAI Spring Symposium on Reasoning with Diagrammatic Representations*. 225-228. Stanford 1992.

Hernández, D., K. Zimmermann: *Default reasoning and the qualitative representation of knowledge.* Report, Technische Universität München.

Lee, S.-Y., Hsu, F.-J., Picture algebra for spatial reasoning of iconic images represented in 2D C-string. *Pattern Recognition Letters* 12 (1991) 425-435.

Lynch, K.: *The Image of the city.* The MIT Press & Harvard University Press, Cambridge/Mass. 1960.

Mukerjee, A., Joe, G.: A qualitative model for space, *Proc. AAAI-90*, 721-727.

Pribbenow, S.: Interaktion von propositionalen und bildhaften Repräsentationen. In: C. Freksa & C. Habel (eds.) *Repräsentation und Verarbeitung räumlichen Wissens*, 156-174, Springer, Berlin 1990.

Randell, D.: Analysing the familiar: a logical representation of space and time. *3rd International Workshop Semantics of Time, Space, and Movement,* Toulouse 1991.

Schätz, B.: Ein konnektionistischer Ansatz zur verteilten Darstellung von Objektpositionen. In: W. Hoeppner (ed.) *Workshop 'Räumliche Alltagsumgebungen des Menschen',* 147-158, Koblenz 1990.

Schlieder, C.: Anordnung. Eine Fallstudie zur Semantik bildhafter Repräsentation. In: C. Freksa & C. Habel (eds.) *Repräsentation und Verarbeitung räumlichen Wissens*, 129-142, Springer, Berlin 1990.

Vilain, M., H. Kautz, P.G. van Beek: Constraint propagation algorithms for temporal reasoning: a revised report. In D.S. Weld and J. de Kleer, eds. *Readings in qualitative reasoning about physical systems,* 373-381, Morgan Kaufmann, San Mateo, CA, 1989.

Zimmermann, K.: A proposal for representing object sizes. *Proc. of the workshop 'Spatial concepts: Connecting cognitive theories with formal representations'* at the ECAI 1992, Wien.