A qualitative model for space

Amitabha Mukerjee
Gene Joe

Department of Computer Science
Texas A&M University
College Station TX 77843-3112

Abstract

Most geometric models are quantitative, making it difficult to abstract the underlying spatial information needed for tasks such as planning, learning or vision. Furthermore, the precision used in a typical quantitative system often exceeds the actual accuracy of the data.

In this work we describe a systematic representation that builds spatial maps based on local qualitative relations between objects. It derives relations that are more "functionally relevant" - i.e. those that involve accidental alignments, or can be described based on such alignments. In one dimension, interval logic [Allen 83] provides a mechanism for representing these type of relations; in this work we propose a formalism that enables us to perform alignment-based reasoning in two and higher dimensions with objects at angles. The principal advantages of this representation is that

a) it is free of subjective bias, and

b) it is complete in the qualitative sense of distinguishing all overlap/ tangency/no-contact geometries.

In addition, the model is capable of handling uncertainty in the initial system (e.g. "the fuse box is somewhere behind the compressor") by constructing bounded inferences from disjunctive input data. Two kinds of uncertainty can be handled - those arising from deliberate imprecision in the interest of compactness ("down the road from"), or those caused by an inadequacy of data (sensors, spatial descriptions, or maps).

Keywords: Spatial reasoning, qualitative vision, path planning, natural language.

1. INTRODUCTION

In this paper, we present a systematic method for extracting meaningful symbolic descriptions from geometric data. For example, for the map in figure 1, consider the query "How do I get from the Post office to the Civic Auditorium?". While conventional geometric modeling systems can represent this map with great precision, the problem of reasoning requires a measure of abstraction which is difficult to obtain from these large databases of geometric coordinates. In this paper we propose a abstraction for spatial relations based on the hypothesis that relations involving tangency are "more important" than others in the categorization of spatial relations, and we show how the representation can be used to successfully generate many of the predicates that have been used in modeling space.

Very often, one assumes that the problem of spatial abstraction can be solved somehow, and the input to the model is a list of carefully chosen symbolic descriptors. The difficulty of this problem has long been known [McCarthy 77,epistemological problem 4]:

> "A robot must be able to express knowledge about space, and the locations, shapes and layouts of objects in space. Present programs treat only very special cases. Usually locations are discrete - block A may be on block B but the formalisms do not allow anything to be said about where block on B it is, and what shape space is left on block B...A formalism capable of representing the geometric information that people get from seeing and handling objects has not, to my knowledge, been approached."

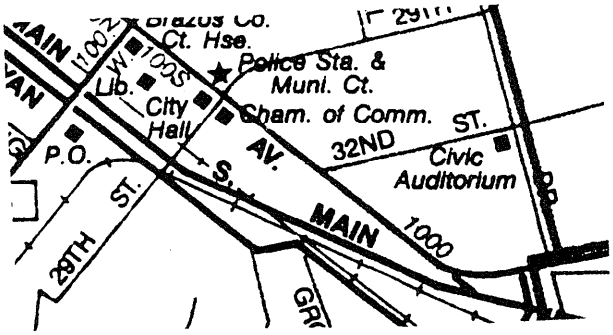


Figure 1. A map of Bryan, TX.

Ten years later, the basic issue addressed in this excerpt still remains a fundamental problem. In practice, the predicates chosen to describe a particular task already contains considerable personal bias and insight. A crucial problem that one faces when one formulates spatial predicates is that one loses the underlying geometric information that relate these spatial predicates together. For example, consider the predicates ON(A,B) and LEFT-OF(A,B) in figure.2; the underlying geometric similarity is revealed when we change our viewpoint, yet there is no way of preserving this relationship at the symbolic level. Furthermore the spatial semantics of ON(A,B) has varied widely in AI literature depending on the implementation. One of the benefits of our representation is that it provides a language for defining relations such as ON(A,B) with one's own particular semantics, while preserving the underlying geometric realities.

ON(A,B)   LEFT-OF(A,B)

Figure 2: If we look at the LEFT-OF figure from the right-hand side of the page, it is clear that the underlying geometric information in both configurations is the same.

1.1 What is functionally important?

This paper is concerned mostly with spatial representation itself, rather than with any particular application. A new and powerful representation is developed that can be used to describe and infer geometric relations under conditions of complete or partial information. The principal thrust is to describe positions relative to other objects, as opposed to descriptions in terms of global coordinates. Such descriptions arise naturally in human spatial reasoning, and many cognitive models are based on some particular orientation of the viewer or in relation to other objects (e.g. "behind the tree") [Dennett 75].

precision/detail ←	→ flexibility/abstraction
QUANTITATIVE MODELS	ABSTRACT MODELS
[Ambler & Popplestone 76]	[Winston 75]
[Requicha 80]	[Kuipers & Byun 88]
	[Peuquet &Ci-Xiang 87]
	[Retz-Schmidt 88]

Figure 3. Given two objects, their position and orientation with respect to each other can be described either in terms of a series of numbers (coordinates), or in terms of some less quantitative measure (e.g. "cross the Safeway store and veer right at the fork"). The two descriptions involve a tradeoff between precision/detail and flexibility/ abstraction.

The formal representation developed here can describe one, two, and three-dimensional models, involving orthogonal or angular relations. In the past, spatial relations have been developed with specific domains in mind. A large body of work has dealt with the orthogonal domains in the blocks world (e.g. [Winston 75]). Geographic data systems need to be able to answer spatial queries [Peuquet & Ci-Xiang 87]. Natural language generators and interpreters need to be able to reason about spatial prepositions [Retz-Schmidt 88]. In robot path planning, a number of qualitative systems have been developed, such as sequence of view-frames [Kuipers & Byun 88], and polyhedral approximations [Ernest 86]. In all these instances, the carefully selected domain-specific vocabulary is likely to result in brittle behavior when exposed to broader problems.

Our approach differs from these prior attempts in that we obtain the set of spatial primitives based on a categorization of all possible configurations that are qualitatively distinct; this is the sense in which it is *complete*. The concepts that we focus on involve tangency - Are the given objects are aligned at some face, line, or point? If not, then where is one object with respect to the other? At the same time, there are spatially relevant concepts such as "near" and "far" which we cannot model, since these are both in the same qualitative category (no-contact). Figure 4 shows some examples to illustrate the importance of this kind of reasoning in human thinking.

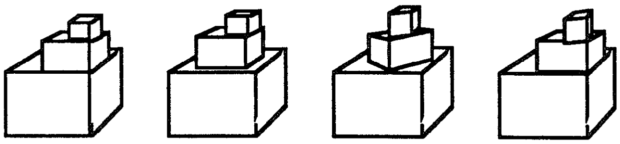


Figure 4a: These arrangements of blocks are distinguished by the accidental alignments between face, edge, and vertices. Usually, such alignments indicate a functional aspect: either the alignment was artificially created, or it is the result of some physical process in accordance with physical laws. For example, along the z-axis, all the objects shown have a face tangency relation, indicating the physical reality of support. In addition to alignments, the other issue that is immediately apparent is that of size - in all three images, the blocks are graded in size.

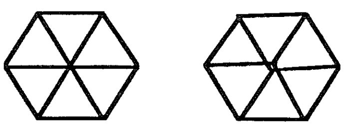


Figure 4b: A number of psychological tests bear witness to the fact that the human cognitive process emphasizes accidental alignments. A well known example is that of the hexagonal cube, where the accidental alignment precludes us from seeing the three-dimensional shape.

One of the more useful aspects of this representation is the ability to model uncertainties in the description of the model - situations where spatial knowledge is not precise enough to quantify through coordinates. Two kinds of uncertainty can be handled - those arising from deliberate imprecision in the interest of compactness ("down the road from the laundry"), or those caused by the inadequacy of data (resolution or storage limits in sensors, spatial descriptions, or maps). In most of these instances, bounded inferences can be made depending on the uncertainty involved.

Another motivation behind a formal theory of spatial relations is to provide an objective set of spatial primitives, so that one can remove the influence associated with selecting the predicates used in a spatial inference system (e.g., in learning arches, "touches" is more important than "overlaps"). At the same time, a formal representation, by providing a known domain over which it is complete, unburdens the designer of many of the problems involved in ensuring that his/her vocabulary is powerful enough to describe all the possible descriptions that can arise. This is particularly appropriate for acquiring new concepts, where the completeness of the vocabulary and capability of hierarchical abstraction become important.

In order to define a systematic approach towards this abstract yet comprehensive representation, we begin by identifying the qualitatively different aspects of the relative positions of two objects in space, i.e. the basic vocabulary of binary spatial relations.

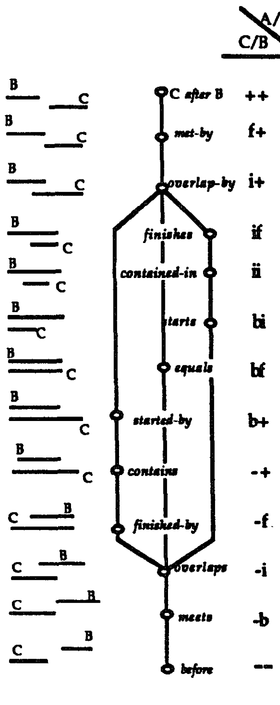
2. ONE-DIMENSIONAL RELATIONS : INTERVAL LOGIC

Let us first consider objects along a single dimension. This case has been investigated in depth in the study of time. We consider here only one of the formalisms proposed, the interval logic model [Allen 83] and [Malik & Binford 83] in which relations are typically defined only between locally related events.

The description predicates are chosen by a simple, comprehensive process. To start with, we realize that objects in one-dimension can be either points or intervals. Also, we assume that the points are ordered along some direction, which may be due to a physical fact (e.g. time), due to some abstract notion (e.g.positive numbers), or some object feature (e.g.direction of motion). There are three possible cases for relations between two objects A and B:

i) *Both A and B are points.* In this case there can be only three relations - A can be behind B (-), it can be the same point as B (=), or it is ahead of B (+).
ii) *A is a point and B is an interval.* In this case, there are five qualitatively distinct cases: A is behind B (-), A is at the same point as the back of B (b), A is inside B (i), A is at the front of B (f), and A is ahead of B (+).
iii) *Both A and B are intervals.* This is the most interesting case, and has been dealt with in some detail in the study of temporal events [Allen 83]. In general thirteen relations are possible. These are discussed below.

In the two-interval case, if we consider an endpoint of an interval C then this point can be before, inside or after B. In addition, there are two more cases of interest - coincidence with either the front or the back boundary of B, which constitute the *tangency* cases. Altogether there are five regions of interest for this endpoint of C: +, f, i, b, - (ahead, front, interior, back, and posterior respectively). Each endpoint of C may be at one of these five qualitative positions, subject to the constraint that the front endpoint (or *head*) must be ahead of the rear endpoint (*tail*), leading to a total of thirteen (5+3+3+1+1) positions for C with respect to B. The relation between two intervals C and B can then be expressed for example, as C (++) B, which would mean that C is *after* B. These relations are shown along the left hand margin of figure 5.



C/B \ A/B	+	f	i	b	-
++	?	-	-	-	-
f+	>	b	-	-	-
i+	>	i	<	-	-
if	+	f	<	-	-
ii	+	+	?	-	-
bi	+	+	>	b	-
bf	+	f	i	b	-
b+	>	i	i	b	-
-+	>	i	i	i	<
-f	+	f	i	i	<
-i	+	+	>	i	<
-b	+	+	+	b	<
--	+	+	+	+	?

Figure 5. One-dimensional interval relations and the transitive inference table. The diagrams on the left show the relations between the intervals C and B as C moves leftwards from "after" B to "before" B. The graph next to it shows the progression of relations during the movement; the three branches in the continuum represent the cases where C is longer than, equal to, or shorter than B. The table on the right shows the relationship of an endpoint of A with respect to the interval C if the relations of both A and C are known with respect to another interval B. The symbol ? denotes that the relation A/C can be any of the five possible relations, > implies that A/C can be one of {i,f,+}, and < implies that the relation is in {-,b,i}.

Given the local relations between "neighboring" intervals (A/B) and (B/C), the transitive relation (A/C) is often disjunctive. The inference relations shown in figure 5 are more compact than Allen's: 5x13 instead of 13x13 since they exploit the independence between the two ends of an interval. For example if A was an interval and we knew that A was *overlap-by* B {A/B = i+} and that C *starts* B, {C/B = bi} then we can conclude based on the transitivity table that with respect to C, the rear end of A is > or {i,f,+} and the front end of A is +. Therefore the relation A/C is either i+ (*overlap-by*), f+ (*met-by*), or ++ (*after*). This establishes a constraint in the possible positions of A with respect to C.

This decomposition also depicts a continuum that exists between relations, shown in the elongated vertical graph in the figure. C is initially ahead of B (++), and as it moves towards B, it intersects B, and can have relations along one of three branches, e.g., if C is smaller than B, then only the relations {if, ii, bi} are possible. This notion allows us to compare and represent the relative size of objects, which is an important qualitative distinction. Thus if we define a ***flush-translation operator*** ϕ for moving A until it is flush with B, then by observing the relation between ϕA and B one can determine whether A is longer, equal or shorter than B. Another benefit of the continuum concept is that it lets us define hierarchical relations, which can be used, e.g., as a measure for inductive bias in learning (see section 6).

3. MULTI-DIMENSIONAL SPACES : ORTHOGONAL DOMAINS

One-dimensional interval logic can be easily extended to multi-dimensional cases where each object is either oriented with the axes, or is enclosed in a box which is so oriented (if the axes are orthogonal, then the box becomes a cuboid). Here we represent the relations along each of N axes as one element in a N-dimensional relation. For objects that are not rectangular, one can associate a "front" direction, which can be used to determine the enclosing rectangle.

In this logic, the "atomic element" is the triple (X,Y,reln), where X and Y are objects with finite extent in each dimension and reln is a n-vector each element of which is in the set of one-dimensional relations outlined above. The "disjunctive element" is the triple (X,Y,complex-reln), where complex-reln is an n-vector where each element is a disjunction from the same relation-set. The interpretation for this syntax is fairly straightforward. Figure 6 shows an example of a transitive inference.

Orthogonality arises quite "naturally" in human thought (e.g. left, right, front, back, east, north). One problem is that each object often has a different "natural" orthogonal system, so that no one representation can model all of them. Another significant problem is that for non-aligned objects, the rectangular enclosures often overlap when the actual objects are disjoint.

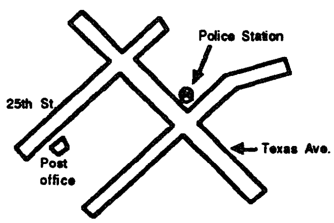


Figure 6. Given the two-dimensional qualitative relations between the post office (A) and Texas Avenue (B) and Texas Avenue and the Police Station (C), we can transitively obtain the disjunctive relation (B,C):

$$(A/B) = \begin{pmatrix} -- \\ ii \end{pmatrix} \wedge (C/B) = \begin{pmatrix} ++ \\ ii \end{pmatrix} \Rightarrow (A/C) = \begin{pmatrix} -- \\ ?? \end{pmatrix}$$

Given only its position with respect to Texas Avenue, one cannot determine the y-relation of the Police Station w.r.t. the post office. Thus without knowing which side of 25th St the police station is, one would not be able to docide which direction to turn into Texas Avenue. However, adding the 25th St information results in the output (A/C) = $\begin{pmatrix} - - \\ - - \end{pmatrix}$, which contains sufficient information to make this decision..

Nonetheless, the orthogonal representation system is actually quite powerful. By adding operators into the system, many qualitatively interesting questions can be answered. For example, let us consider a reflection operator ρ which reflects the object about a +45 degree line through its bottom left corner. Now, by comparing ρA with A, one can answer a query of the form "Is it a square?". Earlier we introduced a flush-translation operator; by repeating translations equal to an object's own dimensions, we can define a "integer-multiple" vector operator λ; e.g. $\begin{pmatrix}2\lambda\\3\lambda\end{pmatrix}A$ would imply a rectangle twice as large as A in x and thrice as large in y, and with the same bottom left corner. This can be used, along with the flush-translate operator described earlier to compare the relative sizes of two objects: "A is three times as large as B in x". We can also build a simple model for shape. For example, we can compare A with $\rho\begin{pmatrix}2\lambda\\\lambda\end{pmatrix}A$ to establish if "the aspect ratio for A is greater than 2 or less than 1". It is clear that even concepts such as near and far can be actually represented in terms of this extended system.

Finally, the axes of reference for these objects need not be perpendicular at all, indeed, they need not be straight even. Thus, lines of latitude and longitude, sectors on a polar plot, patches defined by parametric functions on a family of curves, can all be represented by the orthogonal model, since in each case, there are two clearly independent axes along which the interval relations can be used.

4. OBJECTS AT ARBITRARY ANGLES

When objects are not oriented orthogonally, relationships become more complex. One of the primary problems is that the relationships are no longer meaningful with respect to some absolute coordinate frame but must be expressed in terms of one object or the other. The road veers to the left; it does not "make an angle of 25.77 degrees to the 59th parallel of longitude".

We begin by constructing a comprehensive mapping of the relations between two objects at an angle, in the qualitative contact sense. Since angular relations of B w.r.t A are dependent on A's direction, which is generally not related to B's direction, the operators in this formalism are non-commutative and do not have well-defined inverses, i.e., given the position (A/B), the (B/A) position cannot be determined.

Let us consider the relations with respect to a single object A, which has a designated front. This defines four angular quadrants with respect to A, and the "front" for the other object may be oriented in any of these quadrants.

The spatial relation between two objects at an angle has two attributes:

a. the relative direction (internal angle), and
b. the relative positioning.

In this representation, we have tried to determine a naturally arising set of attributes for capturing the qualitative description for both of these attributes.

4.1. Representing direction

When moving from the orthogonal domain into the angular domain, a mechanism is needed for the representation of the angular information. In our approach, we base all angular relations from a predefined "special" direction, called the "front" (figure 7).

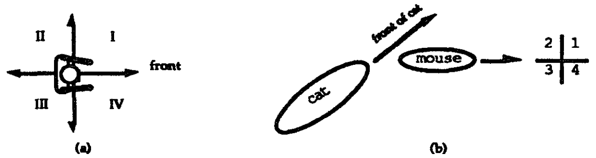


Figure 7. *Qualitative directions.* Based on a special direction or a "front" direction, we can define four principal directions (front,left,right,back) as well as four quadrants. The cat, for example, is behind the mouse and facing in the mouse's first quadrant.

Most objects have a special direction or a front. A car, chair, cat, person, house, etc. all have a "front" direction. Where an object is symmetrical and has no "special" direction, one can assign a front to it. This "front" direction is crucial to our modelling of orientation, since it provides a reference direction for all other objects and their "fronts." For example, the church *faces* the library-signifies that their fronts are at 180°; the cat pounced on the mouse from the ***behind*** and to the ***right*** implies that the front of the cat was in the mouse's first quadrant (figure 7(b)).

In two dimensions, the assignment of a single "front" direction enables us to immediately identify several qualitatively different zones. First of all, "front" defines the orthogonal directions "left," "right," and "back." In addition, it defines four possible quadrants (e.g. quadrant I is from front to left), thus defining eight qualitative angular relations between any two directions. In three-dimensions, one would need to define a second "special" direction, one that could perhaps be called "up." This defines twenty-six qualitative regions based on an octant decomposition.

In figure 7(b) above, the cat is pointing in the mouse's quadrant I. This information can be written as *dir*(Cat/Mouse) = I where the *dir()* relation is a representation of the direction information. Note that *dir()* is invertible, i.e. given dir(A/B), dir(B/A) is uniquely defined.

4.2. Representing size/position

Another consideration of spatial relations is the relative sizes and positions of objects. One method for modeling this may be to construct an enclosing box around the objects based on the "front" direction, and then extend the boundaries of this box to create 8 regions and 8 boundaries as in figure 8(a). The extensions of the lines in the forward direction are called the "lines of travel" as shown in figure 8(b). The representation shown in (a) is somewhat more powerful than that in (b), but is significantly more complex, and does not add any extra information regarding the actual intersection of objects. Furthermore, representation (b) collapses into the orthogonal representation when the two directions are orthogonal, or the

Figure 8 The extension of the sides parallel to the front direction forms the two lines of travel.

linear case when the two directions are parallel. In this research therefore we have adopted the representation based on the lines of travel and ignored the width lines. This representation is much more compact yet preserves information with respect to the front direction of the object.

As in the linear case, we consider each object as defined by the two endlines instead of the interval; eventually this leads to a smaller transitivity table.

4.3. Positional Relations: The Collision Parallelogram

The support lines of two objects, when intersecting at an angle, will form a parallelogram, called the collision parallelogram (CP), which defines the area that is common to the lines of travel of both B and A.

As an endline of A slides along its line of travel from "behind" the CP to "after" the CP, it passes through the "inside" of the CP and also two important qualitative points-"back" of the CP and the "front" of the CP (figure 9(a)).

Altogether there are three pieces of information needed to completely describe the relationship of two objects (A and B):
1)the quadrant information of dir(A/B),
2)where object A is located with respect to object B: pos(A/B), and
3)where object B is located with respect to object A: pos(B/A).

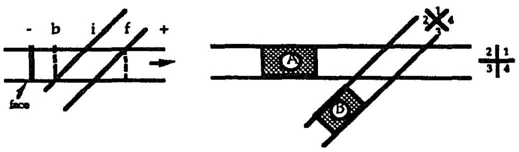


Figure 9. *The Position Relation.* The endline in (a) is located in the "–" region. The relations for two such endlines define the position relation for an object. In (b), the relation between A and B is defined as dir(A/B) = IV, pos(A/B) = —, and pos(B/A) = —. Note that pos(B/A) is computed based entirely on the collision parallelogram, and is independent of pos (A/B); therefore it is necessary to maintain both pieces of information.

Note that once the quadrant information of A/B is known then the quadrant information of B/A can be derived. To derive the relationship between two objects, each endline is considered one at a time. Finding the relationship between two objects is similar to one-dimensional intervals. The face of an object is labeled depending on which region it is located in with respect to the parallelogram formed by the travel lines of another object.

Another aspect of this representation is the interrelations between the direction relation dir() and the position relation pos(). In particular, the front directions of the objects are key to the position relations obtained - changing A's front direction can affect the pos(A/B) as well as the dir() relation. However, note that pos(B/A) is not affected by changing A's direction since the CP remains the same. In the next section we exploit the interrelations between the dir() and pos() predicates to reduce the size of the transitivity tables.

5. TRANSITIVE RELATIONS

If the relations between (A,B) and (B,C) are known, what type of information can be inferred about the relationship between objects A and C? Figure 10 shows an example of such a transitive operation. Let us now investigate the nature of the transitive table for arbitrarily angular objects, which is seen to be significantly larger and more complex than the representation of the one-dimensional transitive table of figure 5.

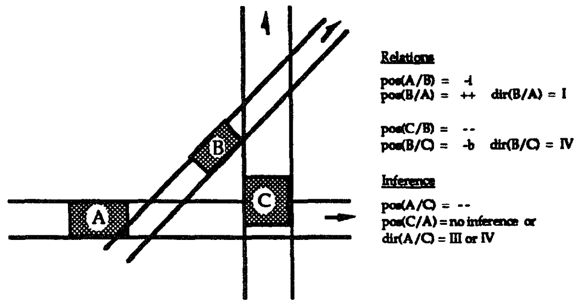
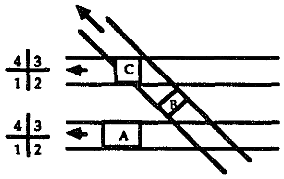


Figure 10. *Transitivity.* An example of transitive inference given the relation between (B,A) and (B,C), we infer that A must be - - with respect to C. The position of C w.r.t A is uncertain since C's position w.r.t B "- -" does not constrain it in anyway w.r.t. the lines of travel of A. In many human contexts also, such decisions are difficult to make in the absence of local information between A and C.

The transitive information relates known information about the spatial relations (A,B) and (C,B) to make inferences about (A,C). Again, by using only the endline information, we can achieve a 5/13 savings. Figure 11 shows a sample entry in the tables. Each quadrant group (e.g. dir(B/A) = I, dir(B/C) = I) contains 13x13 such tables.

In this formalism, there is a 90^{o} uncertainty in each angular relation, so when we perform a transitive operation, the uncertainty in the output is 180^{o} or two quadrants. This is seen in the chart of figure 11, where two quadrants rows are shown for each of A/C and C/A.



B/C – I	quad	+	f	b	i	-
A/B=>A/C	I	+	+	>	>	?
	IV	?	-	-	-	-
C/B=>C/A	I	?	<	<	-	-
	IV	+	+	+	+	?

Figure 11. An entry from the transitive table. This represents the case where pos(B/A) = ++, dir(B/A) = I, pos(B/C) = --, dir(B/C) = I. The A/B and A/C positions are indicated by their endline position in the top row of the table.

To use this table five parameters have to be known: 1) pos(B/A), 2) dir(B/A), 3) pos(B/C), 4) dir(B/C), and 5)the relationship of one endline of A or C with respect to B. Note that the resulting relations inferred from the transitive table may be disjunctive:

e.g. <u>if the input tuples are:</u>
B / A = <I, ++,i+>
B / C = <I,--,bf>
<u>what is inferred from the table is:</u>
A/C = I: -? = <--,-b,-i,-f,-+> 5
IV: >+ = <i+,f+,++> 3
C/A= I: -< = <--,-b,-i> 3
IV: ++ = <++> 1

which is actually a disjunction of (5*1) + (3*3) = 14 possible relations between A and C (since when dir(A/C) = I, dir(C/A) = IV). The number of disjunctions can be reduced if additional constraints are placed on A and C, e.g. via fourth object D.

5.1. Size of the transitive table

Each relation between two objects has 676 possible results (13x13x4). Therefore, for two objects, there are 676 x 676 =

456,976 entries in the table. However, the number of entries in the table can be reduced by realizing that the quadrant information and the position information are interrelated - for example, if the direction of any object is reversed, the position changes in a certain manner. Similarly certain properties are preserved when we consider configurations that are related by reflection. Another point to note is that pos(A/C) is independent of pos(C/A) and vice versa. In particular, pos(A/B) affects only pos(A/C) while pos(C/B) affects only pos(C/A). This permits the transitive table to partition C/A and A/C into two spaces.

Formally, the set of quadrants in the transitive tables form two groups under the operations of reflection and direction inversion defined below. The arrangement of the groupings illustrates the structure of the relationship that exists between the groupings. For example, the group dir(B/A) = I and dir(B/C) = I, when reflected, results in dir(B/A) = III and dir(B/C) = III. Thus by determining the effects of the reflection operator, it becomes sufficient to maintain only one of these two quadrant groups.

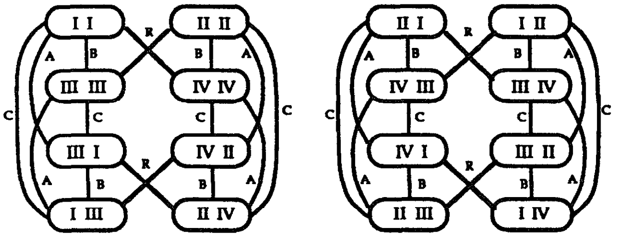
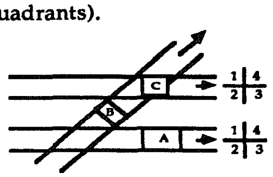


Figure 12. *Quadrant interrelations.* Each node represents the transitivity tables for two quadrants for B/A and B/C. The arcs between the nodes represent the operations that map one quadrant group into another. (R = reflection; A,B,C = direction change on A, B and C). As a result of these interrelations it is sufficient to maintain the tables for only two quadrants, say (I,I) and (II,I).

First let us consider the operation of reflection. For example, if we consider the table in figure 13, then under reflection (looking from behind the page), the configuration is the same as that in figure 11. Clearly the transitive tables are not independent; in this instance the position relations are essentially the same (since reflection does not affect the pos() relations), but the rows are interchanged (it does affect the quadrants).



B/C – I	quad	+	f	b	i	-
A/B=>A/C	I	?	-	-	-	-
	IV	+	+	>	>	?
C/B=>C/A	I	+	+	+	+	?
	IV	?	<	<	-	-

Figure 13. *Reflected version of the table in Figure 11.* The positional information remains unchanged, which is reflected in the fact that rows 1 and 3 of this table is the same as rows 2 and 4 of Figure 11. The quadrants are switched due to reflection.

Another operation that reduces table size is directional inversion. When an object reverses direction, the physical location of the object does not change. In a transitive relationship involving objects A, B, and C, any one of the objects can perform a directional change. The directional change of two objects (e.g. A and B) will result in the same quadrant group as that obtained by changing the direction of the third object alone (i.e. C). Thus, directional changes form a cycle of operations in the group.

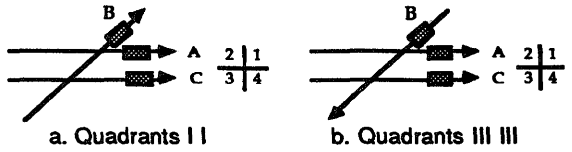


Figure 14. *Direction change.* When the direction of B is reversed, four relations are affected: dir(B/A) and dir(B/C), pos(B/A) and pos(B/C). The pos() relations are reversed, e.g. "- -" becomes "++". The changes in dir() result in going from one quadrant group to another, (e.g. from I I to III III).

Let us consider a transitive inference involving A, B, and C. When the direction of B is changed to its opposite direction, the physical locations of objects A, B, and C remain exactly the same with respect to each other. The only things that change are the quadrant relations of B/A and B/C and B's positional relations with respect to A and C. Since there has been a quadrant change for B/A and B/C, the result is a movement from one quadrant group to another. In this case, from quadrant group I I to III III. This type of quadrant group interrelation allows us to store only two quadrant tables, a saving of 2:16, or one-eighth.

This representation has also been extended to three dimensions, where a nominal frame can be attached to each object based on the "front" and "up" directions mentioned earlier. Rectangular enclosures do not represent sufficient feature information, and a generalized cylinder model is used to represent shape, with the cross-section being modeled with a qualitative version of the medial-axis transform. This model has been used to construct qualitative geometric models for visualization and recognition tasks [King and Mukerjee 90].

6. CONCLUSION

In this paper we have presented a spatial representation scenario that is useful for extracting symbolic representations from geometric models. It can be used to represent relations in any dimensions for objects that are either aligned to the reference frame, or are at arbitrary angles.

The essence of this approach is that it preserves the information of contact, no-contact and tangency. Given two objects, one can determine if they are flush along some face or line, and this contains all the information required to identify accidental alignments and relative positions of objects. Such relations, which describe the properties at the boundaries of an object, are critical in the analysis of many systems such as VLSI, circuits, mechanisms, structures - indeed almost anywhere geometrical relations are important. The representation has been applied to generating directions in city maps, acquiring data from graph diagrams in texts, to path planning for an indoor mobile robot and for learning concepts related to spatial knowledge.

6.1. Explanation-Based Learning

Another capability inherent in the system is building commonsense theories for explanation-based learning. One can define "naive" notions such as the need for support against gravity, and that solids cannot physically intersect. The following predicate formulation in this logic can be used to ensure that every relationship in the model passes the no-intersection condition:

gravity rule:

$$(\forall x)\ \{\ \sim\text{Lies-on-floor}\,(x)\ \Rightarrow\ (\exists y)\ y \begin{pmatrix} \text{INT} \\ \text{INT} \\ \text{-b} \end{pmatrix} x\}$$

no-intersection rule:

$$(\forall x)(\forall y)\ \{\ \sim x \begin{pmatrix} \text{INT} \\ \text{INT} \\ \text{INT} \end{pmatrix} y\ \}$$

where INT = {-i, bi, ii, -f, bf, if, -+, b+, i+}, and implies a degree of overlap.

In essence, the gravity rule states that if an object A is not lying on the floor, then there must be some other object such that it overlaps A in the x and y directions, and supports it ("-b") in z. Similarly the no-intersection rule prohibits two objects from overlapping in all three axes. Note that the INT relation {-i, bi, ii, -f, bf, if, -+, b+, i+} is a hierarchical abstraction for the nine relations contained in it and can be thought of as a generalized overlap. Many similar conceptual clusterings are possible, for example: A smaller-than B = {bi,ii,fi}; touch-contact = {-b, +f}; no-contact= {++, --}; flush-overlap = {-f, b+, bi, if}, etc.

Furthermore, there are built-in continuities that can be used to formulate powerful "inductive bias"es for learning. This is shown in the continuum graph of figure 5, which shows the progression of states as objects move relative to one another. One could use inductive bias to conclude, for example for the well-known arch structure, that if A *overlaps* B and A *contained-in* B are both valid rules, then all in-between relations are also valid rules: A(-i, bi, ii)B are all valid constructs, and A(bf, -f, -+, b+)B, etc., are invalid rules (since A is on the smaller-than-B branch of the continuum).

This model has been used for spatial learning using a robot/teach pendant setup for recognizing the structure of geometric assemblies (with a overhead camera to obtain the part geometry information). Three-dimensional structures are created using the teach pendant and these are identified by the user as positive or negative examples. The system then obtains the underlying concept behind the examples shown [Mukerjee and Bratton 90].

The discussion on transitivity, and the very nature of binary spatial relations raises the spectre of combinatorial explosion. A little consideration reveals several ameliorating aspects. One of the motivating factors behind this model is that all relations should be *locally relevant*. If Galveston is near Houston and Richardson is a suburb of Dallas, then the relation between Richardson and Galveston does not need to be represented directly. For simple path planning problems, it can be shown that for non-contact relations, only nearest neighbors need be modeled, and objects such as roads, three relations need to be stored. Thus the total number of relations is sharply reduced proportional to the number of neighbors and intersections (m). Occasionally, one may store additional information- for example, global as well as local orientation, to constrain the angle further. As the number of objects (N) increases, the number of nearest neighbors per object remains constant. While the number of intersections may increase, it is usually small, and is bounded by the resolution in the domain. This means that the storage requirements are O(mN), and if m can be bounded, we obtain O(N) storage.

At this point it would be appropriate to note that clearly there are a number of spatial relations of interest that cannot be represented well using this mechanism. One such instance is that of "near" or "large", which require some degree of quantitative information which this model does not provide for, although one can extend the logic with the notion of operators that would permit one to model such attributes also. However, it must be realized that for these two predicates at least, there is considerable ambiguity involved in the semantics, and to model these may introduce a degree of arbitrariness that would defeat one of the principal objectives of this work. More precise definitions can be obtained by representing concepts such as these as a predicate in terms of the translation and rotation operators mentioned above.

In conclusion, we have discussed a simple yet powerful mechanism for representing the spatial relations between objects. This technique offers expressive power and logical transitivity, and is capable of dealing with imprecision in the spatial knowledge.

References

[Allen 83] Allen, James F., Maintaining knowledge about temporal intervals, Communications of the ACM, vol.26(11), November 1983, pp.832-843.

[Ambler and Popplestone 75] Ambler, A.P., and R.J. Popplestone, Inferring the positions of bodies from specified spatial relations, Artificial Intelligence, v.6:129-156.

[Dennett 75] Dennett, David C., Spatial and temporal uses of english prepositions - an essay in stratificational semantics, Longman Group, London 1975.

[[Ernest 86] Ernest, Davis, Representing and acquiring geographic knowledge, Morgan Kaufman, Los Altos, CA, 1986, 223 pages.

[King and Mukerjee 90] King, Scott, and Mukerjee, Amitabha, Inexact Visualization, IEEE Conference on Visualization in Biomedical Computing (Also Texas A&M TR 90-05), May 22-25, 1990, Atlanta.

[Kuipers and Byun 88] Kuipers, Benjamin J., and Yung-Tai Byun, A robust qualitative method for robot spatial learning, AAAI 88, p.774-779.

[Malik and Binford 83] Malik, Jitendra and Thomas O. Binford, Reasoning in Time and Space, IJCAI-1983, p. 343-345.

[McCarthy 77] McCarthy, John, Epistemological problems of artificial intelligence, Proceedings IJCAI-77, Cambridge MA, 1977, p.1038-1044.

[Mukerjee and Bratton 90] Mukerjee, Amitabha, and Samuel E. Bratton, Qualitative Spatial Reasoning and the learning of structural descriptions, Texas A&M University Technical Report 90-010.

[Peuquet and Ci-Xiang 87] Peuquet, Donna J., and Zhan Ci-Xiang, An algorithm to determine the directional relationship between arbitrarily-shaped polygons in the plane, Pattern Recognition, v.20(1):65-74, 1987.

[Requicha 80] Requicha, A.A.G., Representation for Rigid solids: Theory, methods, and systems, ACM Computer Surveys, Dec. 1980.

[Retz-Schmidt 88] Retz-Schmidt, Gudula, 1988. Various views on spatial prepositions, AI Magazine, Summer 1988, p. 95-105.

[Winston 75] Winston, Patrick Henry, 1975. Learning structural descriptions from examples, The Psychology of Computer Vision, ed. Patrick Henry Winston, McGraw Hill, 1975, p.157-209.