Title

A Framework for the Definition of Topological Relationships and an Algebraic Approach to Spatial Reasoning Within this Framework (91-7)

Permalink

Authors

Egenhofer, Max J.
Herring, John R.
Smith, Terence
et al.

Publication Date

1991-09-01

NCGIA

National Center for
Geographic Information and Analysis

A Framework for the Definition of
Topological Relationships and an Algebraic
Approach to Spatial Reasoning
Within this Framework

Max J. Egenhofer
National Center for Geographic Information & Analysis
Orono, ME 04469

John R. Herring
Intergraph Corporation
Huntsville, AL 35807-2174

Terence Smith
National Center for Geographic Information & Analysis
Santa Barbara, CA 93106-4060

Keith K. Park
Department of Computer Science and Geography
Santa Barbara, CA 93106-4060

Technical Report 91-7
September 1991

Simonett Center for Spatial Analysis
University of California
35 10 Phelps Hall
Santa Barbara, CA 93106-4060
Office (805) 893-8224
Fax   (805) 893-8617
ncgia@ncgia.ucsb.edu

State University of New York
301 Wilkeson Quad, Box 610023
Buffalo NY 14261-0001
Office (716) 645-2545
Fax   (716) 645-5957
ncgia@ubvms.cc.buffalo.edu

University of Maine
348 Boardman Hall
Orono ME 04469-5711
Office   (207) 581-2149
Fax   (207) 581-2206
ncgia@spatial.maine.edu

1 Introduction

Queries in spatial databases, such as Geographical Information Systems (GIS), image data bases, or CAD/CAM systems, are often based on the relationships among spatial objects. For example, in geographical applications typical spatial queries are "Retrieve all cities within 5 miles of the interstate highway I-95" or "Find all highways in the states adjacent to Maine." Current commercial database query languages do not sufficiently support such queries, because these languages provide only tools to compare equality or order of simple data types, such as integers or strings. The incorporation of spatial relationships over spatial domains into the syntax of a spatial query language is an essential extension beyond the power of traditional query languages, such as SQL [Roussopoulos 1988] [Egenhofer 1988]. Some experimental spatial query languages support queries with spatial relationships; however, the diversity, semantics, completeness, and terminology of these relationships vary dramatically.

Besides the *formulation* of queries with spatial conditions, their *processing* is also of importance. Spatial queries can be easily solved if all spatial relationships between the objects of interest are explicitly stored; however, such a scenario is unrealistic, even for relatively small data collections [Davis 1986]. Impediments are the vast amount of storage space to keep the large variety of spatial relationships between any two objects and the complexity of maintaining such a setting with every update of the geometry of an object. Instead, it is necessary to derive the spatial relationships from their geometry or spatial location. Such a concept needs, of course, a thorough understanding of *what* possible spatial relationships are and *how* they can be determined.

To help clarify the users' diverse understandings about the semantics of spatial relations and enable the processing of spatial queries it is proposed to formally describe the relationships. Users can then examine whether a specific implementation concurs with their expectations and system designers and engineers have formal guidelines for their implementations.

Helpful for such an approach are the linguists' observations about natural language terms for the description of spatial relationships. The use of spatial relationships (in the English language) is independent of the size and material of the reference objects, yet context in which a specific relationship occurs is essential for the selection of the correct terms [Talmy 1983].

Various formal approaches have been proposed. One such formalism uses the primitives *distance* and *direction* in combination with the logical connectors *AND*, *OR*, and *NOT* [Peuquet 1986]. This derivation of topology from metric is conceptually doubtful and leads to implementation problems in computers due to the finiteness of the underlying number system [Franklin 1984] [Egenhofer 1989]. A definition of topological relation-

ships in terms of set operations upon point-sets attempts to describe topological relationships [Güting 1988]; however, it does not distinguish between the topologically distinct parts of point-sets. The point-set approach has been augmented by the distinction of *boundary* and *interior* for some relationships [Pullar 1988a]. In a more systematical approach, the comparison of boundaries with boundaries and interiors with interiors allows for the distinction of four topological relationships, still missing the distinction of some significantly different situations [Wagner 1988].

This paper presents a comprehensive theory for binary topological relationships between n-dimensional spatial objects embedded in an n-dimensional space. The classification of topological relationships is based upon the comparison of all possible combinations of boundaries and interiors of two objects. This approach uses purely topological means to distinguish different topological relationships and provides complete coverage, i.e., any possible constellation between two spatial objects is described by exactly one of the sixteen relationships identified. The previous presentation of relationships between 1-dimensional intervals [Pullar 1988b] is a special case within this framework.

The remainder of this paper is organized as follows: in the next section, different types of spatial relationships are discussed, focussing on topological relationships. Then point-sets are introduced as the underlying model for spatial objects to investigate topological relationships between them. Section 3 presents our theory of binary topological relationships between point-sets in terms of the intersections of their boundaries and interiors. The subsequent investigations provide an answer to the question "Which relationships can be realized in a two-dimensional space?" and show geometric interpretations. Finally, the conclusions in section 4.

2   Spatial Relationships

2.1   A Classification of Spatial Relationships

The entire domain of spatial relationships is too complex and diverse to be treated by a single method in a single attempt. It appears rather favorable to define a framework within which the existence of relationships can be investigated. The identification of similar relationships and the discrimination of dissimilar ones will be supported from the foundation upon mathematical principles of such an approach.

A helpful approach is the categorization of spatial relationships according to different spatial concepts on which they rely. The following classification distinguishes three fundamental types of relationships, the properties of which correspond to the three fundamental mathematical concepts *topology*, *order*, and *algebra*. This classification is not complete since it does not consider *fuzzy* relationships, such as *close* and *next_to* [Robinson 1987], or rela-

tionships which are expressions about the motion of one or several objects, such as *through* and *into* [Talmy 1983]. These types of relationships are not fundamental and rather combine several independent concepts. Motion, for example, can be seen as the combination of spatial and temporal aspects. It appears natural for each category to develop independent formalisms describing the relationships [Pullar 1988b] [Kainz 1989].

- Topological relationships are invariant under topological transformations, such as translation, scaling, and rotation. Examples are concepts like *neighbor* and *disjoint*.

- Spatial order relationships rely upon the definition of order or strict order. In general, each order relation has a converse relationship. For example, *behind* is a spatial order relation based upon the order of *preference* [Freeman 1975] with the converse relationship *in_front*.

- Metric relationships exploit the existence of measurements, such as distances and directions. For instance, "within 5 miles from the interstate highway I-95" describes a corridor based upon a specific distance.

2.2   Point-Set Topology

Topological notions include the concepts of continuity, closure, interior, and boundary, which are defined in terms of neighborhood relations. In this context, topological equivalence is considered a crucial criterion for the comparison of relationships among objects. Topological properties often conflict with metric ones. It is important to keep in mind that topological equivalence does not preserve distances; therefore, the subsequent investigations are based upon continuity which is described in terms of *coincidence* and *neighborhood*.

The data model for spatial regions is based on the classical point-set model and the point-set topological notions of *interior* and *boundary* [Spanier 1966]. The *interior* of a point-set Y, denoted by Y°, is defined to be the union of all open sets that are contained in Y. The *closure* of Y, denoted by $\overline{Y}$, is defined to be the intersection of all closed sets that contain Y. The *boundary* of Y, denoted by ∂Y, is then the intersection of the closure of Y and the closure of the complement of Y, i.e., $\partial Y = \overline{Y} \cap \overline{X - Y}$.

The concepts of separation and connectedness are crucial for establishing the forthcoming topological spatial relationships between point-sets. Let $Y \subset X$. A *separation* of Y is a pair A, B of subsets of X satisfying the three conditions $A \neq \emptyset$ and $B \neq \emptyset$; $A \cup B = Y$; and $\overline{A} \cap B = \emptyset$ and $A \cap \overline{B} = \emptyset$. If there exists a separation of Y then Y is said to be *disconnected*, otherwise Y is said to be *connected*. A *region* is then a non-empty connected set X in $\mathbb{R}^2$.

The *dimension* of the space be defined as the number of independent vectors which are the base elements of the corresponding vector space. Examples of 1-dimensional spaces are

a line, the border of a circle, and its topological images; common 2-dimensional spaces are the open and the closed disks, and their topological images. An important property of an n-dimensional space is that it may embed elements of dimension at most n. This property gives rise to the definition of the dimension of an object. An object has the same dimension n as its embedding space if the object exists in this space, but there is no homeomorphic mapping for the object into a space of dimension n-1. A region, for instance, exists in a two-dimensional space and there is no homeomorphic mapping which may transform a region into a one-dimensional space. Hence, a region is of dimension 2. The standard definitions are: a node is of dimension 0, an edge of dimension 1, a region of dimension 2, etc.

The *codimension* defines the difference between the dimension of the embedding space and the dimension of an object. For example, codimension 1 for a region describes that it is located in a 3-dimensional space. The above definitions imply that the codimension can be never less than zero, and is zero if and only if the object and the space are of the same dimension.

3   A Theory of Topological Relationships

First, a framework for the definition of binary topological relationships will be introduced, consisting of the intersections of boundary and interior of the two objects to be compared. The intersections are analyzed according to their content (i.e., empty or non-empty) which leads to sixteen different specifications for topological relationships. The investigations of the existence of the sixteen relationships demonstrate that only nine occur between two n-dimensional objects with codimension 0. A subset of eight relationships can be identified if the boundary of each object is connected.

3.1   Hypothesis

Definition 1 *The topological relationship R between two spatial objects o1, o2 is based upon the comparison of the intersections of the boundary and interior of o1 with the object parts of o2.*

Boundary and interior can be combined to form the four fundamental criteria of spatial relationships. These are: (1) common boundary parts as the intersection of *boundary*, denoted by $\partial\partial$, (2) common interior parts ($^{\circ\circ}$), (3) boundary as part of the interior (∂°), and (4) interior as part of the boundary ($^{\circ}\partial$). Subsequently, $\partial\partial$ and $^{\circ\circ}$ will be referred to as the two *corresponding* intersections, and ∂° and $^{\circ}\partial$ as the two *opposite* intersections.

Different topological relationships may be identified by comparing *topological invariants* of the intersections. Topological invariants are properties which are preserved under topological transformations.

Definition 2 *Topological invariants of the intersections of the object parts characterize the topological relationship between the objects.*

In this context, the following topological invariants are considered:

- the content (i.e., emptiness or non-emptiness) of the intersection;

- the number of separate boundary intersections; and

- the dimension of the intersection.

The content of the intersections is selected as the fundamental criterion for topological relationships because

- it describes a closed set of relationships with complete coverage; and

- more detailed relationship can be considered a subset of it.

With the binary values empty ($\emptyset$) and non-empty ($\neg\emptyset$) a total of sixteen different specifications is given which provide the basis for the formal definition of the spatial relationships (table 1).

3.2 Existence of Relationships

The targets in this paper are those 2-dimensional objects that are homeomorphic images to connected point-sets with non-empty interiors and connected boundaries. They will be referred to as *regions*.

Not all sixteen specifications exist between two regions with codimension zero. The proof has two parts both eliminating a set of relationships due to a condition among the intersections.

Lemma 1 *The relationships r_4, r_5, r_8, r_9, r_{12}, and r_{13} do not exist between two regions with codimension 0.*

Proof: Any point in the interior of a region must have a 2-cell surrounding it which is also contained within the object. Any two 2-cells surrounding a point must contain a 2-cell containing the point in their intersection and any point on the boundary of an object must be arbitrarily close to some point in the interior. Thus if X is a point in $\partial A \cap B^\circ$, then there is another point Y, close to X, in $A^\circ \cap B^\circ$. This proves that if the boundary of a region A intersects the interior of another region B, then there is a point interior to both. In terms of the four intersections, if at least one of the opposite intersections is non-empty, then the intersection of both interiors must be non-empty as well. This theorem eliminates the

	$\partial\partial$	$\circ\circ$	$\partial\circ$	$\circ\partial$
r_0	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
r_1	$\neg\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
r_2	$\emptyset$	$\neg\emptyset$	$\emptyset$	$\emptyset$
r_3	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$	$\emptyset$
r_4	$\emptyset$	$\emptyset$	$\neg\emptyset$	$\emptyset$
r_5	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$	$\emptyset$
r_6	$\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$
r_7	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$
r_8	$\emptyset$	$\emptyset$	$\emptyset$	$\neg\emptyset$
r_9	$\neg\emptyset$	$\emptyset$	$\emptyset$	$\neg\emptyset$
r_{10}	$\emptyset$	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$
r_{11}	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$
r_{12}	$\emptyset$	$\emptyset$	$\neg\emptyset$	$\neg\emptyset$
r_{13}	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$	$\neg\emptyset$
r_{14}	$\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$
r_{15}	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$

Table 1: The sixteen specifications of binary topological relationships based upon the criteria of empty and non-empty intersections of boundaries and interiors.

relationships r_4, r_5, r_8, r_9, r_{12}, and r_{13} which have empty interior-interior intersections and at least one of the boundary-interior and interior-boundary intersections is empty as well. $\square$

Lemma 2 *The relationships r_2 and r_{14} do not exist between two regions with codimension 0.*

Proof: This proof is based upon the Jordan-Brower separation theorem [Spanier 1966]:

A 1-sphere, embedded in Euclidean 2-space, separates that space into two regions.
The sphere is then the common boundary of the two separated regions.

For any region it holds true that if the boundaries of two regions in 2-dimensional space are disjoint, then the interiors are either disjoint or one point-set is completely contained within the interior of the other. In terms of the four intersections, if the boundary intersection is empty, then either all other intersections are empty as well; or the interior intersection and one of the two boundary-interior intersections are non-empty as well. This restriction

eliminates five of the eight specifications with an empty boundary intersection, namely r_2, r_4, r_8, r_{12}, and r_{14}. □

As a result, only the eight relationships r_0, r_1, r_3, r_6, r_7, r_{10}, r_{11}, and r_{15} exist between two spatial regions with codimension zero. If the boundary of a region need not be connected, i.e., the object may have holes, then r_{14} would be a possible topological relationships [Egenhofer 1990].

3.3 A Geometric Interpretation

A geometric interpretation of the abstract definition will be given below. The interpretations refers to prototype relationships presented for regions with codimension 0. It is not a matter of the definition of terms for the relationships—a systematic terminology $r_0 \ldots r_{15}$ would provide the same service. Nevertheless, it is felt that meaningful names improve the understanding of the abstract definitions of the relationships.

Definition 3 *If all four intersections among all object parts are empty, then the two objects are* disjoint *(figure 1a).*

Disjoint is linear, such that two objects are either disjoint or they are not. The specification for *not_disjoint* follows immediately from the definition above.

Definition 4 *If the intersection between the boundaries is not empty, whereas all other 3 intersections are empty, then the two objects* meet *(figure 1b).*

The nature of *meet* is such that it only matters that the two objects share at least a common part of the boundary.

Definition 5 *Two objects* overlap *if they have common boundaries and interiors, and the boundaries have common parts with the opposite interiors* (figure 1c).

Definition 6 *An object A* covers *another object B if both objects share common boundaries and interiors; B's interior intersects with the boundary of A; and none of A's interior is part of B's boundary.*

Covers has a converse relationship *covered_by* which has the reverse definition of the boundary-interior intersections (figure 1d).

Definition 7 *An object A is* inside *of another object B if (1) A and B share common interiors, but not boundaries, (2) A's boundary intersects with the interior of B, and (3) none of B's boundary coincides with A's interior.*

Like *covers*, *inside* has a converse relationship, called *contains*, with corresponding specifications which are the same except for the reverse opposite intersections (figure 1e).

Definition 8 *Two objects are equal if both intersections of boundary and interior are not empty while the two boundary-interior intersections are empty* (figure 1f).
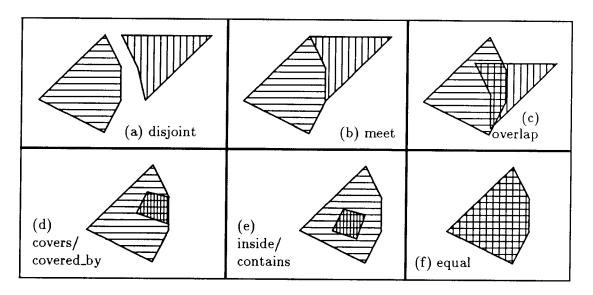

Figure 1: Examples of the relationships between two regions in a 2-dimensional space.

3.4 Dimensions of the Intersections

More details about topological relationships may be expressed by considering other topological invariants in addition to the emptiness/non-emptiness of object part intersections. Here, it will be investigated how the *dimension* of the boundary intersections allows for the definition of more detailed topological relationships.

The dimension of the boundary is defined as the largest dimension of all faces. The dimension of the intersection of two boundaries is then the largest dimension of the faces being part of the intersection. This gives rise to the differentiation of various detailed definitions for *meet*, *overlap*, and *covers/covered_by* based upon the dimension of the common boundaries. The other relationships are excluded from this consideration because they have empty boundary intersections (*disjoint*, *inside/contains*).

Two n-dimensional objects can *meet*, *overlap*, and *cover/be covered_by* in n different ways. These detailed relationships can be distinguished according to the dimension p of the common boundary, and are called *p-meet*, *p-overlap*, and *p-cover/p-covered_by*. For example, the common boundary of two regions can be of dimension 1 if they share one or several 1-faces. Then the relationship is called *1-meet*. The second *meet* relationship in

2-D, *0-meet*, requires that the dimension of the common boundary is 0 (i.e., the common bounding parts are only nodes). Figure 2 shows examples of the differences between the 0- and 1-relationships for *meet*, *overlap*, and *cover*.
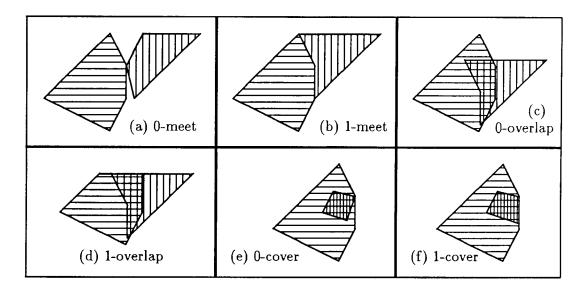

Figure 2: Examples of the detailed relationships between regions in 2-D considering different dimensions in the boundary intersections.

4   Conclusion

A formalism for the definition of topological relationships has been presented. It is based upon purely topological properties and thus independent of the existence of a distance function. The topological relationships are described by the commonality of boundary and interior with the binary values empty and non-empty for these intersections, which gives rise to sixteen mutually excluding specifications.

The investigation of the relationships was restricted to those which yield between two spatial regions in a two-dimensional space. Eight of the sixteen relationships do not exist under this restriction. The remaining eight relationships serve as the framework for more detailed topological relationships (table 2). It can be extended by considering further topological invariants, such as the dimension of the boundary intersections or the number of separate boundary intersections.

Compared to the results of previous investigations of the relationships between one-dimensional, connected objects in 1-D [Pullar 1988b], almost the same set of relationships exists. The only difference is that in 1-D the relationship r_{14}, overlap with disjoint boundaries, exists and r_{15} does not, while in 2-D this is reverse [Egenhofer 1990]. Ongoing work

		$\partial\partial$	^{oo}	∂°	$^{\circ}\partial$
r_0	disjoint	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
r_1	meet	$\neg\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
r_3	equal	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$	$\emptyset$
r_6	inside	$\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$
r_7	coveredBy	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$
r_{10}	contains	$\emptyset$	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$
r_{11}	covers	$\neg\emptyset$	$\neg\emptyset$	$\emptyset$	$\neg\emptyset$
r_{15}	overlap	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$	$\neg\emptyset$

Table 2: The eight specifications of topological relationships between two spatial regions in 2-D.

investigates the application of this theory for codimensions greater than zero and to describing the relationships between spatial objects of different dimensions.

Acknowledgement

References

[Davis 1986] E. Davis. Representing and Acquiring Geographic Knowledge. Research Notes in Artificial Intelligence, Morgan Kaufmann Publishers, Inc., Los Altos, CA, 1986.

[Egenhofer 1988] M. Egenhofer and A. Frank. Towards a Spatial Query Language: User Interface Considerations. In: D. DeWitt and F. Bancilhon, editors, 14th International Conference on Very Large Data Bases, pages 124–133, Los Angeles, CA, August 1988.

[Egenhofer 1989] M. Egenhofer et al. A Topological Data Model for Spatial Databases. In: A. Buchmann and O. Gunther, editors, Symposium on the Design and Implementation of Large Spatial Databases, Springer-Verlag, New York, NY, July 1989.

[Egenhofer 1990] M. Egenhofer and R. Franzosa. Point-Set Topological Spatial Relations. submitted for publication, 1990.

[Franklin 1984] W.R. Franklin. Cartographic Errors Symptomatic of Underlying Algebra Problems. In: International Symposium on Spatial Data Handling, pages 190–208, Zurich, Switzerland, August 1984.

[Freeman 1975] J. Freeman. The Modelling of Spatial Relations. Computer Graphics and Image Processing, 4:156–171, 1975.

[Güting 1988] R. Güting. Geo-Relational Algebra: A Model and Query Language for Geometric Database Systems. In: J.W. Schmidt et al., editors, Advances in Database Technology—EDBT '88, International Conference on Extending Database Technology, Venice, Italy, pages 506–527, Springer Verlag, New York, NY, 1988.

[Kainz 1989] W. Kainz. Order, Topology, and Metric in GIS. In: ASPRS-ACSM Annual Convention, pages 154–160, Baltimore, MD, March 1989.

[Peuquet 1986] D. Peuquet. The Use of Spatial Relationships to Aid Spatial Database Retrieval. In: D. Marble, editor, Second International Symposium on Spatial Data Handling, pages 459–471, Seattle, WA, 1986.

[Pullar 1988a] D. Pullar. Data Definition and Operators on a Spatial Data Model. In: ACSM-ASPRS Annual Convention, pages 197–202, St. Louis, MO, 1988.

[Pullar 1988b] D. Pullar and M. Egenhofer. Towards Formal Definitions of Topological Relations Among Spatial Objects. In: D. Marble, editor, Third International Symposium on Spatial Data Handling, pages 225–242, Sydney, Australia, August 1988.

[Robinson 1987] V.B. Robinson and R.N. Wong. Acquiring Approximate Representations of Some Spatial Relations. In: N.R. Chrisman, editor, AUTO-CARTO 8, Eighth International Symposium on Computer-Assisted Cartography, pages 604–622, Baltimore, MD, March 1987.

[Roussopoulos 1988] N. Roussopoulos et al. An Efficient Pictorial Database System for PSQL. IEEE Transactions on Software Engineering, 14(5):630–638, May 1988.

[Spanier 1966] E. Spanier. Algebraic Topology. McGraw-Hill Book Company, New York, NY, 1966.

[Talmy 1983] L. Talmy. How Language Structures Space. In: H. Pick and L. Acredolo, editors, Spatial Orientation: Theory, Research, and Application, Plenum Press, New York, NY, 1983.

[Wagner 1988] D. Wagner. A Method of Evaluating Polygon Overlay Algorithms. In: ACSM-ASPRS Annual Convention, St. Louis, MO, 1988.

An Algebraic Approach to Spatial Reasoning

Terence R. Smith

National Center for Geographic Information and Analysis
and Departments of Computer Science and Geography
University of California at Santa Barbara
Santa Barbara, CA. 93106. U.S.A.

and Keith K. Park

Department of Computer Science and Geography
University of California at Santa Barbara
Santa Barbara, CA. 93106. U.S.A.

Abstract. A simple, exemplary system is described that performs reasoning about the spatial relationships between members of a set of spatial objects. The main problem of interest is to make sound and complete inferences about the set of all spatial relationships that hold between the objects, given prior information about a subset of the relationships. The spatial inferences are formalized within the framework of relation algebra and procedurally implemented in terms of constraint satisfaction procedures. Although the approach is very general, the particular example employs a new "complete" set of topological relationships that have been recently described in the literature. In particular, a relation algebra for these topological relations is developed, and a computational implementation of this algebra is described. Systems with such reasoning capabilities appear to have many applications in geographical analysis, and could be usefully incorporated into GIS and related systems.

1 Introduction

The primary intent of this paper is to show how a recently described, "complete" set of topological relationships (Egenhofer and Franzosa 1991) may be used to serve as the basis for a procedure that performs *spatial reasoning*. A secondary intent is to encourage further research along similar lines, in terms of the development of more expressive sets of spatial relationships and in terms of more robust procedures for reasoning.

It is increasingly recognized that database and knowledge-base systems may be viewed as systems that contain models of some set of phenomena in the real world (Smith, Ramakrishnan and Voisard, 1991). It is therefore reasonable to develop procedures that access such systems by deriving inferences concerning the phenomena represented in the models. In the particular case of GIS, such inferential capabilities typically relate to entities that occupy space and change over time, and involve both the spatial and non-spatial attributes and relationships of the entities.

Important issues associated with such computational modeling activities are the formalism in terms of which the domain knowledge is represented and the inferential procedures for reasoning about the entities in the domain. Typical approaches to modeling include the development of:

1. conceptualizations of some set of phenomena of interest in the world in terms of a set of objects, properties of the objects and relationships between the properties of objects;

2. representations of this conceptualization using a formal language with a well-defined syntax and declarative semantics. The objects and the relations conceptualized in the preceding step serve as the domain of terms for the expressions of this formal language;

3. sets of inference procedures which perform "reasoning", and which are axiomatized for the domain of objects and relationships.

If we are to employ this approach in developing GIS with deductive reasoning capabilities which are logically *sound* and *complete*, we need to develop models of spatio-temporal objects and their interrelationships, and formal languages in which these may be represented. We also need to define axiomatic theories of the language to give a procedural semantics of our intended model. In this paper, we describe one example of such a model and an associated inference system.

1.1 The Problem

We assume that we are given an database in which a set of spatial objects and their interrelationships are explicitly represented. We are frequently faced with a situation in which such a database contains *implicit* information about many of the objects and spatial relationships that are referred to in the database (i.e., information not in the explicit form that we may require). Hence we may ask whether:

1. we may correctly infer all the facts about the objects and their relationships that are implicitly represented in the database;

2. the current database of objects and relations is consistent, and whether it will remain consistent if we add new objects or relationships.

These two questions are closely related and clearly require a reasoning process of some form which is based upon an appropriate model of the objects and their relationships and which can automated in a sound and complete manner. The *soundness* of a reasoning procedure guarantees that the procedure leads only to correct inferences. Hence, for example, any inferred spatial relationships are always logically implied by the initial set of relationships. The *completeness* of a reasoning procedure guarantees that the procedure leads to all of the correct inferences. Hence, for example, it will eventually produce all of the correct inferences about relationships that are logically implied by the initial set of relationships. The initial set of relations that is provided to the procedure must, of course, be consistent to make the discussion of soundness and completeness non-trivial (Genesereth and Nilsson 1987).

We examine answers to these questions in terms of an exemplary system that is able to make sound and complete inferences about the spatial relationships that hold between sets of spatial objects. In the remainder of the paper, we discuss some of the research background that is relevant to these questions. We then describe a simple conceptualization of spatial objects and their interrelationships, and an abstraction by means of which we may represent the objects and their relationships and which we may employ in making inferences concerning the spatial relationships of the objects. We follow this with a description of general approaches to implementing such reasoning schemes and provide a description of a particularly efficient representation. We also provide simple examples to clarify our general approach. It is to be emphasized that a major goal of the paper is to stimulate further research towards systems that are capable of making inferences about complex spatial objects represented in large spatial databases.

1.2 Research Background

The fundamental basis for our approach to problems of spatial reasoning was provided by Allen in his research concerning reasoning about relationships among elements of a set of intervals of time (Allen 1983). Allen introduced the notion of temporal interval as a primitive (analogous to a spatial pointset) and developed a formalism for expressing relationships between temporal intervals. Allen considered thirteen relationships, including *overlap, meet, before, equal, during, start, finish* and their converses. Allen expressed knowledge about a given set of intervals in terms of a special graph called a binary constraint network, in which nodes represent intervals and labeled edges between nodes represent any subset of the thirteen relationships that may hold between intervals. This network may be viewed as a database system that allows one to draw logical inferences concerning implicit relationships between different time intervals. The database is maintained in such a way that whenever

a new relation is entered into the network, a constraint satisfaction procedure is executed to preserve the consistency of the network. The constraint satisfaction procedure essentially computes the transitive closure of temporal relations. The novel feature of Allen's approach is that by giving *relations* the status of first class entities, the domain of temporal intervals with infinite real metrics is provided with a computational reasoning scheme.

There have been many attempts to define the nature and classes of spatial relationships that frequently arise in many applications and there have also been attempts to define the nature of reasoning about such relationships. We refer the reader to material referenced, for example, by Frank (1991), Kuipers (1978) and Gopal and Smith (1990) for general treatments of relationships and reasoning. Few of these studies have, however, attempted to structure the concepts in terms of formal systems, and of those studies of which we are aware, a significant proportion have been based on Allen's research, although there has been some research (such that of Chang, Jungert and Li, 1990) that employs a diffent approach. In the research that has adopted an approach analogous to that of Allen, spatial pointsets replace the temporal intervals of Allen's work and specific classes of spatial relationships replace the 13 relationships defined by Allen. Guesgen (1989), for example, extended Allen's approach to one-dimensional spatial domains, and attempted to extend the one-dimensional version to two spatial dimensions, but with limited success because of ambiguities in the representations. Maddux defined a class of compass algebras on points in Euclidean space, and showed that constraint satisfaction in any compass algebra is NP-hard (Maddux 1989). Malik specified spatial relationships in terms of linear inequalities between boundary surfaces of objects and based spatial reasoning on an application of linear programming that employed the simplex algorithm (Malik and Binford 1983).

We have chosen to adopt an approach that is analogous to that of Allen, since it is remarkably simple and appears to have great potential for generalization. We believe that the formal basis that we propose has not yet been applied with respect to the particular concept of spatial objects and the particular set of spatial relationships that are described in the following sections.

1.3 Spatial Relations based on Pointset Topology

For describing and reasoning about the most general classes of spatial phenomena, we require a relatively sophisticated conceptualization of the world and a powerful formal language in which we may define and manipulate arbitrarily complex spatial objects and their interrelationships. While such languages are currently under development (Smith, Ramakrishnan and Voisard, 1991), it suffices *for present purposes* to consider a very simple conceptualization of a set of spatial phenomena and a very simple formal language for representing such objects and their relationships.

We assume that a spatial object is defined as some entity possessing an essential projection onto some geometrical space, and that this projection may be represented as a pointset with a well defined interior and a connected boundary. We also assume that any spatial rela-

tionship of interest between objects may be defined in terms of the eight, binary topological relationship discussed in Egenhofer and Franzosa (1991). We consider only binary spatial relationships since n-ary relations may be equivalently represented in terms of a conjunction of binary relationships.

The set of relationships discussed by Egenhofer and Franzosa is simple but "complete" set of relationships that is based on ideas developed in elementary pointset topology (Munkres 1975). If A is a pointset, then:

- the interior of A, denoted by A^i, is defined to be the union of all open sets that are contained in A

- the closure of A, denoted by A^c, is defined to be the intersection of all closed sets that contain A

- the boundary of A, denoted by A^b, is defined to be the intersection of the closure of A and the complement of A

One may then enumerate a total of 16 binary topological relationships (including their converses) for any two pointsets A and B according to the pattern of intersections between the boundaries and the interiors of A and B, i.e., all possible truth values for the 4-tuple $[A^b \cap B^b, A^i \cap B^i, A^b \cap B^i, A^i \cap B^b]$. It is clear that these relations are *mutually exclusive*, and form a *partition* of the set of all relations. It is these two properties of this set of topological relations that allow us to develop an algebra on these relations and to carry out algebraic manipulation based on this algebra (Appendix contains a brief description of relation algebra.)

Egenhofer and Franzosa (1991) provided formal proofs that only 8 of the 16 topological relations can actually occur among pointsets of 2- or higher dimensions with co-dimension 0 between the object and the embedding space. This is the case of interest for most GIS, and in this case, the 8 relations form a partition of the set of all possible relations. We may view each of the 8 relations as being *atomic*, i.e., each may be viewed as a singleton set containing the smallest non-zero element in the appropriate relation algebra. These atomic relations and their defining conditions are listed in Table 1 together with their denotational symbols. Figure 1 provides a geometric interpretation of the relations.
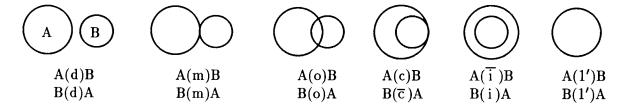

| A(d)B | A(m)B | A(o)B | A(c)B | A($\bar{\imath}$)B | A(1')B |
| B(d)A | B(m)A | B(o)A | B($\bar{c}$)A | B(i)A | B(1')A |

Figure 1: Geometric Interpretations of Atomic Spatial Relations

$A^b \cap B^b$	$A^i \cap B^i$	$A^b \cap B^i$	$A^i \cap B^b$	Relation	Symbol	Converse
ϕ	ϕ	ϕ	ϕ	A and B are disjoint	d	self-converse
$\neg\phi$	ϕ	ϕ	ϕ	A meets B	m	self-converse
$\neg\phi$	$\neg\phi$	$\neg\phi$	$\neg\phi$	A and B overlaps	o	self-converse
$\neg\phi$	$\neg\phi$	ϕ	ϕ	A equals B (identity)	1'	self-converse
$\neg\phi$	$\neg\phi$	ϕ	$\neg\phi$	A covers B	c	$\overline{c}$
ϕ	$\neg\phi$	$\neg\phi$	ϕ	A is inside of B	i	$\overline{i}$
$\neg\phi$	$\neg\phi$	$\neg\phi$	ϕ	B covers A	$\overline{c}$	c
ϕ	$\neg\phi$	ϕ	$\neg\phi$	B is inside of A	$\overline{i}$	i

Table 1: Eight Atomic Spatial Relations and Their Symbols

2 Relation Algebra and the Composition of Spatial Relations

2.1 Abstraction of Spatial Knowledge in Constraint Network

We now provide an *application* of the set of topological relationships described above as a basis for a particular form of spatial reasoning. Consider a pair of spatial objects, and assume that we have reason to believe that the relations holding between these objects lie in some subset of this set of atomic relations. We may interpret such a subset as a *disjunction* of relations, in the sense that any of the relations occurring in the subset may hold between the pair of objects and, since the relationships are disjoint, no more than one relationship in the disjunction will hold. We employ the infix notation $o_i(R)o_j$ to represent the fact that the topological relationship R holds between the spatial projections (or pointsets) of objects o_i and o_j . The converse of R is represented by $\overline{R}$, thus $o_i(R)o_j \equiv o_j(\overline{R})o_i$. The disjunction, or sum, of two relations R and S over the domain of spatial objects is denoted $(R + S)$, where $o_i(R + S)o_j \equiv o_iRo_j \vee o_iSo_j$ and $\vee$ denotes disjunction ("or"). For example, if "c" denotes the relation *covers* and "o", the relation *overlaps*, respectively, then $o_1(c + o)o_2$ represents "*either* o_1 covers o_2 *or* o_1 overlaps o_2 ". In general, there are 2^8 different disjunctions of 8 atomic relations that may be used to characterize the feasible relationships between a given pair of spatial objects.

We may represent our knowledge about a given set of objects in terms of a graph, which we term a *binary spatial constraint network* (BSCN). Figure 2 provides an example of such a representation for the case of three spatial objects o_i , o_j , o_k , and their relationships.

Each spatial object is mapped into a distinct node in the network, and the binary relationships among them are represented by the *direction* and the *label* of the edges. An edge is directed to distinguish between a binary relation and its converse. The solid edges represent relations that are *explicitly* specified in some database representing objects and their relationships. Since the converses of our relations are well defined from Table 1, the relationships between any pair of objects is represented by labeling a *single* directed edge. The edge R^b_{ji}
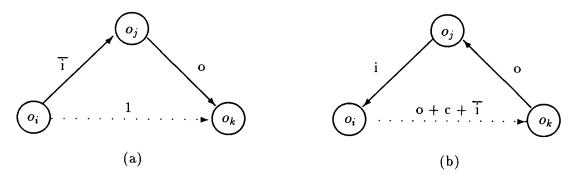
Figure 2: Spatial Knowledge Abstracted in BSCN

in Figure 2-b, for example, is the converse of R_{ij}^a in Figure 2-a, and both representations conveys the same information. Dotted edges, such as R_{ik}^a, represent relations that are not explicitly specified in the database. We thus initialize the label for such an edge with the *universal* relation "1", which is simply the disjunction of all possible relations, namely ($1'$ + d + m + o + c + $\bar{c}$ + i + $\bar{i}$). The derivation of the label ($\bar{i}$ + c + o) of R_{ik}^b in the Figure is described below.

The fact that the relations that are expressed as a disjunction indicates that our explicit knowledge of the relationships is incomplete. Hence we may view a disjunction of relations as a constraint on our knowledge of possible relationships. Within this framework, reasoning may be viewed in terms of "pruning" the set of feasible relations, or tightening the constraints on the labels of the underspecified edges. We may perform such reasoning if the *combined* information of the labels of the edges R_{ij} and R_{jk} (the *composition* of them), implies some information about the edge R_{ik} that is not explicitly encoded in the label. Our focus in this paper is on a formal basis for reasoning in this manner about the implied relationships.

We may formalize the concept of such reasoning by providing a formal representation for the BSCN and the associated reasoning. As in the case of relational databases, (see, for example, Ullman 1989), our formal representation may take the form of a calculus or an algebra. We have chosen to take the algebraic point of view because it appears to offer, in the present case, a greater degree of simplicity. In relation to the *calculus* approach, we note that a BSCN may be represented formally as a *logical conjunction* of terms, each of which is a *disjunction* of binary *predicates* representing the relationships (Ladkin 1990). The BSCN in Figure 2-b, for instance, may be represented as

$$o_j \,(\text{i})o_i \,\wedge\, o_k \,(\text{o})o_j \,\wedge\, [o_i \,(\text{o})o_k \,\vee\, o_i \,(\text{c})o_k \,\vee\, o_i \,(\bar{\text{i}}\,)o_k \,]$$

where $\wedge$ and $\vee$ represent logical conjunction ("and") and disjunction("or"), respectively. Hence we may translate the BSCN that represents a set of objects and their relationship into a set of formulae of first order logic, and reasoning may be viewed in terms of the predicate calculus with a well defined declarative semantics, from the model-theoretic point of view or with a well-defined procedural semantics from the proof-theoretic point of view (see Enderton, 1972). Hence if a BSCN were formulated as a logic program, it would require

formal theorem-proving techniques in order to make inferences. On the other hand, the *algebraic* approach that we describe in detail below, employing relation algebra, allows us to manipulate directly the symbolic expressions involving *relations as terms*. In this case, a BSCN may be represented as a set of algebraic equations where the domain of the constants and variables are relations. Formal reasoning with respect to the BSCN is then based on generating solutions to this set of equations using the algebraic *equality* defined over the terms of algebra.

2.2 Composition of Relations as the Basis of Reasoning

We first provide a few basic definitions that permit us to deal with spatial relations as terms that may be manipulated algebraically. We shall view a binary spatial relationship R, in its explicit form, as a set of pairs of spatial objects whose pointset projections are in a given relationship to each other (e.g., one of the 8 relations defined above). Given binary relations R and S, the *composition* of them denoted by $R \circ S$ is

$$R \circ S \equiv \{ <\mathbf{o}_i, \mathbf{o}_k > \mid \exists \mathbf{o}_j \text{ s.t. } <\mathbf{o}_i, \mathbf{o}_j > \in R \text{ and } <\mathbf{o}_j, \mathbf{o}_k > \in S \} \tag{1}$$

or, equivalently using infix notation,

$$\mathbf{o}_i (R \circ S) \mathbf{o}_k \iff \exists \mathbf{o}_j [\mathbf{o}_i (R) \mathbf{o}_j \wedge \mathbf{o}_j (S) \mathbf{o}_k]$$

Consider the case where $\mathbf{o}_i$ and $\mathbf{o}_j$ are related by R_{ij} and $\mathbf{o}_j$ and $\mathbf{o}_k$ are related by R_{jk}. Our focus is on finding the true relation R_{ik} between $\mathbf{o}_i$ and $\mathbf{o}_k$ given an initial assertion R_{ik}^+ and the composition $R_{ij} \circ R_{jk}$. R_{ik}^+ may be viewed as disjunctions that include R_{ik} with another (possibly empty) set of spurious relations R^s, i.e., $\mathbf{o}_i (R_{ik} + R^s) \mathbf{o}_k$. The essential axiom on which we base our system of reasoning is then: Given $<\mathbf{o}_i, \mathbf{o}_j > \in R_{ij}$ and $<\mathbf{o}_j, \mathbf{o}_k > \in R_{jk}$,

$$<\mathbf{o}_i, \mathbf{o}_k > \in R_{ik} \implies <\mathbf{o}_i, \mathbf{o}_k > \in (R_{ik}^+ \cap (R_{ij} \circ R_{jk})) \tag{2}$$

In order to obtain insight into this axiom, let us suppose that $\mathbf{o}_i$ and $\mathbf{o}_k$ are connected by a relation R_{ik}, i.e., $<\mathbf{o}_i, \mathbf{o}_k > \in R_{ik}$. It is clear that $\mathbf{o}_i$, $\mathbf{o}_j$, $\mathbf{o}_k$ form a triangle, and it immediately follows by the definition of (1) that $<\mathbf{o}_i, \mathbf{o}_k > \in R_{ij} \circ R_{jk}$. Further, $<\mathbf{o}_i, \mathbf{o}_k >$ is in R_{ik}^+ since $<\mathbf{o}_i, \mathbf{o}_k > \in R_{ik}$ by assumption and $R_{ik} \subseteq R_{ik}^+$. Now since $<\mathbf{o}_i, \mathbf{o}_k >$ is in both R_{ik}^+ and $(R_{ij} \circ R_{jk})$, it is clearly also in their intersection. We may also rephrase the above axiom in terms of its contraposition: if $<\mathbf{o}_i, \mathbf{o}_k >$ is *not* in the intersection of R_{ik}^+ and $(R_{ij} \circ R_{jk})$ for some j, then it *cannot* be in R_{ik} either. This contraposition is the rationale behind the process of *eliminating* infeasible pairs of objects from R_{ik} (or, from the dual point of view, R_{ik} cannot be in the true relations between the pair of objects), and the elimination is safe since a relation will not be pruned unless *proven* infeasible. In other words, R_{ik} is further constrained *indirectly* by the composition of $R_{ij} \circ R_{jk}$, and the intersection operation guarantees a *monotonic* decrease of R_{ik} viewed as a set of relations. Our goal of reasoning is to find the *least upper bound* of the sets of relations in the whole BSCN network, i.e., to

make the strongest possible assertions on the labels of edges using all the indirect constraints from the composed edges. The so called path-consistency algorithms iteratively impose the above contraposition in order to achieve this goal algebraically.

To manipulate algebraic equations involving the composed terms, the allowed interpretations for the algebraic composition *operator* need to be defined. We will derive, step by step, the algebraic equalities of composition applied to atomic relations as building blocks. The particular definitions should reflect the inherent properties of the set of spatial topological relations and their composition. We will reuse the original symbolism of Tarski (1952) by adopting ";" as the algebraic operator for the computation of the composition, "o".

Consider, for example, the composition of R_{ij}^a and R_{jk}^a which derives R_{ik}^b of Figure 2, i.e.,

$$(\bar{\imath} \ ; o) \stackrel{\text{def}}{=} \bar{\imath} + c + o$$

To show algebraic equality of the above composition, we must show that both the following conditions hold:

$$(\bar{\imath}) \circ (o) \subseteq (\bar{\imath} + c + o) \tag{3}$$

$$(\bar{\imath}) \circ (o) \supseteq (\bar{\imath} + c + o) \tag{4}$$

Condition (3) will be satisfied if

$$\forall \ o_i, o_j, o_k \quad [o_i (\bar{\imath}) o_j \wedge o_j (o) o_k \implies o_i (\bar{\imath}) o_k \vee o_i (c) o_k \vee o_i (o) o_k]$$

The derivation of it is shown graphically in Figure 3, where the bottom row exhausts all possibilities between o_i and o_k modulo topological homomorphism. Condition (4) will be satisfied if

$$\forall \ o_i, o_k \quad [o_i (\bar{\imath}) o_k \vee o_i (c) o_k \vee o_i (o) o_k \implies \exists \ o_j \quad [o_i (\bar{\imath}) o_j \wedge o_j (o) o_k]]$$

Clearly, for any pair of pointsets, $<o_i, o_k>$, which are related by one of the relationships in the bottom row of Figure 3, we can *always* place some o_j such that it satisfies both relations in the top row. (Technically, this is because the topological space of spatial pointsets is of dense order, such as $Q \times Q$, where Q is the set of rational numbers which is countably infinite.)

Similarly, it is relatively straightforward to define the remaining compositions of the 8 atomic relations (AC). We present the results of making all pairwise compositions in Table 2. The element AC_{ij} is the composition of the ith row as the first argument and the jth column as the second argument for $i, j = 1, \ldots, 8$. The set of terms of composition in our algebra, however, are computed from the collection of 2^8 distinct topological relations, and we must, in principle, also define all of the $2^8 \times 2^8$ binary compositions, a seemingly prohibitive task. The semantics of the compositions, fortunately, can be defined *axiomatically* employing the standard theory of relation algebra. We have already defined the compositions among the 8
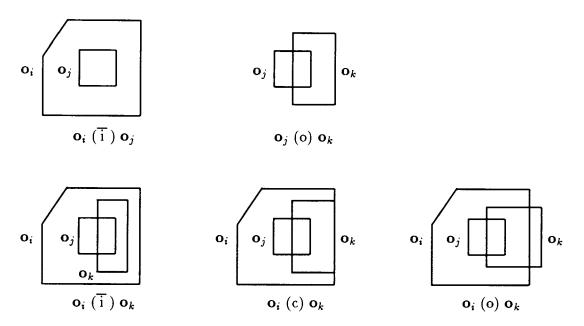
Figure 3: Composition of ($\bar{\text{i}}$) and (o)

atomic relations. Then their *closure*, $\mathcal{A}$, under composition is defined using the distributive law of composition over sum:

$$\forall r_x, r_y, r_z \in \mathcal{A} \quad [(r_x + r_y); r_z = (r_x; r_z) + (r_y; r_z)] \tag{5}$$

With this, the composition of any two sums (i.e., disjunctions) of relations is computable via the sum of pairwise compositions involving only the pairwise composition of *atomic* relations as defined in Table 2.

Note that in the above axiom of the algebraic approach, the spatial relations are directly manipulated as terms (i.e., variables). An axiomatization of the composition of relations in the predicate calculus, in contrast, would be complicated, since the relations are represented as predicate symbols and the terms in the calculus are spatial objects (with pointset projections) rather than relations between objects. As such, a direct translation of the above axiom into predicate calculus results in high-order formulae in which the predicate symbols themselves (i.e., relations) become terms and variables. To avoid the undesirable complications associated with high-order logic, the complete first-order axiomatic system must inevitably enumerate axioms for each composition in which the total number of axioms for composition is exponential in the number of underlying atomic relations. Hence our main reason for adopting a algebraic approach rather than a logical one in this paper is based on an appeal to simplicity, in which the closure of compositions may be parsimoniously defined, and in which the inferential procedures may also be specified with great simplicity.

	1'	d	m	o	c	i	c̄	T̄
1'	1'	d	m	o	c	i	c̄	T̄
d	d	1	d m o i c̄	d m o i c̄	d	d m o i c̄	d m o i c̄	d
m	m	d m o c T̄	1' d m o c c̄	d m o i c̄	d m	o i c̄	m o i c̄	d
o	o	d m o c T̄	d m o c T̄	1	d m o c T̄	o i c̄	o i c̄	d m o c T̄
c	c	d m o c T̄	m o c T̄	o c T̄	c T̄	o i c̄	1' o c c̄	T̄
i	i	d	d	d m o i c̄	d m o i c̄	i	i	1
c̄	c̄	d	d m	d m o i c̄	1' d m o c c̄	i	i c̄	d m o c T̄
T̄	T̄	d m o c T̄	d m	o c T̄	T̄	1' o c i c̄ T̄	o c T̄	T̄

Table 2: Atomic Compositions in Algebra of Spatial Topological Relations

3 Implementation of Spatial Inference

3.1 Constraint Inference via Path Consistency Procedures

We have represented a BSCN in terms of relation algebra, with an associated inference procedure that involves finding the least upper bound of the set of the relations R_{ij} for all i, j. That is, the partially or unspecified (universal relation) edges in the BSCN is monotonically reduced to its least upper bound by *eliminating* those relations which are *not implied* by the axiom of relation composition. This axiom for reduction is algebraically

$$R_{ij}^{new} \leftarrow R_{ij} \cdot (R_{ik}; R_{kj})$$

Path-consistency computations, applied to each of the triangles of the BSCN edges, will effectively produce the inferences that we desire.

We now describe the data structures for a procedural implementation of spatial inferences on a database consisting of n spatial objects and their topological relations. To simplify the illustration of the idea behind the procedure, we first extend the BSCN abstraction of our database of objects to a *strongly-connected* directed graph, such that there exists an edge for every pair for each direction (for example, if there is no edge, we add an edge labeled with the universal relation "1"). Second, we map this extended BSCN into a square matrix data structure, M of dimension n on which we carry out the path consistency procedure. We will call this matrix M a *relation matrix*, and the entry M_{ij} contains the relation R_{ij} between the object o_i and o_j. The invariant properties and conventions of this representational scheme include:

- R_{ij} asserted in the database initializes M_{ij}.

- M_{ii} is the identity relation, $1'$.

- M_{ij} is the converse of M_{ji}, and *vice versa*.

- If there is no relation asserted between o_i and o_j where $i \neq j$, then M_{ij} is initialized by the universal relation, 1.

We now present an algorithm schema for a path-consistent reduction of BSCN.

$\quad$ **Procedure** *mostGeneralPathConsistency*
$\quad$ **Input** $M_{n \times n}$ *initial relation matrix*
$\quad$ **Output** $M_{n \times n}^{new}$ *inferred relation matrix*
$\quad$ begin
(1) $\qquad M^{new} \Leftarrow M$
(2) $\qquad$ repeat
(3) $\qquad\qquad M \Leftarrow M^{new}$
(4) $\qquad\qquad$ for $i = 1$ to n do
(5) $\qquad\qquad\qquad$ for $j = 1$ to n do
(6) $\qquad\qquad\qquad\qquad$ for $k = 1$ to n do
(7) $\qquad\qquad\qquad\qquad\qquad M_{ij}^{new} \leftarrow M_{ij} \cdot (M_{ik}; M_{kj})$
(8) $\qquad\qquad\qquad\qquad\qquad$ if $M_{ij}^{new} = 0$ then
(9) $\qquad\qquad\qquad\qquad\qquad\qquad$ report inconsistency among (i, j, k)
(10) $\qquad\qquad\qquad\qquad\qquad$ endif
(11) $\qquad\qquad$ until $M \leq M^{new}$
$\quad\quad$ end

The operator $\Leftarrow$ in line (1) and (3) denotes the element-by-element copy of matrices; lines (4) through (6) represents iteration of the path consistent reduction on every combination

of triangles in the given BSCN; line (7) is the key step in the procedure and implements the essential axiom of inference described earlier. An entry of M becomes the empty relation "0", which may occur during the intersection operation, if the given BSCN contains inconsistencies for the triangle i, j, k. For instance, an assertion, "o_i covers o_k, and o_j is inside of o_k, and o_i either meets or overlaps with o_j ", is topologically inconsistent and thus results in the empty set at step (7). The procedure may then be forced to terminate at this point in (9), since an inconsistent BSCN does not have any *model*.

The symbol, $\leq$, appearing in the termination condition at (11) represents the partial order defined by the subset relation *on relations*. This condition will be met when there are no more reductions in M, i.e., at the least upper bound. Taking matrices as arguments,

$$ M \leq M^{new} \equiv \forall_{i,j} \ M_{ij} \subseteq M_{ij}^{new} \equiv \forall_{i,j} \ (M_{ij} \cdot M_{ij}^{new} = M_{ij}) $$

We illustrate this procedure in detail with a simple example of a 4 node BSCN of spatial objects, o_1 through o_4, in Figure 4. The relation matrices for the associated strongly connected BSCN are shown in Table 3.
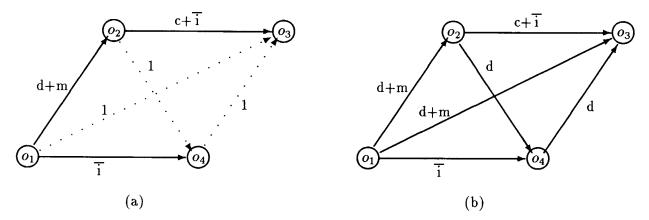

Figure 4: Algebraic Reduction of Spatial Constraints for 4-Node BSCN

M^{new} in Table 3 shows the result of the minimal reduction of M at the termination of the procedure, containing inferences for the three initially unknown relations. A detailed trace of computation on the edge M_{13} and M_{24} is given below where the iterations of "·" are concatenated into a single expression:

$$
\begin{aligned}
M_{13}^{new} &= (M_{11}\,;M_{13}\,) \cdot (M_{12}\,;M_{23}\,) \cdot (M_{13}\,;M_{33}\,) \cdot (M_{14}\,;M_{43}\,) \\
&= (1';1) \cdot ((d+m);(c+\overline{1}\,)) \cdot (1;1') \cdot (1';1) \\
&= 1 \cdot ((d;c) + (d;i) + (m;c) + (m;\overline{1}\,)) \cdot 1 \cdot 1 \\
&= 1 \cdot ((d) + (d) + (d+m) + (d)) \cdot 1 \cdot 1 \\
&= 1 \cdot (d+m) \cdot 1 \cdot 1 \\
&= (d+m) \\
M_{24}^{new} &= (M_{21}\,;M_{14}\,) \cdot (M_{22}\,;M_{24}\,) \cdot (M_{23}\,;M_{34}\,) \cdot (M_{24}\,;M_{44}\,)
\end{aligned}
$$

	o_1	o_2	o_3	o_4
o_1	$1'$	d m	1	$\overline{1}$
o_2	d m	$1'$	c $\overline{1}$	1
o_3	1	$\overline{c}$ i	$1'$	1
o_4	i	1	1	$1'$

M

	o_1	o_2	o_3	o_4
o_1	$1'$	d m	d m	$\overline{1}$
o_2	d m	$1'$	c $\overline{1}$	d
o_3	d m	$\overline{c}$ i	$1'$	d
o_4	i	d	d	$1'$

M^{new}

Table 3: Relation Matrices of Figure 4

$$
\begin{aligned}
&= ((d + m); (\overline{1}\ )) \cdot (1'; 1) \cdot ((c + \overline{1}\ ); 1) \cdot (1; 1') \\
&= ((d; \overline{1}\ ) + (m; \overline{1}\ )) \cdot 1 \cdot 1 \cdot 1 \\
&= (d + d) \cdot 1 \cdot 1 \cdot 1 \\
&= d
\end{aligned}
$$

To this point, we have merely presented a *schema* for pruning constraints in order to provide intuition into the class of procedures for enforcing path-consistency, solving a simultaneous set of algebraic equations of constraints. The schema does not involve optimization in terms of its space and time complexity, since the "blind" iterations over every triangle results in a worst case time complexity of $O(n^5)$. Algorithms employing heuristic backtracking and a queue of triangles and having a worst case time complexity of $O(n^3)$ are, however, known (Mackworth and Freuder 1985). In terms of storage, the scheme with slight modifications will generate the same reduction using a triangular matrix of size of $M/2$, although the space complexity is still $O(n^2)$. A discussion of these computationally efficient variants of path-consistency procedure are beyond the scope of the present paper, and readers are referred to the literature for the well established algorithms.

Our primary concern in relation to the path consistency procedures is with the *logical* notions of *soundness* and *completeness* of procedural reasoning scheme. The example of Figure 4 involves a relatively trivial case of inference, in the sense that the same inferences may be made by hand in a few minutes. With BSCNs containing a dozen or more nodes, such an approach becomes infeasible and the automated approach that we suggest becomes indispensable, since it is both logically sound and complete. The termination of the procedure is guaranteeded because of the monotonic reduction in line (7), and any inconsistencies in the initial BSCN are detected in (8). With the proof of its *correctness* (Montanari 1974), the procedure satisfies the soundness requirement. This follows from the fact that the composition of two connected edges in a triangle forms a path of length 2 and from the theorem of Montanari (1974) that if every path of length 2 of a complete underlying graph of a network

is path-consistent, then the network is path-consistent. In other words, the correct solutions with respect to only triangles are also the logical consequences (i.e., sound inferences) with respect to the entire set of assertions in BSCN.

The *exact* reductions, namely a BSCN connected by edges labeled with single atomic relations, are theorems of the algebraic model. Once the reduction of the BSCN is complete, backtracking search through alternative atomic relations among disjunctive labels will achieve an exact reduction. In essence, the procedure checks for path-consistency in each combination of atomic relations chosen from the (reduced) disjunction on each edge of the BSCN (we refer the reader to Malik and Binford, 1983, for details of the algorithm) The exact reduction of the path-consistent network is an NP-hard problem in general (Maddux 1989), but NP-hardness is a computational issue of tractability and not a logical issue. In other words, the path consistency procedure will never prune correct relations and it will eventually find *all* the correct ones, and as such, is a logically complete reasoning procedure for BSCN.

3.2 An Efficient Implementation of Algebras of Relations

In the preceding section, we were concerned with the implementation of the inference procedure. We now focus on an efficient implementation of the relation algebra itself, and in particular, for cases in which the number of atomic spatial relationships is relatively small. A computationally efficient implementation of relation algebra and its operators is critical since they appear in the innermost loop of inference procedures.

We first describe a bit-field encoding scheme for the *terms* of our algebra, namely the spatial relations. Our algebra is *finite* since it involves the finite closure (2^8 relations) of the set of 8 atomic relations under composition and converse, including the empty and universal relation. It is, therefore, possible that a single 8-bit byte will uniquely "code" any one of the 2^8 distinct terms of our algebra. That is, each atomic relation may be associated with a distinct bit field within a byte, and the non-atomic (i.e., disjunctive) relations may then be encoded in a single byte, with appropriate bits set to 1, and the remaining bits to 0. This encoding scheme provides a compact and uniform representation of relations, as well as an efficient implementation of the set-theoretic operators, "+", "·" and the converse "$^-$" of relation algebra. The 8-bit code is clearly the most compact data structure, and the entries of the composition table (and relation matrices) will now be uniformly stored in single bytes. The union operation of relations may now be implemented as bit-wise OR operations which we denote by $\vee_b$. Similarly, the intersection of the algebra is replaced by bit-wise AND operations, $\wedge_b$. We have effectively replaced the set-theoretic operators of our algebra by the single machine instructions available in programming languages such as C. The converse "$^-$" is simply a bit-masking operation using $\vee_b$ and $\wedge_b$, where the necessary bits (in our case, i, c and their converses) are toggled.

We now describe the implementation of the composition operator, ";". Recall that the computation of compositions of non-atomic relations is based on the 8 × 8 atomic composi-

tion table (AC), since composition distributes over the sum in relation algebra. If we build a complete table of compositions (CC) among the entire set of relations, we may avoid the necessity of distributing compositions over sums, looking up each pairwise atomic composition in AC and summing the result. That is, a composition of any term of our algebra may be computed by a *single* lookup of the complete table CC of 256×256. This table, using the bit-field encoding system, involves a relatively small amount (64K) of main memory. Note also that the size of CC is independent of the size of the BSCNs.

This coding scheme may be generalized in a straightforward manner for algebras involving larger (but finite) numbers of atomic relations. A typical 4-byte integer, for example, permits the encoding of an algebra with 32 atomic relations, and the concatenation of machine words permits even larger codes. It is clear, however, that the space complexity is of exponential order in the number of atomic relations, and the practicality of constructing a full-sized CC diminishes rapidly as this number grows.

The following algorithm "systematically" computes the CC for the general case where the set of atomic relations has cardinality of N, which in our algebra is 8.

Procedure *constructCompleteCompositionTable*
Input $AC_{N \times N}$ *atomic composition table*
Output $CC_{2^N \times 2^N}$ *complete composition table*
begin
(1)   for $i = 1$ to 2^N do
(2)   for $j = 1$ to 2^N do
(3) $r \leftarrow 0$
(4)   for $k = 1$ to N do
(5)   if $((2^k \wedge_b i) = 2^k)$ then
(6)   for $l = 1$ to N do
(7)   if $((2^l \wedge_b j) = 2^l)$ then
(8) $r \leftarrow r \vee_b AC_{kl}$
(9)   endif
(10)   endif
(11) $CC_{ij} \leftarrow r$
    end

Lines (3) through (10) construct the code "r", for the composition of CC_i and CC_j; lines (5) and (7) perform pairwise computation, where 2^k and 2^l is the predetermined bit position for each of the N atomic relations. The occurrences of the atomic pairs between the disjunctive relations CC_i and CC_j are checked by the bit-masking operations, and the composition for the pair is looked up using AC, and then summed by $\vee_b$ in line (8).

One advantage of the encoding scheme for relations with a CC constructed as above is that the code of the relations may be used *directly* in indexing the CC table, since our coding scheme is also an integer system (independent of its byte order) enumerated from 0 to 255.

We now reconsider the essential step in computing any type of algebraic inference with a BSCN abstraction, namely

$$M_{ij}^{new} \leftarrow M_{ij} \cdot (M_{ik} ; M_{kj})$$

With the bit-field encoding system and the complete composition table CC, we may rewrite this step as:

$$M_{ij}^{new} \leftarrow M_{ij} \wedge_b CC_{M_{ik}M_{kj}}$$

Hence relation algebraic operators ($\cdot$ and ;) are *literally translated* by $\vee_b$ and a single table lookup, since the relation algebra and its operators are subsumed by Boolean bit operations. This is not surprising since relation algebra is an extension of Boolean algebra. The single atom "1" in the Boolean algebra is extended to a number of atoms (i.e., the atomic relations N). The disjunction "+" and conjunction "·" of relation algebra is nothing but the Boolean algebra performing "+" and "·" *in parallel* for each bit position, where the total number of bit positions is the same as the number of atomic relations under consideration.

4   Conclusion

Our discussion of the soundness and completeness of a spatial reasoning procedure is based on the assumption that our abstraction of the world and the task of reasoning are expressible in terms of a formal language, in which the domain specific theories are axiomatizable and in which procedural semantics may be given to each expression. As an example of our approach, we have described an algebra in which we may represent topological relations among spatial objects and the composition of such relationships. The particular set of spatial relations that may be represented in this simple algebra is based upon a set of 8 atomic topological relations that Egenhofer and Franzosa (1991) defined on spatial pointsets. A non-trivial set of spatial reasoning problems are solvable within this algebraic framework, which takes the set of topological relations as the *terms* of an algebraic equation system and applies path consistency procedures.

We employed a binary spatial constraint network (BSCN) as an abstraction for the database of spatial knowledge. The BSCN has both a declarative and procedural semantics. Reasoning via the algebra of spatial relations within the context of constraint satisfaction procedures permits us to provide a sound and complete procedural semantics for a given spatial database represented in terms of a BSCN. We provided an implementation of both the theory and the operators of the algebra in terms of a table lookup scheme and bit-field encodings.

While the spatial reasoning system that we have described is intended to be purely exemplary in nature, it nevertheless provides an indication of the general class of systems that one might develop with a more expressive set of spatial relations. We believe that the reasoning systems similar to ours, in which qualitative spatial relations are a primitive data type, may well play an important role that is complementary to the role of systems that are based on numerical computations which preserve the metric properties of the object

space involving, for example, distance and direction. We believe that there is a fruitful area of research in which the prime objectives of investigation involve finding alternative sets of spatial relationships and more general inference procedures. We suggest that it may be of great value to incorporate such reasoning functionality into a variety of computational systems that support spatial analysis and modeling, such as GIS and spatial database and modeling systems in general.

References

[1] ALLEN, J.F., 1983, Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, **26**, pp. 832–843.

[2] CHANG, S.K., JUNGERT, E. AND Y. LI, 1990, The Design of Pictorial Databases Based upon the Theory of Symbolic Projections. *in Design and Implementation of Large Spatial Databases, ed A.Buchmann, O.Gunther, T.R.Smith and Y-F.Wang,* Lecture Notes in Computer Science, Springer-Verlag, Berlin.

[3] EGENHOFER, M.J., AND FRANZOSA, R.D.,1991, Point-set Topological Spatial Relations. *International Journal of Geographical Information Systems*, **5**:2, pp. 161–174.

[4] ENDERTON, H.B., 1972, *A Mathematical Introduction to Logic*, Academic Press, New York, NY.

[5] FRANK, A., 1991, Qualitative Spatial Reasoning about Distances and Directions in Geographic Space. *Technicial Report, NCGIA, Orono, Maine.*

[6] GENESERETH, M.R., AND NILSSON, N.,1987, *Logical Foundations of Artificial Intelligence*, Morgan Kauffman Publishers, Inc., Palo Alto, Ca.

[7] GOPAL, S. AND SMITH, T.R., 1990, Human Way-finding in an Urban Environment:a Performance Analysis of a Computational Process Model. *Environment and Planning*, **22**, pp. 169-191.

[8] GUESGEN, H.W., 1989, Spatial Reasoning based on Allen's Temporal Logic. Technical Report TR-89-049, International Computer Science Institute, Berkeley, Ca.

[9] JONSSON, B., AND TARSKI, A.,, 1952, Boolean Algebras with Operators II. *American Journal of Mathematics*, pp. 127–162.

[10] KUIPERS, B., 1978, Modelling Spatial Knowledge. *Cognitive Science*, **2**, pp. 129-153.

[11] LADKIN, P.B., 1990, Constraint Reasoning with Intervals: A tutorial, Survey, and Bibliography. Technical Report, International Computer Science Institute, Berkeley, Ca.

[12] MACKWORTH, A., AND FREUDER, E., 1985, The Complexity of Some Polynomial Network Consistency Algorithms for Constraint Satisfaction Problems. *Artificial Intelligence*, **25**, pp. 65-73.

[13] MADDUX, R.D., 1989, Some Algebras and Algorithms for Reasoning about Time and Space. Internal paper, Department of Mathematics, Iowa State University, Ames, Iowa.

[14] MALIK, J., AND BINFORD, T.O., 1983, Reasoning in Time and Space. In *International Joint Conference on Artificial Intelligence*, pp. 343-345.

[15] MONTANARI, U., 1974, Network of Constraints. *Information Science*, **7**, pp. 95-132.

[16] MUNKRES, J.R., 1975, *Topology: A First Course.* Prentice-Hall, Englewood Cliffs, NJ.

[17] SMITH, T.R., RAMAKRISHNAN, R. AND VOISARD, A., 1991, A Data Model and a Deductive Language for Spatio-temporal Database Applications. Procedings of the Basic GOODS Workshop on Spatial Databases, Capri, Italy. Lecture Notes in Computer Science, Springer Verlag.

[18] SMITH, T.R., PEUQUET, D., MENON, S., AND AGARWAL, P., 1987, KBGIS-II: A Knowledge-based Geographical Information System. *International Journal of Geographical Information Systems*, **1**, pp. 149–172.

[19] ULLMAN, J.D., 1989, *Principles of Database and Knowledge-base systems*, Computer Science Press, Rockville, Md.

Appendix: Relation Algebra

Material in this appendix has appeared in Ladkin (1990) and elsewhere. Relation algebra is an extension of Boolean algebra, which may be viewed as a 6-tuple $< \mathcal{B}, +, \cdot, -, 0, 1 >$, where $\mathcal{B}$ is a set (equations in algebraic theory) and $+, \cdot, -, 0,$ and 1 have the standard meanings. A standard collection of axioms for Boolean algebra includes equations such as:

$$x + x = x$$

$$x \cdot x = x$$

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z)$$

A standard partial ordering for $x, y \in \mathcal{B}$ may be defined by:

$$x \leq y \Leftrightarrow x \cdot y = x$$

An *atom* in a Boolean algebra is a smallest non-zero element in the algebra with respect to this ordering $\leq$. The additional binary operators and a constant which make a Boolean algebra into a Relation algebra are described below using the notation of Josson and Tarski(1952) :

;   binary operator for composition of relations

¯   unary operator for forming the converse of a given relation

$1'$ constant denoting the identity relation

These operators in relation algebra operate on *relations,* thus the topological relations are the objects of the algebra in our case. Note that every object in this algebra is the union of the singleton sets, i.e., 8 atomic relations in our case. "0" represents the empty relation and "1" represent the universal relation, i.e., the union of all relations. A relation algebra $\mathcal{R}$, then, is a structure $< \mathcal{R}, +, \cdot, -, 0, ;, 1', ^{-} >$ that satisfies following additional axioms, where $x, y, z \in \mathcal{R}$:

$$(x; y); z = x; (y; z)$$

$$(x + y); z = x; z + y; z$$

$$x; 1' = x$$

$$\overline{(\overline{x})} = x$$

$$\overline{(x + y)} = \overline{x} + \overline{y}$$

$$\overline{(x; y)} = \overline{y}; \overline{x}$$

$$\overline{x}; -(x; y) + -y = -y$$