Definitions of Line-Line Relations for Geographic Databases*

Max J. Egenhofer
National Center for Geographic Information and Analysis
and
Department of Surveying Engineering
Department of Computer Science
University of Maine
Orono, ME 04469-5711
max@mecan1.maine.edu

Abstract

Query languages for spatial databases need appropriate tools to inquire about spatial data and provide access to the relations among spatial objects. These spatial relations are more complex than conventional predicates comparing equality or order. Examples are such spatial predicates as "neighbor," "intersect," and "inside." A formal definition of spatial relations is necessary to define the semantics of an appropriate set of spatial predicates in query languages and to provide a basis for spatial query processing. We have extended a model, initially designed for binary topological relations between 2-dimensional objects, to treat 1-dimensional objects in $\mathbb{R}^2$ as well. The approach used is based upon algebraic topology and compares the interiors, boundaries, and exteriors of the lines. A total of 33 different topological relations between two simple lines has been identified formally, for which geometric interpretations are given.

1 Introduction

Queries in Geographic Information Systems (GISs) and image databases are often based on the relations among spatial objects. For example, in geographic applications typical spatial queries are, "Retrieve all roads that lead to interstate highway I-95" and "Find all electric power lines that run across a river." Current database query languages, such as SQL and Quel, do not sufficiently support such queries, because they provide only tools for comparing equality or order of such simple data types as integers or strings. The incorporation of spatial relations over geometric domains into a spatial query language has been identified as an essential extension beyond the power of traditional query languages [5, 17]. Some experimental spatial query languages support queries with one or the other spatial relation; however, the diversity, semantics, completeness, and terminology vary dramatically [4, 10]. With the advent of extensible database query languages such as SQL3, it will become increasingly important to have (consistent) models and formalizations of relations.

Previous investigations developed a formal model for *binary topological relations* for co-dimension 0 (i.e., if the dimension between the embedding space and the objects is 0). It applies to relations between regions (2-dimensional objects) embedded in $\mathbb{R}^2$ [7] or lines in $\mathbb{R}^1$ [16]. For these settings, we have promoted a comprehensive formalism [6], which generalizes to n-dimensional objects embedded in $\mathbb{R}^n$. Others have used this model, for instance to formalize topological relations among spatial objects in 3-D [11, 15], studies of the use of spatial predicates in natural languages [14], or as a

---

[*] This research was partially funded by NSF grant No. IRI-9309230 and grants from Intergraph Corporation. Additional support from NSF for the NCGIA under No. SES-880917 is gratefully acknowledged.

basis for the integration of topological constraints into spatial query languages [3, 13, 18]. Here, we extend the formalizations of topological relations to line-line relations in $\mathbb{R}^2$.

The remainder of this paper is organized as follows: Section 2 briefly summarizes the model used for topological relations with co-dimension 0 and demonstrates why it is necessary to extend it for objects with co-dimension 1, such as lines in $\mathbb{R}^2$. Section 3 introduces the 9-intersection as our model to formalize binary topological relations between lines in $\mathbb{R}^2$. Section 4 investigates which relations can be realized between two lines in $\mathbb{R}^2$. The conclusions in Section 5 discuss future research activities.

2 4-Intersection

Binary topological relations between two objects, A and B, are defined in terms of the four intersections of A's boundary (∂A) and interior (A°) with the boundary (∂B) and interior (B°) of B [6], called the *4-intersection* [6]. Topological invariants of these four intersections, i.e., properties that are preserved under topological transformations, are used to categorize topological relations. Examples of topological invariants applicable to the 4-intersection are the content (i.e., emptiness or non-emptiness) of a set, the dimension, and the number of separations [9]. The content invariant is the most general criterion as any other invariants can be considered refinements of non-empty intersections. By considering the values empty ($\emptyset$) and non-empty ($\neg\emptyset$) for the four intersections, one can distinguish sixteen binary topological relations, nine of which can be realized for 2-dimensional objects (including objects with holes), called *regions*, if the objects are embedded in $\mathbb{R}^2$ [7], and eight between two *lines* in $\mathbb{R}^1$ (Fig. 1) (which correspond to Allen's interval relations [2] if the order of the 1-dimensional space is disregarded) . The difference is due to the fact that regions may have connected boundaries, while lines have disconnected boundaries.

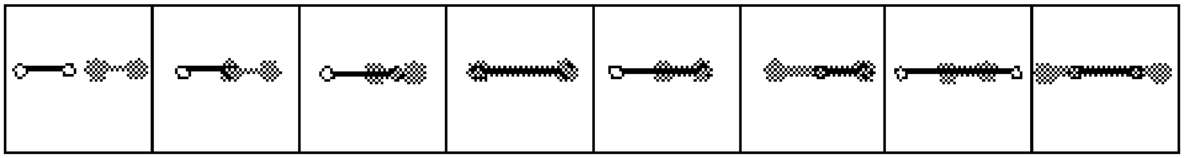

Figure 1: Examples of the eight topological relations between two lines in $\mathbb{R}^1$.

The same relations also exist when the objects are mapped into a higher-dimensional space; however, due to the greater degree of freedom, the objects may take configurations that are not represented by one of the relations between objects with co-dimension 0. For example, if two lines "cross," they have non-empty interior-interior intersections, while the other three intersections are empty. Such a 4-intersection could not be realized for two lines in $\mathbb{R}^1$. On the other hand, some 4-intersections may have ambiguous geometric interpretations (Fig. 2). From a practical point of view, there are certainly situations in which one would like to distinguish between them when querying

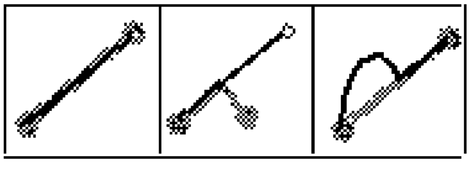

Figure 2: Examples of topological relations between two lines in $\mathbb{R}^2$ that have the same 4-intersection ($A^\circ \cap B^\circ = \neg\emptyset$; $A^\circ \cap \partial B = \emptyset$; $\partial A \cap B^\circ = \emptyset$; and $\partial A \cap \partial B = \neg\emptyset$).

a geographic database. These observations motivated the development of an extended model to account also for relations between n-dimensional objects that are embedded in an m-dimensional space, $m > n$.

3 9-Intersection for Binary Topological Relations between Lines

The definition of a line is based on 1-cells, i.e., the connections between two geometrically independent nodes. A line is a sequence of $1 \ldots n$ connected 1-cells such that they neither cross themselves nor form a cycle. Nodes at which exactly one 1-cell ends will be referred to as the *boundary* of the line. Nodes that are an endpoint of more than one 1-cell are *interior nodes*. The *interior* of a line is the union of all interior nodes and all connections between the nodes. The *closure* of a line is the union of its interior and boundary. Finally, the *exterior* is the difference between the embedding space and the closure of the lines. We will call a sequence of 1-cells a *simple line* if it has exactly two boundary nodes. Lines that would have less than two boundary nodes would include cycles, which are excluded by definition.

The 4-intersection is extended by considering the location of each interior and boundary with respect to the other objects exterior; therefore, the binary topological relation R between two lines, A and B, in $\mathbb{R}^2$ is based upon the comparison of A's interior (A°), boundary (∂A), and exterior (A^-) with B's interior (B°), boundary (∂B), and exterior (B^-). These six object parts can be combined such that they form nine fundamental descriptions of a topological relation between two lines and be concisely represented by a 3×3-matrix, called the *9-intersection*.

$$R(A,B) \quad = \quad \begin{pmatrix} A^\circ \cap B^\circ & A^\circ \cap \partial B & A^\circ \cap B^- \\ \partial A \cap B^\circ & \partial A \cap \partial B & \partial A \cap B^- \\ A^- \cap B^\circ & A^- \cap \partial B & A^- \cap B^- \end{pmatrix}$$

Each intersection will be characterized by a value *empty* ($\emptyset$) or *non-empty* ($\neg\emptyset$), which allows one to distinguish $2^9 = 512$ different configurations. Only a small subset of them can be realized between two lines in $\mathbb{R}^2$.

4 Existing 9-Intersections Between Lines in $\mathbb{R}^2$

In order to identify which of the 512 different 9-intersections may be realized between two lines in $\mathbb{R}^2$, we formalize a set of properties as conditions for binary topological line-line relations, that must hold between the parts of any two lines. These properties can be expressed as consistency constraints in terms of the 9-intersection [8], such that by successively eliminating from the set of 512 relations the relations that would violate a consistency constraint, one retains the candidates for those 9-intersections that can be realized for the particular spatial data model. The existence of these relations is then proven by finding geometric interpretations for the corresponding 9-intersections.

Condition 1: The exteriors of two lines always intersect with each other.

Condition 2: A's boundary intersects with at least one part (interior, boundary, or exterior) of B, and vice-versa.

Condition 3: Each boundary of a simple line intersects with at most two parts of another line.

Condition 4: If A's boundary is a subset of B's boundary, then the two boundaries coincide, and vice-versa.

Condition 5: If A's interior does not intersect with B's exterior then the two interiors must intersect as well, and vice-versa.

Condition 6: If A's interior does not intersect with B's exterior then A's boundary must not intersect with B's exterior, and vice-versa.

Condition 7: If A's interior does not intersect with B's exterior then A's interior must not intersect with B's boundary, and vice-versa.

Condition 8: If A's closure is a subset of B's interior then either A's exterior intersects with both B's boundary and B's interior, or not at all, and vice-versa.

Based on these conditions, we find 33 relations between two simple lines, 13 of which are symmetric (examples of geometric interpretations are shown in the top two rows of Fig. 3) and the remaining ones form 10 pairs of converse relations (bottom two rows of Fig. 3 show one example of each pair of the converse relations).

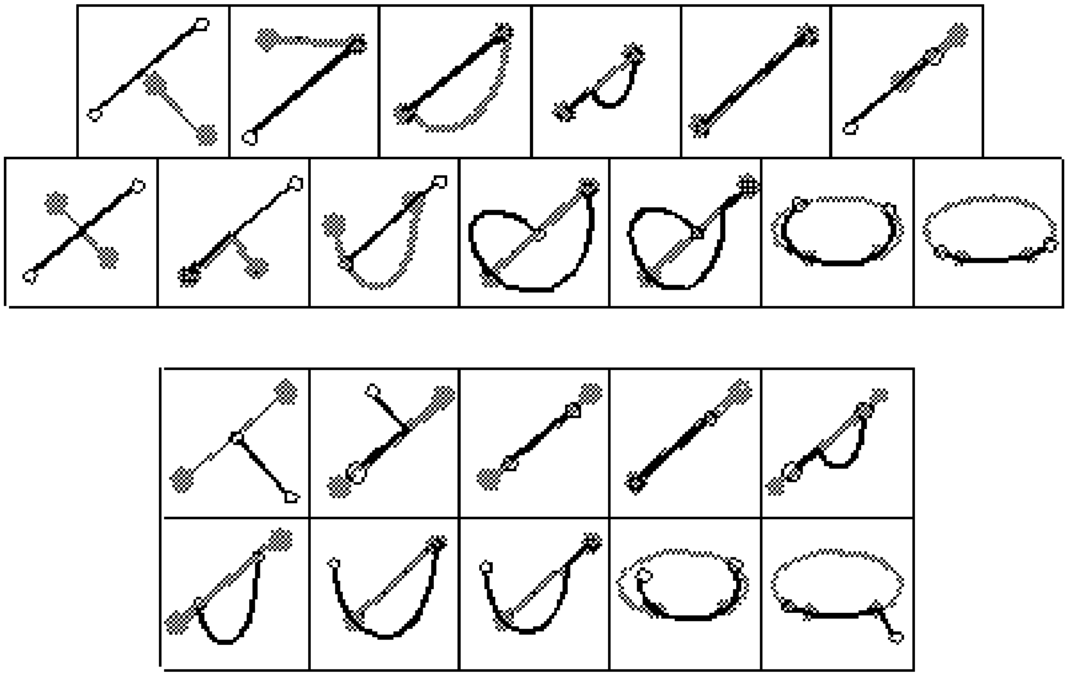

Figure 3: Geometric interpretations of the topological relations between two simple lines in $\mathbb{R}^2$.

5 Conclusions

A formalism for the definition of binary topological relations between linear spatial objects embedded into $\mathbb{R}^2$ has been presented. Binary topological relations are described by the 9-intersection, i.e., the set intersections between the interiors, boundaries, and exteriors of the two lines. The criterion for distinguishing different topological relations is the *content* of the 9-intersections, i.e., whether the intersections are empty or non-empty. Using these criteria to classify topological relations, it was shown that in $\mathbb{R}^2$ there are only 33 line-line relations that can be realized. This set of relations provides a solid framework to process spatial queries. While the content invariant provided a very generic criterion for classifying line-line relations, there are a number of other topological invariants that reveal more detailed differences about line relations. Examples of these invariants are the number of non-empty interior-interior intersections, their dimension, their type (crossing or touching), and the order in which these invariants are encountered [3, 9, 12]. Each of them increases the number of spatial relations a user can choose from when querying a geographic database.

The results of this paper have some serious implications for the design of spatial query languages. Obviously, there is a large number of different relations that users might want to distinguish. Although one may be tempted to group them into categories of conceptually similar relations [3], human subject tests with line-region relations have indicated that some users actually distinguish the entire array of different relations offered by the 9-intersection model [14]. The diversity of spatial configurations that may be distinguished and sometimes necessary for users to make decisions, may be greater than what can be easily described in a traditional database query language. For example, natural (English) language offers only a small, limited amount of spatial predicates to describe spatial configurations [19]. Without comprehensive tests of how humans judge spatial relations in different situations in which the context or the semantics of the objects involved change, any decisions about the use of vocabulary in spatial query languages will be speculative.

As a consequence for spatial query language design, detailed spatial relations may need to be described by other means than predicates, e.g., graphically as a sketch, as annotated drawing, or by selecting combinations of legal and illegal configurations from a set of given prototypes. Graphical renderings appear to be very natural in interacting with spatial data, however, there are some major problem one faces with specifying spatial constraints as sketches. For example, complex Boolean combinations of spatial constraints—particularly disjunctions and negation—are very difficult to draw. Likewise, any graphical rendering contains inevitably more spatial constraints than what is expressed by a single spatial predicate. For example, a sketch of a topological relation that A is disjoint from B carries additional information about the directions between the objects, their shapes, relative sizes and distances. Query languages that incorporate an *interview process* during which the user is involved in resolving overspecified or underspecified constraints, may be promising approaches to overcome the limitations of traditional query languages based on the question-answer paradigm.

References

[1] P. Alexandroff, *Elementary Concepts of Topology.* New York, NY: Dover Publications, Inc., 1961.

[2] J. Allen, "Maintaining knowledge about temporal intervals," *Communications of the ACM*, vol. 26, no. 11, pp. 832–843, Nov. 1983.

[3] E. Clementini, P. Di Felice, and P. van Oosterom, "A small set of formal topological relationships suitable for end-user interaction," in *Third International Symposium on Large Spatial Databases, SSD '93,* (Singapore), pp. 277–295, Lecture Notes in Computer Science, vol. 692, New York, NY: Springer-Verlag, June 1993.

[4] M. Egenhofer, *Spatial Query Languages.* PhD thesis, University of Maine, Orono, ME, 1989.

[5] M. Egenhofer and A. Frank, "Towards a spatial query language: user interface considerations," in *14th International Conference on Very Large Data Bases*, (D. DeWitt and F. Bancilhon, eds.), (Los Angeles, CA), pp. 124–133, Aug. 1988.

[6] M. Egenhofer and R. Franzosa, "Point-set topological spatial relations," *International Journal of Geographical Information Systems*, vol. 5, no. 2, pp. 161–174, 1991.

[7] M. Egenhofer and J. Herring, "A mathematical framework for the definition of topological relationships," in *Fourth International Symposium on Spatial Data Handling,* (K. Brassel and H. Kishimoto, eds.), (Zurich, Switzerland), pp. 803–813, July 1990.

[8] M. Egenhofer and J. Sharma, "Topological relations between regions in R^2 and Z^2," in *Third International Symposium on Large Spatial Databases, SSD '93,* (Singapore), pp. 316-336, Lecture Notes in Computer Science, vol. 692, New York, NY: Springer-Verlag, June 1993.

[9] R. Franzosa and M. Egenhofer, "Topological spatial relations based on components and dimensions of set intersections," in *International Society for Optical Engineering, SPIE's OE/Technology '92Vision Geometry,* (R. Melter and A. Wu, eds.), (Boston, MA), vol. 1832, pp. 236–246, Nov. 1992.

[10] O. Guenther and A. Buchmann, "Research issues in spatial databases," *SIGMOD RECORD*, vol. 19, no. 4, pp. 61–68, 1990.

[11] N.W. Hazelton, L. Bennett, J. Masel, "Topological structures for 4-dimensional geographic information systems," *Computers, Environment, and Urban Systems*, vol. 16, no. 3, pp. 227–237, 1992.

[12] J. Herring, "The mathematical modeling of spatial and non-spatial information in geographic information systems," in *Cognitive and Linguistic Aspects of Geographic Space*, (D. Mark and A. Frank, eds.), pp. 313–350, Dordrecht: Kluwer Academic Publishers, 1991.

[13] S. de Hoop and P. van Oosterom, "Storage and manipulation of topology in Postgres," in *Third European Conference on Geographical Information Systems, EGIS '92*, (Munich, Germany), pp. 1324–1336, Apr. 1992.

[14] D. Mark, "Cognitive image schemata and geographic information: relation to user views and GIS interfaces," in *GIS/LIS '89*, (Orlando, FL), pp. 551–560, Nov. 1989.

[15] S. Pigot, "Topological models for 3d spatial information systems," in *Autocarto 10*, (D. Mark and D. White, eds.), (Baltimore, MD), pp. 368–392, March 1991.

[16] D. Pullar and M. Egenhofer, "Towards formal definitions of topological relations among spatial objects," in *Third International Symposium on Spatial Data Handling*, (D. Marble, ed.), (Sydney, Australia), pp. 225–242, Aug. 1988.

[17] N. Roussopoulos, C. Faloutsos, and T. Sellis, "An efficient pictorial database system for PSQL," *IEEE Transactions on Software Engineering*, vol. 14, no. 5, pp. 630–638, May 1988.

[18] P. Svensson and H. Zhexue, "Geo-SAL: a query language for spatial data analysis," in *Advances in Spatial Databases—Second Symposium, SSD '91*, (O. Günther and H.-J. Schek, eds.), pp. 119–140, Lecture Notes in Computer Science, vol. 525, New York, NY: Springer-Verlag, Aug. 1991.

[19] L. Talmy, "How language structures space," in *Spatial Orientation: Theory, Research, and Application*, (H. Pick and L. Acredolo, eds.), pp. 225–282, New York, NY: Plenum Press, 1983.