A Mathematical Framework for the Definition of Topological Relationships

(Extended Abstract)

Max J. Egenhofer*
National Center for
Geographic Information and Analysis
and
Department of Surveying Engineering
107 Boardman Hall
University of Maine
Orono, ME 04469, USA
MAX@MECAN1.bitnet

John R. Herring
Intergraph Corporation
Mail Stop IW17A2
One Madison Industrial Park
Huntsville, AL 35807-2174, USA
ingr!b17a!mobius!jrh@uunet.uu.net

Abstract

A new theory of binary topological relationships between n-dimensional spatial objects is presented. Unlike previous approaches, it provides a complete coverage, i.e., any possible constellation between two spatial objects can be described by exactly one of the relationships defined. The formalism is based upon fundamental concepts of algebraic topology and set theory. Spatial regions are modeled as point-sets and the binary topological relationships are then defined in terms of the intersections of the boundaries and interiors of two point-sets. Sixteen potential relationships are identified by considering empty and non-empty intersections. Prototypes are shown for the eight relationships that actually exist between two point-sets embedded in a two-dimensional space. More detailed relationships as refinements of these eight relationships are identified by considering other criteria, such as the number of the individual segments of the four intersections or their dimensions.

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

*This work was partially funded by grants from NSF under No. IST 86-09123, Digital Equipment Corporation, and Intergraph Corporation, principal investigator Andrew U. Frank. The support from NSF for the NCGIA under grant number SES 88-10917 is gratefully acknowledged.

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 relationships 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*.[1] 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*.

[1] This classification is not complete since it does not consider *fuzzy* relationships, such as *close* and *next_to* [Robinson 1987], or relationships 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.

- 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).

	$\partial\partial$	$^{\circ\circ}$	∂°	$^{\circ}\partial$
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.

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 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. □

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.[2]

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).

[2]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].

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).

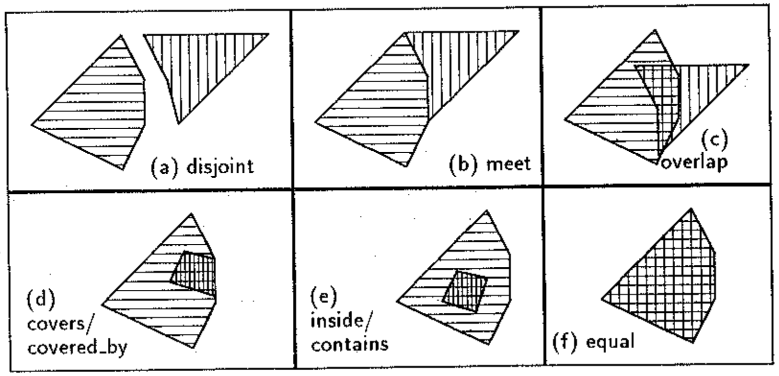

Figure 1: Examples of the relationships between two regions in a 2-dimensional space.

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*.

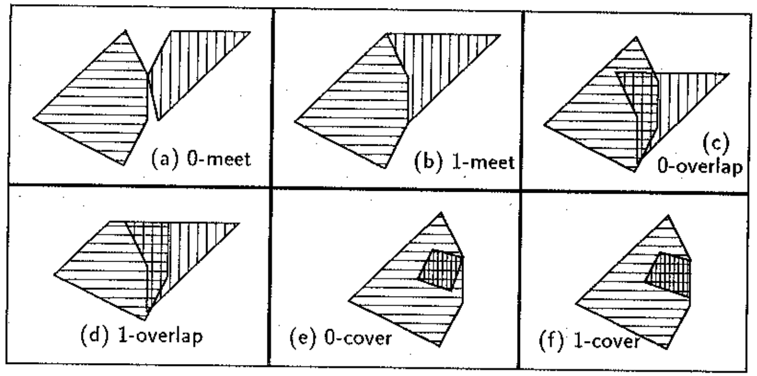

Figure 2: Examples of the detailed relationships between regions in 2-D considering different dimensions in the boundary intersections.

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.

		$\partial\partial$	$^{\circ\circ}$	∂°	$^{\circ}\partial$
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.

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

Acknowledgement

Robert Franzosa improved our understanding of combinatorial topology and helped with the proofs. Andrew Frank's expertise was a valuable contribution to this paper.

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.