Preserving Local Topological Relationships

Mark McKenney, Alejandro Pauly, Reasey Praing & Markus Schneider *
University of Florida
Department of Computer Science & Engineering
Gainesville, FL 32611, USA
{mm7,apauly,rpraing,mschneid}@cise.ufl.edu

ABSTRACT

Topological relationships between objects in space are of great importance in many disciplines. Recently, topological relationships have been defined for complex spatial objects. However, this definition only expresses topological relationships between complex spatial objects as a whole (global view); therefore, topological information between the individual components (local view) that compose the objects is lost. In this paper we propose a novel, hybrid model of topological relationships for composite regions that provides access to the global topological relationships as well as the local topological relationships that exist between the simple regions that are the components of the composite regions involved.

Categories and Subject Descriptors

H.2.8 [**Information Systems**]: Spatial databases and GIS

General Terms

Design

Keywords

Spatial databases, GIS, composite regions, topological relationships, dominance, composition

1. INTRODUCTION

The exploration of topological relationships between objects in space is known to be of great importance within many disciplines such as spatial databases, geographic information systems (GIS), CAD, robotics, cartography, and geoinformatics, to name a few. Topological relationships are used to characterize the relative positions of spatial objects. They are purely qualitative and are preserved under continuous transformations. Numerous developments have been made to formalize the definitions of topological relationships for use in spatial queries as well as spatial data retrieval and analysis. The goals of most of these developments center at improving conceptual clarity, processing efficiency, and object complexity.

Initially, research efforts focused on the definition of models for topological relationships between *simple spatial objects*. The most popular models include the 9-intersection model, which is based on point-set topology, and the RCC model, which is based on spatial logic. The emergence of *complex spatial objects* (roughly defined as homogeneous sets of simple spatial objects, possibly containing holes) led to a redefinition of topological relationships based on the 9-intersection model so that they can handle the more complex objects. However, these definitions consider each complex spatial object as a whole (*global view*) and are not interested in characterizing any relationships between the individual components (*local view*) that compose the objects. As a result, some topological information is lost.

We observe that the use of a global view of topological predicates has the side effect that some relationships between the simple components of complex objects can sometimes be lost in the generalized definition of the global relationship between the entire complex objects themselves. This is because the global view exhibits *dominance* properties among the topological relationships as defined by the 9-intersection model. For example, while building roads between two adjacent countries, one might be interested to know that there is a disjoint island in one of the countries for which a bridge to the other country is required. The disjointedness in this case is overshadowed or dominated by the existing *meet* (adjacent) situation between the countries' mainlands. We characterize the existence of these dominance properties as the *dominance problem* of the 9-intersection model.

A second observation of the global view of topological predicates using the 9-intersection model is that they exhibit *composition* properties. In other words, the global predicate may indicate a certain relationship between two objects that does not exist locally. For example, consider two complex regions that have individual faces that satisfy the *inside*,*covers*, and *meet* predicates. Globally, this configuration satisfies the overlap predicate even though no faces *overlap* locally. We denote this as the *composition problem* of the 9-intersection model.

This paper focuses on a formal treatment of topological

*This work was partially supported by the National Science Foundation under grant number NSF-CAREER-IIS-0347574.

GIS'06, November 10–11, 2006, Arlington, Virginia, USA.

predicates between composite regions (i.e. complex regions without holes) that do not suffer from the dominance or composition problems. Because composite regions can essentially be modeled as collections of simple regions, we are able to utilize the well known eight topological relationships between simple regions as a basis to build our model of predicates without dominating or composition properties. This allows us to lay a precise groundwork for topological predicates that preserve information about local topological relationships between components of composite regions that can later be generalized to a complete model of topological predicates without dominance and composition properties between complex regions and complex spatial objects in general.

The goal of this paper is to provide a model that leverages the local view of topological relationships between composite regions to provide a global view of topological relationships that does not exhibit the dominance and composition problems. In essence, we provide a *hybrid* view that enables access to the global view of topological relationships between composite regions without sacrificing access to the local view. This hybrid view presents new opportunities for querying at all views. This effectively allows the user to fully customize the level of granularity at which to operate; furthermore, it precludes the need for clustering methods that are currently the only option for making large sets of topological predicates accessible to end-users.

In the next section, we review related work on spatial data types and topological relationships. In Section 3, we define the topological relationships between composite regions. Section 4 classifies existing models of topological predicates as being global, local, or hybrid views of topological relationships, and describes the dominance and composition problems in detail. We present a new model of *localized topological predicates* in Section 5 that does not suffer from the dominance or composition properties. Section 6 details the new *hybridized topological relationship* model which is a more expressive extension of the localized model. In Section 7, we demonstrate the use of the refinement when embedded into a popular database query language. Finally, in Section 8, we provide conclusions and future work.

2. RELATED WORK

In this section, we briefly introduce the basic concepts that are necessary for the understanding of the ideas developed in the rest of the paper.

2.1 Spatial Data Types for Regions

Extensive work has focused on the definition of the data types *point*, *line*, and *region* as the central concept for spatial data models and query languages. These types provide fundamental abstractions for modeling the structure of geometric entities and their relationships. We distinguish three *generations* of spatial data types for regions (Figure 1). The first generation, known as *simple regions*, defines a region as a two-dimensional point set in $\mathbb{R}^2$ that is topologically equivalent to a closed disk. Increased application requirements lead to the second, more expressive, generation of *composite regions*, which defines a region as a finite collection of simple regions that may be *disjoint* or *meet* at articulation points. The need to model more complex features of geographic reality finally lead to the third generation of regions, *complex regions*, which consist of a finite number of faces that may be *disjoint* or *meet* at articulation points, possibly containing a finite number of holes that may be *disjoint* or *meet* at articulation points.

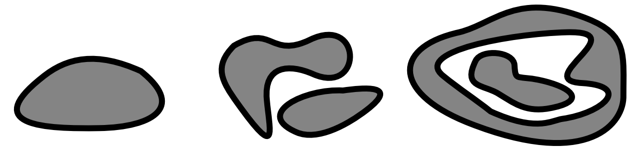

(a) (b) (c)

Figure 1: Sample illustrations for a simple region(a), a composite region(b), and a complex region(c).

2.2 Topological Relationships

Topological relationships refer to a qualitative description of the relative position of two spatial objects. Topological predicates are used to discover such relationships in spatial databases and answer questions such as: "Do countries A and B *overlap*?" or " Are those two roads *disjoint*?". Well known models such as the RCC model [5] based on spatial logic and the 9-intersection model (9IM) [2] based on point set topology focus on the definition of topological relationships between simple regions.

The 9IM characterizes the topological relationship between two spatial objects by evaluating the non-emptiness of the intersection between all combinations of the interior ($^\circ$), boundary (∂) and exterior ($^-$) of the objects involved. A unique 3×3 matrix with values filled as illustrated in Figure 2 describes the topological relationships between each pair of spatial objects.

$$\begin{pmatrix} A^\circ \cap B^\circ \neq \varnothing & A^\circ \cap \partial B \neq \varnothing & A^\circ \cap B^- \neq \varnothing \\ \partial A \cap B^\circ \neq \varnothing & \partial A \cap \partial B \neq \varnothing & \partial A \cap B^- \neq \varnothing \\ A^- \cap B^\circ \neq \varnothing & A^- \cap \partial B \neq \varnothing & A^- \cap B^- \neq \varnothing \end{pmatrix}$$

Figure 2: The 9-intersection matrix for topological relationships.

The set of topological relationships between two simple regions identified by the 9IM is {*overlap*, *meet*, *equal*, *disjoint*, *inside*, *contains*, *covers*, *coveredBy*}.

Extensions to the 9IM for simple regions have been proposed in order to handle simple regions with holes [3], complex spatial objects [6], and *composite* regions (i.e., complex regions without holes) [1]. The latter model characterizes the topological relationships between composite regions as the conjunction of topological relationships between their underlying simple regions. For two composite regions A and B with n and m components respectively, a matrix of $n \cdot m$ elements represents the topological relationship between A and B. This means that the number of topological relationships between two composite regions is dependent on the number of components in the regions, resulting in an infinite number of predicates. Therefore, we identify the finite set of topological relationships between composite regions based on the 9IM in Section 3 that is independent of the number of components and that we use as a basis for the rest of this paper.

3. TOPOLOGICAL RELATIONSHIPS BETWEEN COMPOSITE REGIONS

In this section, we define composite regions and derive their topological relationships using a similar approach as for complex regions in [6]. Let *sreg* denote the set of all simple regions. A simple region is defined as a closed two-dimensional subset of $\mathbb{R}^2$ that is topologically equivalent to a closed disc. A composite region A is defined as a homogeneous set of simple regions [1]. For a composite region $A = \{A_0, A_1, \ldots, A_n\}$ the following properties are enforced:

(i) $\forall A_i \in A : A_i \in sreg$

(ii) $\forall A_i, A_j \in A, i \neq j : A_i^\circ \cap A_j^\circ = \varnothing$

(iii) $\forall A_i, A_j \in A, i \neq j : |\partial A_i \cap \partial A_j|$ is finite

Although holes are not allowed in composite regions, "hole-like" configurations can exist if two components of one region touch at a single point of their boundaries at two different locations (see Figure 3). We denote the set of all composite regions as *creg* so that $X \in creg \Rightarrow X \subset sreg$. Note that not all subsets of *sreg* belong to *creg* due to properties (ii) and (iii) above.

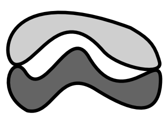


Figure 3: Sample composite region with two components presenting a hole-like structure.

In order to define topological predicates for composite regions, we consider the complete set of topological predicates between complex regions which is given in [6]. The predicates are derived by the *proof by constraint and drawing* technique in which constraints are given to eliminate 9-intersection matrices that represent invalid topological configurations. All remaining configurations are then proved to be valid by providing prototypical examples of them. Because composite regions can only express a subset of what complex regions can express (i.e., multi-component regions without holes), the topological configurations possible between two composite regions are a subset of those between complex regions. The predicate constraints provided in [6] for complex regions also apply to composite regions. Therefore, the complete set of topological predicates between composite regions can be obtained by posing additional constraints that eliminate matrices representing topological configurations that cannot be drawn with composite regions. It turns out that only a single additional constraint expressed by the following lemma is needed.

Lemma 1 *If the boundary of a composite region A and the exterior of a composite region B do not intersect, then either the interior of A is completely contained in the interior of B or the boundary of A intersects the boundary of B, and vice versa when the boundary of B does not intersect the exterior of A, i.e.,*

$$\partial A \cap B^- = \varnothing \;\Rightarrow\; (A^\circ \subseteq B^\circ \;\vee\; \partial A \cap \partial B \neq \varnothing) \;\vee$$
$$\partial B \cap A^- = \varnothing \;\Rightarrow\; (B^\circ \subseteq A^\circ \;\vee\; \partial B \cap \partial A \neq \varnothing)$$

PROOF. If the boundary of A does not intersect the exterior of B, then it must be contained in the closure $B^\circ \cup \partial B$ of B. If the boundary of A is completely contained in the interior of B, then the interior of A is completely contained in the interior of B. This is guaranteed by the Jordan Curve Theorem and the definition of composite regions in which the boundary of a hole-like structure in B can only intersect the interior of region A if it passes through the boundary of A. The same proof can be applied to the reverse case. □

This constraint eliminates matrices 27 and 28 from the 33 complex region relationships given in Table 6 of [6] leaving 31 valid relationships between composite regions. We omit the prototypical drawings of the resulting configurations due to a lack of space.

Theorem 1 *Based on the 9-intersection model, 31 different topological relationships exist between pairs of composite region objects.*

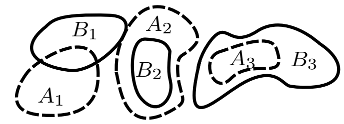


Figure 4: Two composite regions, one shown with a solid line and the other shown with a dashed line.

4. CLASSIFYING MODELS OF TOPOLOGICAL PREDICATES

In this section we introduce a classification of topological predicate models based on their structural view. This classification sheds light on important drawbacks of constraining the representation to one of the views. Section 4.1 examines the behavior of the 9IM model for topological predicates between composite regions as it applies to a scene of objects. We analyze two scenes and the 9IM predicates that characterize them to reveal the dominance (Section 4.2) and composition (Section 4.3) problem of the 9IM model.

4.1 Views of Topological Predicates

Many models of topological predicates have been proposed in the literature, some of which have been derived from the 9IM model for simple objects; however, there has been little attempt to classify the models in any meaningful way. Here we use an example to motivate a classification of topological predicate models into three *views* of predicates. We provide a single scene of composite region objects, Figure 4, and use it to demonstrate the various views.

The first view of topological predicates we consider is the *global view*. Topological predicates that take into account the global view of a scene equate any given scene to a single topological predicate. For instance the 9IM for simple regions is a global view model, as is the 9IM for composite regions, which is an extension of the 9IM for simple regions. The 9IM for composite regions classifies the objects in Figure 4 as satisfying the *overlap* predicate. However, such a global view does not necessarily indicate complete information about a scene. In this case, regions A and B do indeed *overlap*, but a *meet* situation also exists between two components, as well as *inside*, *covers*, and other relationships. While the global view of the 9IM does assign a unique predicate to this scene, it effectively hides other interactions between components of the composite regions.

A second category is the *local view* of topological relationships, which can also be characterized as the *existence view*.

This view focuses solely on the topological relationships between the individual components of multi-component objects. Queries that take advantage of the local view can typically be posed in the form of an existence question; for example, a user may wish to ask "does there exist a *meet* interaction between object A and object B?". The local view is fundamentally different from the other views in that it does not attempt to assign a unique predicate to a given configuration of objects. Instead, it assigns a unique predicate to the interaction of any two components between spatial objects.

Finally, the *hybrid view* combines the information provided by global and local views to allow multiple local interactions in a given scene to be expressed in a unique global predicate. One example of such a view is given in [1], where the authors present a model for topological predicates which describes a scene as a matrix of predicates where each entry in the matrix represents a 9IM predicate between components of each object. Such a view provides a single, although complex, predicate for the entire scene while maintaining the information about all interactions between all faces of the objects. The advantage of the hybrid view is that a global predicate can be assigned to a scene such that local information is not hidden. While the model in [1] is useful, it does not provide a finite set of topological predicates between composite regions. Without a finite set of predicates, implementing and using such a model becomes costly.

4.2 The Dominance Problem

A major drawback to global views of modeling topological relationships is the *dominance problem*. The dominance problem is characterized by a global topological configuration overshadowing the existence of other local topological configurations between a pair of given objects. Figure 4 depicts an arrangement of two composite region objects A and B in the plane. As described above, the regions are globally in an *overlap* configuration due to the overlapping of components A_1 and B_1. However, if the interaction of each face of object A with object B is examined independently i.e., *locally*, then in addition to the *overlap* configuration, a *meet* configuration is observed between A_2 and B_1. Because the global configuration for regions is *overlap*, we say that the *overlap* configuration dominates the *meet* configuration. Other examples are the *inside* configuration between A_3 and B_3 as well as the *disjoint* configuration between B_1 and A_3; they are all dominated by the *overlap* configuration.

In order to cope with the dominance problem, it is necessary to determine which topological predicates dominate which other topological predicates. In other words, we must define a dominance order among the topological predicates. We can determine a dominance order based on the 9-intersection matrices of each topological predicate. For a 9-intersection matrix $(p_{ij})_{0 \leq i,j \leq 2}$ representing a topological predicate, we define the *not-empty-entry* set NE_p to be the set of matrix coordinates corresponding to matrix entries that are equal to 1 (*true*). Conversely, we define the *empty-entry* set E_p to be the set of all matrix coordinates corresponding to entries in matrix $(p_{ij})_{0 \leq i,j \leq 2}$ that are equal to 0 (*false*). Formally, we define:

$$NE_p = \{(i,j) | 0 \leq i,j \leq 2 \wedge p_{ij} = 1\}$$
$$E_p = \{(i,j) | 0 \leq i,j \leq 2 \wedge p_{ij} = 0\}$$

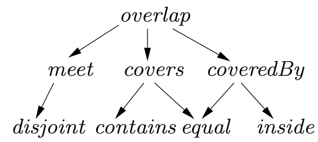


Figure 5: The dominance hierarchy for topological predicates between simple regions. An arrow $p \rightarrow q$ means that p dominates q.

Let T_{creg} be the set of topological predicates between composite regions. For two topological predicates $p, q \in T_{creg}$, we define the *dominates* operator to return *true* if and only if the not-empty-entry set of q is a subset of the not-empty-entry set of p:

$$dominates : T_{creg} \times T_{creg} \longrightarrow \mathbb{B}$$
$$dominates(p,q) := \text{if } NE_q \subset NE_p \text{ then } true \text{ else } false$$

Based on the dominates operator, for illustration purposes we construct the dominance hierarchy for the elements of the set T_{sreg} of topological predicates between simple regions as shown in Figure 5. The dominance hierarchy can be determined for the topological predicates between composite and complex points, lines, and regions, as well as any combination of those.

4.3 The Composition Problem

The composition problem arises due to the observation that global topological relationships can be present without actually existing locally. For example in Figure 6, locally we distinguish the existence of an *equal* between A_1 and B_1, an *inside* (A_2, B_2) and a *coveredBy* (A_3, B_3). If we look at the global 9IM describing the topological relationship between A and B, we notice that it represents the *overlap* predicate even though no such *overlap* exists locally. Whereas the dominance problem relates to the fact that the global view hides information about local relationships that do exist, the composition problem hides information about local relationships that do not exist.

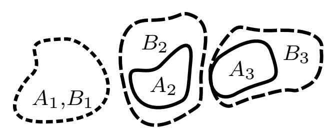


Figure 6: A scene exposing the composition problem of the 9IM.

5. LOCALIZED TOPOLOGICAL PREDICATES

In this section, we present a novel model of *localized topological relationships* for multi-component spatial objects that does not suffer from the dominance and composition problems. The *localized topological predicate* (LTP) model is a hybrid view of topological relationships in that it utilizes local information to provide a global relationship between two objects. We first provide a detailed description of what an LTP is, and then provide a useful, bit vector notation for referring to individual LTPs. Much like deriving topological predicates from the 9-intersection model, once the bit vector notation is characterized, we must identify the set of valid bit vectors.

5.1 The Topological Bit Vector

In order to discover the local topological relationships between composite regions, we define *existence topological predicates* (ETP). For $p \in T_{sreg}$, the ETP $p_e \in ET_{creg}$ is defined as a function $p_e : creg \times creg \to \mathbb{B}$. For any $A, B \in creg$:

$$p_e(A,B) = \begin{cases} true, \ \exists\, A_i \in A, \exists\, B_j \in B : p(A_i, B_j) = true \\ false, \text{ otherwise} \end{cases}$$

Based on this definition, for composite regions we identify a total of eight existence topological predicates that exactly correspond to the eight topological predicates between simple regions, i.e. $ET_{creg} = \{disjoint_e,\ meet_e,\ equal_e,\ inside_e,\ coveredBy_e,\ contains_e,\ covers_e,\ overlap_e\}$.

For composite regions A and B we define the localized topological predicate that describes their relationship as a conjunctive boolean expression with exactly eight clauses, each with a single element that corresponds to a predicate from ET_{creg} or its negation. That is, an LTP characterizes the topological predicate between two composite regions by asserting which ETPs hold and which do not hold between the composite regions. Let $E(A,B)$ and $F(A,B)$ be the sets of ETPs that yield *true* and *false* respectively, for A and B, i.e.

$$\begin{aligned} E(A,B) &= \{e \in ET_{creg} \mid e(A,B) = true\} \\ F(A,B) &= \{f \in ET_{creg} \mid f(A,B) = false\} \end{aligned}$$

We define $l \in LT_{creg}$ (where LT_{creg} is the set of LTPs between composite regions) as:

$$l(A,B) = \left(\bigwedge_{e \in E(A,B)} e(A,B) \right) \wedge \left(\bigwedge_{f \in F(A,B)} \neg f(A,B) \right)$$

Based on this definition we identify the set $\Phi(ET_{creg})$ that contains all possible such conjunctions of elements (and negated elements) in ET_{creg}. Clearly $LT_{creg} \subseteq \Phi(ET_{creg})$ holds. We represent every element in $\Phi(ET_{creg})$ by a *topological bit vector* (TBV), that contains one entry for each element in ET_{creg}. We refer to the topological bit vectors for elements in LT_{creg} as *localized topological bit vectors* and the set of such vectors is denoted $LTBV_{creg}$. Specifically, we define every element in $LTBV_{creg}$ as a bit vector that has the following structure: $[disjoint_e,\ meet_e,\ equal_e,\ inside_e,\ coveredBy_e,\ contains_e,\ covers_e,\ overlap_e]$. For a given pair of composite regions A and B, an entry in the bit vector is set to 1 (0) if the corresponding ETP yields *true* (*false*). The LTP for Figure 6 is represented by the following conjunctive boolean expression:

$$disjoint_e(A,B) \wedge \neg meet_e(A,B) \wedge equal_e(A,B) \wedge inside_e(A,B) \wedge coveredBy_e(A,B) \wedge \neg contains_e(A,B) \wedge \neg covers_e(A,B) \wedge \neg overlap_e(A,B)$$

Its corresponding element in $LTBV_{creg}$ is represented by the 8-bit vector: $[1, 0, 1, 1, 1, 0, 0, 0]$.

5.2 Identifying LTPs

In order to identify the elements that belong to LT_{creg}, we first consider the complete set of 8-bit vectors that can exist (i.e., $2^8 = 256$ such vectors). By considering the semantics of the bits, we can see that not all combinations are possible. A trivial example is represented by the TBV $[0, 0, 0, 0, 0, 0, 0, 0]$ that we can determine to be topologically invalid due to the fact that every pair of (non-empty) simple regions must satisfy at least one of the eight topological predicates. We follow a procedure of *constraint* and *validation* in order to successfully identify the complete set of valid TBVs for composite regions. The first step of the procedure is to eliminate, by way of constraint rules, TBVs that are topologically invalid. The second step entails validating the remaining TBVs. If all remaining TBVs are successfully validated, then we have reached a complete set, otherwise we must identify a new constraint rule. This sequence is repeated until the validation is complete.

Based on the procedure described above we have identified four constraints that eliminate all the invalid TBVs. The first constraint is the trivial case described above. The other three constraints, which are formally defined by Lemmas 2, 3, and 4, eliminate TBVs that do not have the *disjoint* or *meet* bit set when the remaining bits in the vector require one of these to be set. Let $A = \{A_1, A_2, \ldots, A_n\} \in creg$ and $B = \{B_1, B_2, \ldots, B_m\} \in creg$, $1 \le i,k \le n$, $1 \le j,l \le m$, $i \ne k$, and $j \ne l$:

Lemma 2
$inside(A_i, B_j) \ \wedge\ q(A_k, B_l) \Rightarrow disjoint(A_i, B_l)$ where $q \in \{contains, covers, equal\}$

PROOF.

$$\begin{array}{ll} inside(A_i,B_j) & \Rightarrow \partial A_i \subseteq B_j^\circ \ \wedge\ A_i^\circ \subseteq B_j^\circ \\ q(A_k,B_l) & \Rightarrow A_k^\circ \supseteq B_l^\circ \\ A_i^\circ \subseteq B_j^\circ \ \wedge\ A_k^\circ \supseteq B_l^\circ & \Rightarrow A_i^\circ \cap B_l^\circ = \varnothing \\ \partial A_i \subseteq B_j^\circ & \Rightarrow \partial A_i \cap \partial B_l = \varnothing \\ A_i^\circ \cap B_l^\circ = \varnothing \ \wedge\ \partial A_i \cap \partial B_l = \varnothing & \Rightarrow disjoint(A_i,B_l) \end{array}$$

□

Lemma 3
$p(A_i, B_j) \ \wedge\ q(A_k, B_l) \Rightarrow r(A_i, B_l)$ where $p \in \{equal, inside, coveredBy\} \wedge q \in \{contains, covers, equal\} \ \wedge\ r \in \{disjoint, meet\}$

PROOF.

$$\begin{array}{ll} p(A_i,B_j) & \Rightarrow A_i^\circ \subseteq B_j^\circ \\ q(A_k,B_l) & \Rightarrow A_k^\circ \supseteq B_l^\circ \\ A_i^\circ \subseteq B_j^\circ \ \wedge\ A_k^\circ \supseteq B_l^\circ & \Rightarrow A_i^\circ \cap B_l^\circ = \varnothing \\ & \Rightarrow disjoint(A_i,B_l) \ \vee\ meet(A_i,B_l) \end{array}$$

□

Lemma 4
$equal(A_i, B_j) \ \wedge\ q(A_k, B_l) \Rightarrow r(A_i, B_l)$ where $q \in \{overlap, coveredBy, contains, covers, inside\} \wedge r \in \{disjoint, meet\}$

PROOF.

$$\begin{array}{ll} equal(A_i,B_j) & \Rightarrow A_i^\circ = B_j^\circ \\ q(A_k,B_l) & \Rightarrow A_k^\circ \cap B_l^\circ \ne \varnothing \\ A_i^\circ = B_j^\circ \wedge A_k^\circ \cap B_l^\circ \ne \varnothing & \Rightarrow A_i^\circ \cap B_l^\circ = \varnothing \\ & \Rightarrow disjoint(A_i,B_l) \vee meet(A_i,B_l) \end{array}$$

□

These four constraint rules are sufficient to eliminate 90 TBVs, leaving 166 to be validated. In order to validate these TBVs we employ the theory of binary constraint networks and path-consistency. Based on a variation of the method presented in [4], we implemented a program that proves the validity for each of the remaining 166 TBVs by ensuring that there exists at least one corresponding topologically consistent scene.

5.3 Comparing LTPs with 9IM

Although the $LTBV_{creg}$ model of localized topological relationships effectively exposes the local interactions among components of composite regions, it is not yet clear how expressive the model is in comparison to the 9-intersection model. We determine the expressiveness of the LTP model with respect to the 9-intersection model based on the following: if every spatial configuration that is representable and distinguishable by the 9IM model can also be represented and distinguished by the LTP model, then the LTP model is at least as *globally* expressive as the 9-intersection model. Otherwise, we say that the LTP model is *globally* less expressive than the 9-intersection model.

It turns out that a number of spatial configurations are distinguishable by the 9-intersection model (i.e., they have different matrix representations) but not by the LTP model. For example, both scenes illustrated in Figure 7 are represented by different 9-intersection matrices but by the single bit vector $[0, 0, 0, 0, 0, 0, 0, 1]$, i.e., all pairs of components from the respective objects *overlap*. Thus, we conclude that the LTP model is *globally* less expressive than the global 9IM model.

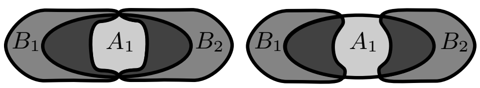


Figure 7: Two spatial configurations with the same topological bit vector

Conversely, we compare the *local* expressiveness of the models to determine whether every spatial configuration that is representable and distinguishable by the LTP model can also be represented and distinguished by the 9IM model. Figures 8a, 8b, and 8c show three spatial configurations which are distinguished by three different bit vectors $[1, 0, 0, 0, 0, 0, 0, 1]$, $[0, 1, 0, 0, 0, 0, 0, 1]$, and $[0, 0, 0, 1, 0, 0, 0, 1]$ respectively. However, all three scenes are represented only by a single 9-intersection matrix. Thus, we conclude that the 9IM model is *locally* less expressive than the LTP model.

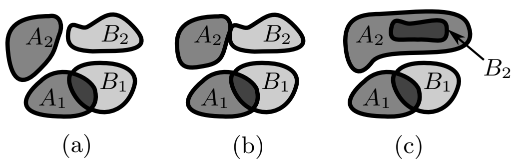


Figure 8: Three spatial configurations with the same 9-intersection matrix

Due to the differences in expressiveness of the LTP model and the 9IM at both the global and local levels we cannot draw any conclusion as to the relative general expressiveness of the models. This is due to the fact that the models are fundamentally incomparable. Our goal is to produce a hybrid model of topological predicates that is at least as globally expressive as the 9IM.

We observe that the scenes in Figure 7 cannot be differentiated due to the fact that a global matrix can include global information that cannot be determined from examining purely local information. For example, the bit vector for the scenes in Figure 7 merely indicates that there exists a local *overlap* between a component of composite region A and a component of composite region B, whereas both scenes are distinguished by two different global matrices because of the difference in the existence of intersection between the boundary of A and the exterior of B. The non-existence of this intersection in Figure 7a occurs because another local *overlap* exists such that the boundary of A is now part of the closure of B. To determine such global information from a bit vector, more than simply local information must be included in the LTP model.

6. HYBRIDIZED TOPOLOGICAL PREDICATES

Since we have shown the incomparability of the expressiveness between the LTP and the 9IM, we must determine a different, hybrid view model that is at least as globally expressive as the 9IM model. As described in Section 5.3, some global information cannot be expressed by the LTP model, which keeps purely local information. In order to handle this issue, the necessary global information must also be represented in the model. In this section, we present a *hybrid topological predicate* (HTP) model that combines both local and global information into a single predicate. We refer to the set of all HTPs between composite regions as HT_{creg}. We begin by determining which global information must be included into the HTP model, and then define a new bit vector for HTPs and use it to derive a complete set of valid HTPs between composite regions. Finally, we compare the 9IM and the HTP model.

6.1 The Hybridized Topological Bit Vector

In order to determine which global information is needed in the HTP model, we examine the effects of local interactions on global interactions, and vice versa. To do this, we identify the collections of matrix entries in the global matrix that must be 1 (non-empty intersection) in order for a given local interaction to be possible. In other words, if an existence predicate yields *true*, we must define how it is expressed in the global matrix. For example, if a *meet* interaction occurs between any pair of components from two composite regions, then we know that in the global matrix representing the scene, the boundaries of the objects must intersect. In general, if any local interaction exists between two composite regions such that the interiors intersect, the boundaries intersect, or the interior intersects the boundary, then the corresponding matrix entry in the global 9-intersection matrix will be 1. This is because if any of these interactions occur locally, the existence of other local interactions cannot cause them to not be reflected at the global level. In contrast, any local interactions involving the exterior of either region may not be reflected in the global matrix because other local interactions can overwhelm them. For example, in Figure 7a, the interiors and boundaries of A_1 and B_1 intersect, and adding other local interactions cannot cause them to no longer intersect; thus, they will intersect in the global matrix. Locally, A_1 and B_1 *overlap*, thus their boundaries and interiors intersect with each other's exteriors. However, the inclusion of B_2 in the scene causes the interaction involving the boundary of A_1 with the exterior of B_1 to not be reflected in the global matrix. This is because globally, the exterior of B_1 is not considered independently, instead the exterior of B is considered as a whole.

We determine that the global information that is required to distinguish the scenes in Figure 7 from each other involves only the interactions of the exteriors of the objects. Based on [6], we know that the intersection between exteriors is always non-empty. Based on the observations detailed above, we must be able to represent the other four exterior interactions (i.e., intersections between interior and exterior, boundary and exterior, and their converses) in order to achieve a *one-to-many* relationship from the 9IM to the HTP model so that all cases that the 9IM can differentiate can also be uniquely identified by the HTP model.

To support the HTP model, we extend the topological bit vector into a new *hybridized topological bit vector* whose bit entries represent both local information and global information. The local information bits represent the eight ETPs between composite regions described earlier, whereas the global information bits represent the global exterior interactions. We denote the four meaningful global interactions involving the exterior of either object as *global interaction predicates* (GIP) and define them as: $ie = (A^\circ \cap B^- \neq \varnothing)$, $be = (\partial A \cap B^- \neq \varnothing)$, $ei = (A^- \cap B^\circ \neq \varnothing)$, and $eb = (A^- \cap \partial B \neq \varnothing)$. GIPs are expressed by four global information bits in the hybridized topological bit vector. Thus, each element of HT_{creg} is represented by a 12-bit vector with the following structure: $[\mathit{disjoint}_e, \mathit{meet}_e, \mathit{equal}_e, \mathit{inside}_e, \mathit{coveredBy}_e, \mathit{contains}_e, \mathit{covers}_e, \mathit{overlap}_e, ie, be, ei, eb]$. A bit of the bit vector is set to 1 if its corresponding ETP or GIP yields *true* and 0 otherwise. The set of all hybridized topological bit vectors is denoted $HTBV_{creg}$. Now, we can distinguish between the scenes in Figure 7: the bit vectors corresponding to Figure 7a and 7b are $[0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1]$, and $[0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1]$ respectively.

6.2 Mapping Between HTP and 9IM

The final step in the definition of hybridized topological predicates is the identification of the set $HTBV_{creg}$ that represents the elements in HT_{creg}. We make the assertion that for the elements in $HTBV_{creg}$, the first 8 bits (local information) of all 12-bit vectors must match one of the 166 elements in $LTBV_{creg}$ identified in Section 5.2. This assertion is based on the fact that the information contained in those bits must be identical to an element in $LTBV_{creg}$ or else it describes a topologically invalid scene. The addition of GIPs only reflects global information and thus does not affect the validity of the topological configuration described by the first 8 bits.

To identify the set of valid 12-bit vectors we provide a two stage method for determining the relationship between the set of 9-intersection matrices and the set of 12-bit vectors: (1) determine all 8-bit configurations that correspond to each 9-intersection matrix, and (2) extend each 8-bit configuration into a 12-bit vector by taking the four bits *ie*, *be*, *ei*, and *eb* of the corresponding 9-intersection matrix from stage one and insert them into the four GIP bits of the 12-bit vector.

We achieve stage one using the matrix templates shown in Table 1. For each ETP we derive a set of template matrices so that each template shows how the existence of a local interaction can be expressed in a global 9-intersection matrix. The values of the templates are based on two observations: first, interactions between the exterior of one component of a composite region and the interior or boundary of a component of another composite region may not be expressed globally due to the existence of other local interactions. As an example of this observation, in Figure 7a, the local interaction between the exterior of B_1 and the boundary of A_1 is not expressed globally because of the existence of B_2. Therefore, for a single ETP we create templates that enumerate all combinations of such possible interactions with the following restrictions:

(1) If the boundary of an object intersects the exterior of another object, then its interior must intersect the exterior of the other object [6].

(2) The templates for inside_e, $\mathit{coveredBy}_e$, $\mathit{contains}_e$, and covers_e require the interior of the containing component to intersect the exterior of the contained component.

The second observation used to derive the templates is that the existence of a local interaction involving the interior or boundary of a component of one composite region and the interior or boundary of a component of another composite region asserts that the corresponding interaction will be reflected in the global matrix describing the scene. Thus, if an ETP is *true* for a scene and it involves one of these four interactions, then all of its corresponding matrix templates will also have a 1 in the corresponding matrix entry. Note that values of 0 in the matrix templates are inferred based on the other matrix templates within each group.

$$\mathit{disjoint}_e \rightarrow \begin{pmatrix} - & - & 0 \\ - & - & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & - & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 0 \\ - & - & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & - & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 0 \\ - & - & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & - & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & - & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & - & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & - & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$\mathit{meet}_e \rightarrow \begin{pmatrix} - & - & 0 \\ - & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 0 \\ - & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 0 \\ - & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} - & - & 1 \\ - & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$\mathit{equal}_e \rightarrow \begin{pmatrix} 1 & - & - \\ - & 1 & - \\ - & - & 1 \end{pmatrix}$$

$$\mathit{inside}_e \rightarrow \begin{pmatrix} 1 & - & - \\ 1 & - & - \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & - & - \\ 1 & - & - \\ 1 & 1 & 1 \end{pmatrix}$$

$$\mathit{coveredBy}_e \rightarrow \begin{pmatrix} 1 & - & - \\ 1 & 1 & - \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & - & - \\ 1 & 1 & - \\ 1 & 1 & 1 \end{pmatrix}$$

$$\mathit{contains}_e \rightarrow \begin{pmatrix} 1 & 1 & 1 \\ - & - & 0 \\ - & - & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ - & - & 1 \\ - & - & 1 \end{pmatrix}$$

$$\mathit{covers}_e \rightarrow \begin{pmatrix} 1 & 1 & 1 \\ - & 1 & 0 \\ - & - & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ - & 1 & 1 \\ - & - & 1 \end{pmatrix}$$

$$\mathit{overlap}_e \rightarrow \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Table 1: Matrix templates for mapping 9IM to HTP. Hyphen entries indicate "*don't care*" values.

Using Table 1, we construct a correspondence between global matrices and 8-bit configurations as follows. First, we consider the set K_m of all matrix templates whose *non-empty-entry* set is a subset of the *non-empty-entry* set of a given 9-intersection matrix m (the *non-empty-entry* set of a template matrix consists of all entries that are explicitly set to 1). Then we consider all elements of the powerset of K_m where their boolean sum[1] is equivalent[2] to m. For each of these elements, we create an 8-bit configuration in which a corresponding ETP has a value of 1 if any of its

[1]If an entry is equal to 1 in any matrix in the set, it is equal to 1 in the boolean sum of the matrices of that set

[2]We consider two matrices (template or global) to be equivalent if their *non-empty-entry* sets are equal.

templates exist in the combination. If the resulting 8-bit configuration does not match one of the 166 topologically valid 8-bit vectors, it is discarded. The non-discarded configurations describe the same topological scene as the given 9-intersection matrix m.

In the second stage, we extend each 8-bit configuration to a 12-bit configuration by assigning the combination of GIPs from each 9-intersection matrix to the global information bits of the corresponding 12-bit configurations. This results in 702 12-bit vectors that represent elements in $HTBV_{creg}$. We know this set to be complete based on two arguments: first, the 8-bit configurations are proved to be complete based on the constraint and validation mechanism. Second, the values of the global information bits in a 12-bit vector are directly derived from the 9-intersection matrices that map to the corresponding 8-bit configuration. Any other combination of values for the global information bits would force a mapping into a global matrix that has been proven topologically invalid in Section 3.

Besides identifying the elements of HT_{creg}, the method described above also identifies a *one-to-many* relationship between a 9-intersection matrix and elements of $HTBV_{creg}$. For this reason we determine that the hybridized topological predicate model can distinguish more topological scenes than the 9IM, and is therefore more *globally* expressive than the 9IM while maintaining local information.

7. QUERYING WITH HYBRIDIZED TOPOLOGICAL PREDICATES

In this section, we provide sample queries using the HTP model. We introduce a simple notation that takes advantage of the bit vector representation. We define the function $htp : HT_{creg} \times creg \times creg \to \mathbb{B}$. The first argument is given as a bit vector template with values of 1 and 0 as well as $*$ for wildcard values (see queries for example syntax). The bit vector values correspond to ETPs in the following order: [$disjoint_e$, $meet_e$, $equal_e$, $inside_e$, $coveredBy_e$, $contains_e$, $covers_e$, $overlap_e$, ie, be, ei, eb].

Figure 9a illustrates two adjacent countries. We assume the authorities of country B want to build highways to directly connect to all of their neighboring countries (one of those is A) and increase the influx of tourism. The following query retrieves all the countries neighboring B (i.e., global *meet* which matches all the countries in the figure).

```
SELECT  X.name FROM countries X, countries Y
WHERE   htp([*,1,0,0,0,0,0,0,*,*,*,*],X.shape,
        Y.shape) AND Y.name = "B";
```

In the previous example, the value for $disjoint_e$ is not required. In this way, countries with an island (like A) will be included in the query result. The authorities also want to know how many bridges they will have to build to connect to large island populations in adjacent countries, so they pose the following query to find the adjacent countries with a local $disjoint_e$ interaction, which retrieves only country A:

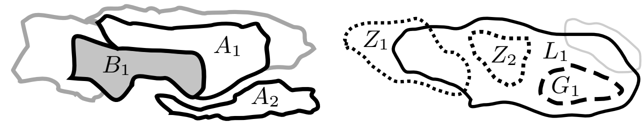
```
SELECT  X.name FROM countries X, countries Y
WHERE   htp([1,1,0,0,0,0,0,0,*,*,*,*],X.shape,
        Y.shape) AND Y.name = "B";
```

Figure 9b depicts roaming areas for wild animals (L: lions, G: gazelles, Z: zebras). An interesting query can discover which herds are facing danger of being wiped out by the lions. Specifically, we are interested in which animals have a component of their roaming area that is completely *inside* the roaming area of the lions. By using the existence predicate $inside_e$, gazelles and zebras are identified. In contrast, a global *inside* predicate only identifies gazelles.

```
SELECT  X.name FROM animals X, animals Y
WHERE   htp([*,*,*,1,*,*,*,*,*,*,*,*],X.roam,
        Y.roam) AND Y.name = "L";
```

Furthermore, we can determine which type of animal is in danger of extinction due to a lack of any safe roaming area outside the lion's roaming area. A query posed using the GIP *ie* identifies gazelles.

```
SELECT  X.name FROM animals X, animals Y
WHERE   htp([*,*,*,*,*,*,*,*,0,*,*,*],X.roam,
        Y.roam) AND Y.name = "L";
```



(a) (b)

Figure 9: (a) Adjacent countries, one with a disjoint component. (b) Animal roaming areas.

8. CONCLUSIONS

Currently, the only models of topological predicates that provide finite sets of predicates between spatial objects support the global view. In this paper, we introduce a hybrid model of topological predicates between composite regions that does not suffer from the dominance or composition problems that appear in the global view models. The concrete application of this model to composite regions lays the necessary groundwork for extending such a model to handle complex regions and complex spatial data types in general. In the future, we plan to precisely define hybrid models of topological predicates between complex points, lines, and regions.

9. REFERENCES

[1] E. Clementini, P. Di Felice, and G. Califano. Composite Regions in Topological Queries. *Information Systems*, 20:579–594, 1995.

[2] M. J. Egenhofer and R. D. Franzosa. Point-Set Topological Spatial Relations. *Int. Journal of Geographical Information Science*, 5:161–174, 1991.

[3] M.J. Egenhofer, E. Clementini, and P. Di Felice. Topological Relations between Regions with Holes. *Int. Journal of Geographical Information Systems*, 8:128–142, 1994.

[4] A. Pauly and M. Schneider. Topological Reasoning for Identifying a Complete Set of Topological Predicates between Vague Spatial Objects. In *Florida Artificial Intelligence Research Society Conference*, pages 731–736. AAAI Press, 2006.

[5] D. A. Randell, Z. Cui, and A. Cohn. A Spatial Logic Based on Regions and Connection. In *International Conference on Principles of Knowledge Representation and Reasoning*, pages 165–176, 1992.

[6] M. Schneider and T. Behr. Topological Relationships between Complex Spatial Objects. *ACM Trans. on Database Systems (TODS)*, 31(1):39–81, 2006.