LNCS 2224, pp27-40, 2001.
(URL: http://www.springerlink.com/link.asp?id=b2m5y03b8mbevxhj)

Modeling Topological Constraints in Spatial Part-Whole Relationships

Rosanne Price[1], Nectaria Tryfona[2], Christian S. Jensen[2]

[1]Department of Computer Science, RMIT University, GPO Box 2476V
Melbourne 3001, Australia
School of CSSE, Monash University, Caulfield East, Vic. 3145 Australia
rosanne@cs.rmit.edu.au
[2]Department of Computer Science, Aalborg University, Fredrik Bajers Vej 7E
DK-9220 Aalborg Øst, Denmark
{tryfona, csj}@cs.auc.dk

Abstract. To facilitate development of spatial applications, we investigate the problem of modeling topological constraints in part-whole relationships between spatial objects, where the related objects may themselves be composite. An example would be countries that belong to a supranational organization, where the countries are themselves composed of states. Current topological classification schemes are restricted to simple, bounded, regular, and/or 0-2D spatial data types; do not support the set-based topological constraints required to describe inter-part relationships such as those between members of a supranational organization; and focus primarily on query rather than design. We propose an approach to modeling topological relationships that allows specification of binary and set-based topological constraints on composite spatial objects. This approach does not depend on restricting the type of spatial objects, can be used to describe part-whole and inter-part relationships, and is at a level of detail suitable for use in conceptual modeling.

1 Introduction

Spatial applications must manage complex relationships between *spatial objects,* objects with associated spatial extents, where the individual spatial objects may be simple (connected) or composite (consisting of disjoint spatial components). Examples are a supranational organization formed from member countries, the division of an administrative region into voting districts, or the structures erected on a building site during construction. Such complex relationships typically involve an asymmetric relationship between spatial objects, where one object—the *whole*—can be used to represent a group of other objects—the *parts*. We refer to this as a *spatial part-whole* (PW) *relationship,* described in detail in [9]. The spatial relationships between the whole and all of its parts and between the individual parts are important for constraint specification during conceptual design as well as in later stages of database development. These characteristics include *orientation*, *metrics*, and *topology*. Of these, topology serves as a particularly useful descriptor in conceptual application modeling both because of its qualitative, thus more intuitive, nature and

H.S. Kunii, S. Jajodia, and A. Sølvberg (Eds.): ER 2001, LNCS 2224, pp. 27-40, 2001.

because—unlike orientation and metrics—it is preserved through many of the common distortions that can occur in representations of real world objects. The focus of this paper is to develop techniques for modeling topological constraints on spatial PW relationships during requirements analysis and conceptual design that are of general applicability in the context of composite spatial objects.

Classification schemes for binary topological relationships have been the subject of extensive study over the years [2,3,4,5,6,7,11]. The research focus is on the development of mathematical formalisms to precisely specify topological relationships. The majority of topological research to date is based on assumptions that are too restrictive for use as a general modeling technique, assuming objects with simple, bounded, and regular regions and lines embedded in 2D space. However, many spatial applications involve semantic entities having holes, discontinuities, and other irregularities. For example, the land mass of a country such as Indonesia or the spatial distribution of different soil types both involve discontinuous spatial extents. Furthermore, applications such as those measuring location in terms of latitude, longitude, and elevation require 3D spatial objects and embedding space.

Recently, researchers have tried to address the challenges of extending topological research to include a wider range of spatial objects, including regions with holes, lines with multiple end-points [5,6], and composite spatial extents [2,3,4,11]. In [5,6,11], topological relations between spatial objects are described based on boundary and interior intersections between object closures (to regularize spatial extents with holes or discontinuities) and object components or discontinuities (e.g. holes or gaps). The most comprehensive work, in terms of the range of spatial object types considered, is described in [4]; including bounded composite spatial objects formed exclusively from either lines (possibly with self-crossings or extra end-points beyond the usual two), points, or regions (possibly with holes). A mutually exclusive and complete set of binary topological relations—*touch, in, overlap, disjoint,* and *cross*—is defined based on boundary, interior, and object intersections and their dimensions. Separate definitions of boundary and interior are used for each type of composite spatial object. In [3], equivalent definitions for the binary topological relations between composite regions are given in terms of relations between their components, i.e. within any component pair composed of a component from each composite object.

A more comprehensive solution to describing topological relationships between composite objects at the component level is proposed in [2]. This is based on a complete set of adverbs that can be used to refine an existing binary classification scheme by extending it to the component level. The adverbs *never, occasionally* (or *partially* if unrelated pairs are disjoint), or *entirely* are used respectively to describe when *no* pair, *some* pair, or *every* pair of composite object components is related by a given binary topological relationship. The adverbs *mostly, mostly$_{rev}$,* or *completely* are used respectively to describe when a component from the other composite object can be found that is related by a given binary topological relationship to (1) each of the second composite object's components, (2) each of the first composite object's components, or (3) both, but using the inverse relationship in the second case. For example, the components of two countries such as Indonesia and the Philippines consisting of island archipelagos should ideally *never* overlap (i.e. no pair of components, one from each country, overlaps). However, if the reality of boundary disputes is considered, the two countries may *partially* overlap (i.e. there may be some component pairs with overlap where there are boundary disputes, but otherwise component pairs are disjoint). For formal definitions of these adverbs refer to [2].

In the context of modeling spatial PW relationships, existing topological research has limitations with respect to the range of spatial data types considered, the understandability of the models proposed, and support for modeling n-ary topological relationships (required to model constraints between spatial parts). To the authors' knowledge, none of the binary topological classification schemes to date explicitly consider spatial extents that are not closed; irregularities such as loops, punctures, and cuts; mixed-dimension composites (e.g. a single composite object consisting of regions, lines, and points); or 3D objects and embedding space. In addition, even when the work considers more complex spatial objects, it is fundamentally based on boundaries and interior intersections [3,4,5,6,11]. Although this allows a high degree of expressiveness in terms of being able to precisely describe a wide range of topological configurations, this comes at the price of increased complexity and reduced understandability. An example is the redefinition of boundary and interior required for each type of composite spatial object in [4] or the identification of topological classes by number [6] or complex conjunctions [11] instead of by name.

A further problem is that the definitions of boundary, interior, and dimension used vary depending on the underlying mathematical model assumed and may not match the intuitive understanding the user has of these concepts. For example, the definition of boundary and interior are formulated only in terms of the spatial object itself in algebraic topology; whereas, in point-set topology, they depend on the embedding space as well. This means that the boundary of a line is its end-points in algebraic topology and for a 1D embedding space in point-set topology, but it is the whole line for a 2D or 3D embedding space in point-set topology. Similarly, the concept of dimension is less intuitive if applied to mixed-dimension composites or their intersections.

In the context of analysis and design of spatial applications, we need a different modeling approach to address application developers' requirements. The level of complexity must be suitable for use in early application development phases and for integration with existing conceptual modeling languages. This potentially means sacrificing, to some degree, the expressiveness of the model (i.e. the level of detail that can be specified for a topological relationship) for the sake of generality (i.e. being able to model the range of different types of spatial objects in spatial applications) and clarity (i.e. based on highly intuitive concepts and classifications).

Furthermore, in order to model spatial PW relationships, we must be able to describe the n-ary topological relationships between the parts. Topological research to date has focussed on binary topological relationships suitable only for describing the relationship between the whole and the geometric union of its parts. In the context of multimedia databases, [8] defines an n-ary temporal relation consisting of an ordered, finite sequence of time intervals where any two adjacent intervals have an identical temporal relation. However, ordering is not suitable for describing topological relationships between a set of spatial objects since, in general (except in the special case of 1D space), there is no inherent linear order in space.

In this paper, a simple framework and modeling constructs intended to facilitate specification of general topological constraints between two or more spatial objects, in the context of spatial PW relationships, are proposed. In Section 2, we describe a simple approach to modeling binary topological relationships based fundamentally on intersection and difference of spatial extents. This is extended to describe n-ary topological relationships in the Section 3. Finally, in Section 4, we apply these methods to the spatial PW relationships discussed in [9], using the proposed binary and n-ary topological relationships to describe constraints on whole-part and part-part relationships respectively. Examples are given to show the applicability and ease of use of the approach adopted.

2 Modeling Binary Topological Relationships

A classification scheme that is both simple (i.e. easy to use) and flexible (i.e. applicable to a wide range of spatial applications) is required to model topological constraints between the whole and the geometric union of the parts in spatial PW relationships. Consider the building site example described in Section 1. It is essential that any structure erected on that site does not extend beyond the site boundary. In an analogous manner, when an administrative region is divided into voting districts, the combined spatial extents of the resulting voting districts must be exactly equal to that of the administrative region. To model these constraints, we require a formal yet simple method of describing binary topological constraints. In this section, we propose a method specifically designed to facilitate conceptual modeling of spatial PW relationships. We first review the assumptions and terminology relevant to the work presented here. For the purposes of this work, it is sufficient to assume an Euclidean model of space (1D, 2D, or 3D) with embedded spatial objects. The classification described here holds under either point-set or algebraic topology; therefore, either can be used as a theoretical basis for discussion. Since the proposed classification scheme is based fundamentally on the set-based concepts of intersection and difference, point-set topology is the more natural choice.

A *spatial extent* is then described as a subset of the points in the embedding space. The spatial extent is considered to be *connected* if any two of its points can be connected by a path consisting entirely of points within the spatial extent and considered to be *disconnected* otherwise. It is *weakly connected* if the same spatial extent becomes disconnected after removal of a finite number of points and *strongly connected* otherwise. A spatial extent that is disconnected is called a *composite* spatial extent consisting of a finite set of *components*, each of which is a connected (weakly or strongly) spatial extent. A *Geometric Union* (GU) of a finite number of spatial extents is the set consisting of all the points from each of the spatial extents including all of their components.

Point-set topology is built from the concept of *neighborhoods*, where there exists a neighborhood both for every point in space and inside the intersection of any two neighborhoods for that point. A *near point* for a spatial extent is one where each of its neighborhoods includes a point in the spatial extent. A spatial extent—whether *connected* or *composite*—forms an *open set* if every point has a neighborhood completely within the spatial extent and forms a *closed set* if it includes all its near points. A spatial extent is called *unbounded* if open, *bounded* if closed, and *partially bounded* otherwise. The largest *open set* in the spatial extent is usually called the *interior* and the rest the *boundary*; however, these terms are not always used consistently in the literature for the reasons discussed earlier. A spatial extent is called *simple* if it is connected and *regular* if it is bounded and contains no irregularities (e.g. no holes, crossings, isolated missing punctures or cuts, extra end-points).

Finally, we adopt the well-known concept of a minimum bounding box, used to approximate an object's location in the embedding space by the smallest rectilinear rectangle completely enclosing that spatial extent. More generally and without restricting the dimension or the type of figure used, the term minimum bounding figure is used here to refer to any simple, regular bounding figure.

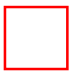
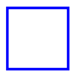
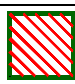
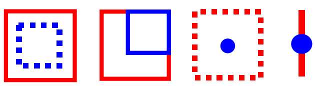
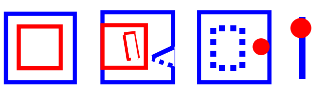
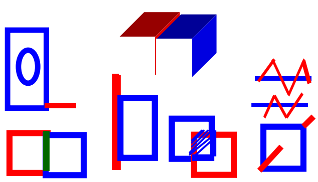
With this foundation, the proposed two-level classification scheme for binary topological relationships (between two spatial extents) can be described. A given spatial extent can consist of a finite number of disconnected or weakly connected parts of the same or different dimensions (from 0D to 3D); can have irregularities such as holes, punctures, cuts, self-crossings, extra end-points and loops; and can be bounded, partially bounded, or unbounded.

The first level of classification is based only on whether the intersection ($\cap$) and difference (-) of the two spatial extents is empty or non-empty, concepts that are easily understandable, intuitive, and not dependent on the dimension of the embedding space. The classification scheme is illustrated in Table 1. Colors are used to distinguish between the two spatial extents and a dotted line used to indicate a partially bounded or unbounded spatial extent. Only simple spatial extents are used in the table for understandability. Composite spatial extents are shown in Fig. 1-3.

After eliminating trivial cases where at least one of the two spatial extents is the empty set, we have the non-intersecting category (the intersection is the empty set) *disjoint* and the intersecting categories (the intersection is not empty) *equal(s)*, *contain(s)*, *inside*, and *connected* as shown in Table 1. So, for example, the equal and contain relationships can be used to model the voting district and building site examples respectively. This set of relationships is complete and mutually exclusive for two non-empty spatial extents, i.e. any topological relationship between two objects falls into exactly one of these categories. For any two non-empty spatial extents, their intersection and differences must be either empty or non-empty. Therefore, by considering exhaustively all the possible permutations (as in Table 1), the resulting categories must be both complete and mutually exclusive. The two non-symmetrical relationships, *contain* and *inside* (i.e. contained-by), can be combined through disjunction into one symmetric *nested* relationship where either the forward difference (A-B) or the reverse difference (B-A), but not both, is the empty set. The connected and disjoint categories have a further level of classification defined when more refinement of the model is required for a specific application.

Connected objects can be further classified based on whether they have a *boundary*, *interior*, or *mixed overlap*, i.e. whether their intersection includes only boundary, only interior, or both boundary and interior points. Since boundary and interior points often represent semantic differences in applications, it is useful to be able to specify whether the intersection involves object boundaries, interiors, or both. For example, in the case of voting districts for a given administrative region, interior points are used to represent administrative jurisdiction and boundary points are used to represent a change in jurisdiction. This example will be discussed further in Section 3, in the context of n-ary topological constraints between spatial parts.

Table 1. Binary Topological Relationships based on Intersection and Difference

Example(s)	**∩**	**A-B**	**B-A**	***R* Name (A *R* B)**
	∅	∅	∅	no name (A,B =∅
	∅	¬∅	∅	no name (B = ∅)
	∅	∅	¬∅	no name (A = ∅)
	∅	¬∅	¬∅	*Disjoint*
	¬∅	∅	∅	*Equal*
	¬∅	¬∅	∅	*Contains (Nested)*
	¬∅	∅	¬∅	*Inside (Nested)*
	¬∅	¬∅	¬∅	*Con-nected*

A crucial aspect of the connected sub-categories is that, in contrast to other proposed topological classifications, the only assumption is that every point in a spatial extent must be either a boundary or interior point, but not both. Further definition is left to the user as appropriate to specific application requirements. This approach supports the intuitive notion that boundary points differ semantically from interior points, but does not dictate further those aspects of the definition that may vary between applications.

To illustrate the sub-categories of *connected* in Fig. 1, we assume a 2D embedding space and point-set boundary and interior definitions dependent on the embedding dimension. Thus a 1D line embedded in 2D space or a single point consists only of boundary points. Note that if we were to assume that the embedding space was 3D instead of 2D, then all the examples shown as *Interior-Overlap* or *Mixed-Overlap* would also become *Boundary-Overlap*. This is because all the points in a 2D area embedded in 3D space are boundary points in point-set topology.

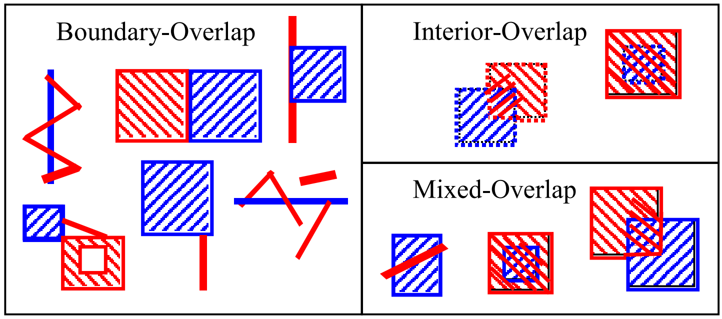


Fig. 1. Connected Sub-Categories

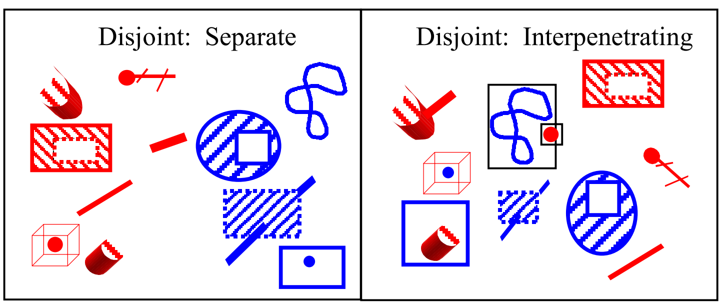


Fig. 2. Disjoint Sub-Categories

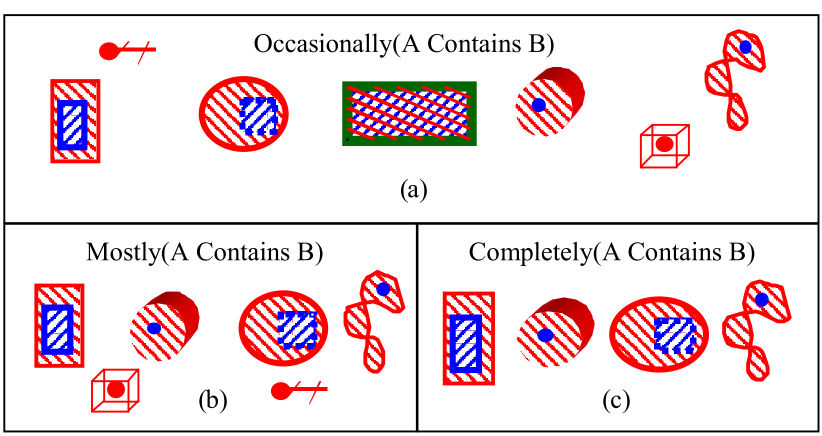


Fig. 3. Using Adverbs to Describe Component-Level Topological Constraints

A disjoint relationship between two spatial entities can be further distinguished based on whether their spatial extents inter-penetrate, i.e. whether the minimum bounding figures of the two objects intersect. The method used to calculate the minimum bounding figure is application dependent, e.g. with respect to the orientation of the axes and granularity. As with the definition of boundary and interior used for the sub-categories of *connected*, the decision as to exactly how to determine the minimum bounding figure is left to the user. *Separate* is a disjoint relationship where the minimum bounding figures of the two spatial extents do not intersect and *interpenetrating* is a disjoint relationship where they do intersect. This distinction is particularly relevant for the applications involving archipelagos, such as the distribution of soil types discussed in Section 1, where the spatial parts in a spatial PW relationship are widely dispersed. Fig. 2 shows examples of separate and interpenetrating disjoint relationships between two composite spatial extents, one colored red and the other blue. Even two simple spatial extents can have an interpenetrating relationship, as shown by the two minimum bounding boxes in black.

The second level of classification consists of complete and mutually exclusive sub-categories in the specific category. For example, every disjoint relationship is either separate or interpenetrating, but not both. That follows logically from the definition of the two categories based on whether the minimum bounding figures of the two spatial extents intersect. Similarly, mutual exclusivity of the *connected* subcategories follows logically from the assumption stated earlier that a spatial extent can be completely partitioned into mutually exclusive sets of boundary and interior points.

Thus the seven categories—separate, interpenetrating, equal, contain, inside, boundary-overlap, mixed-overlap, and interior-overlap—represent a complete and mutually exclusive set of binary topological relationships as shown in the preceding discussions. The more general categories—disjoint, nested, connected, and intersecting—can be derived from these seven relationships. Except for contain and inside, all of the relationships are symmetric.

Although the set of topological relationships is complete and mutually exclusive, there may be certain applications that require a model with greater degree of precision even at the requirements analysis and conceptual modeling phases of system development. For applications requiring a more detailed understanding of the topological relationships between components in pairs of composite objects, the adverbs of [2] can be employed with the binary topological relationships introduced in this section. For instance, Fig. 3 (a), (b), and (c) are all examples of the binary inclusion relationship *A Contains B*. The adverbs from [2] can be used to specify more restrictive constraints and differentiate between these three examples as shown in Fig. 3, with increasing restrictions from (a) to (c). When yet further precision in describing topological relationships is required, e.g. to distinguish between the different cases of boundary-overlap shown in Fig. 1; models involving a more limiting set of assumptions and more complex geometric concepts—such as those described in Section 1—are required.

3 Defining Topological Relationships between *n* Spatial Extents

The binary topological classification described in Section 2 is sufficient to describe topological constraints between a whole and the GU of its parts (i.e. between the *spatial extent* of the whole and the GU of the *spatial extents* of its parts, where the latter is called the part union). However, n-ary topological relationships are required to describe topological constraints between the parts. For example, the voting districts created for an administrative region cannot have overlapping interiors, as this would allow a single constituent to vote in more than one district. In this section, a general method of modeling n-ary topological relationships is described.

Given some binary topological relationship *R* defined for two spatial objects, how can we extend this to *n* spatial objects? For example, how can we extend the definition of boundary-overlap to describe the constraint on the set of voting districts, i.e. that none of the voting districts can share interior points? It follows logically that if a binary topological constraint *R* is extended to *n* spatial objects at least one of three following conditions is true:

1. *R* holds for every pair (i.e. *all*) of the *n* spatial objects.
2. *R* holds for at least one pair (i.e. *some*) of the *n* spatial objects.
3. *R* holds for no pair (i.e. *none*) of the *n* spatial objects.

Although it is clear that this set of three conditions is complete (i.e. given a binary relationship *R* and *n* spatial extents at least one of the three conditions holds), they are not minimal (since *none* can be modeled as ¬ *some*) or mutually exclusive (since condition 2 does not exclude condition 1). However, the conditions are formulated with reference to conceptual modeling with simplicity and ease of modeling as a priority. It is more intuitive to model *none* directly and *some* as *at least one* as evidenced by common usage in natural language. If required, the constraint *at least one but not all* can still be expressed as *some* ∧ *(¬ all)*. This set of conditions is used as the basis for defining modeling constructs to describe n-ary topological relationships. These are defined formally after describing the notation used as follows.

Let $O \stackrel{\text{def}}{=} \{ o_1,\ldots,o_i,\ldots,o_j,\ldots,o_n \} \stackrel{\text{def}}{=}$ a finite set of *n* spatial extents, where n>=2 and i <> j.

Let $R \stackrel{\text{def}}{=}$ a topological expression consisting of:

1. one, a disjunction, or conjunction of the binary relationships from Section 2, or
2. one of the adverbs *mostly*, $mostly_{rev}$, *completely, partially, occasionally, entirely,* or *never* from [2] with (1)
3. a disjunction and/or conjunction of (2).

Let $S \subseteq O$ (a non-empty subset of O) $\stackrel{\text{def}}{=} \{ s_1,\ldots,s_k,\ldots,s_p \} \stackrel{\text{def}}{=}$ a set of *p* spatial extents, where p>=1 and p<=n-2.

We then define the following four modeling constructs for describing n-ary topological relationships, assuming i <> j.

$$all(R, O) \stackrel{\text{def}}{=} \forall\, o_i, o_j \in O\ (o_i \mathrm{R}\, o_j) \quad \textbf{(1)}$$

$$some(R, O) \stackrel{\text{def}}{=} \exists\, o_i, o_j \in O\ (o_i \mathrm{R}\, o_j) \quad \textbf{(2)}$$

$$none(R, O) \stackrel{\text{def}}{=} \neg\exists\, o_i, o_j \in O\ (o_i \mathrm{R}\, o_j) \quad \textbf{(3)}$$

$$linked(R, O) \stackrel{\text{def}}{=} \forall\, o_i, o_j \in O \quad \textbf{(4)}$$

$$(\,(o_i \mathrm{R}\, o_j) \vee (\exists\, S, (\,(o_i \mathrm{R}\, s_1) \wedge ... \wedge (s_{k-1} \mathrm{R}\, s_k) \wedge ... \wedge (s_p \mathrm{R}\, o_j)\,)\,)\,)$$

The first three constructs are based on the three conditions discussed earlier. The last construct, *linked*, describes a special case of *some* where any two spatial extents in the set can be related directly or indirectly by the given topological expression.

Note that the definition of O excludes sets of spatial extents having zero members or one member. If O is empty or has only one member, then *all*, *some*, *none*, and *linked* are defined to be true for all R. If O has two members, then *all* $\Leftrightarrow$ *some* $\Leftrightarrow$ *linked* for all symmetric R.

These modeling constructs allow specification of general topological relationships between the spatial extents—whether simple or composite—of n spatial objects. With the adverbs from [2], the same modeling constructs allow specification of topological relationships between components of pairs of n different composite spatial extents.

There may be some cases where we want to treat a set of composite spatial extents as a set of their individual components. This could be used to model topological constraints between all the individual components of a set of composite spatial extents without any reference to the original composite configurations. To do this, we need to define an additional modeling construct that decomposes a set of spatial extents into the set of all their individual components. That is, given a set O of m composite spatial extents $o_1,\ldots,o_i,\ldots,o_m$ with $n_1,\ldots,n_i,\ldots,n_m$ components respectively and where c_{ik} is the kth component of the ith composite spatial extent o_i, we have the following:

$$decompose(O) \stackrel{\text{def}}{=} \{\ldots,c_{ik},\ldots\} \text{ where } 1 <= i <= m \text{ and } 1 <= k <= n_i \quad \textbf{(5)}$$

We can then use any of the previously defined constructs for n-ary topological relationships, replacing O with *decompose(O)*. For example, consider the case of a national road network, with the entities being individual roads with spatial extents describing their location and geometry. Although a single road usually is a simple polyline, there may be cases where a road may consist of several disconnected segments. For instance, consider a long-distance road that is a freeway for most of the distance, but has a few segments inherited from local road networks that have different names, are not freeways, and may not even be administered by the same transport authority. When modeling the national road network, we want to enforce the constraint that the road network as a whole must be continuous. Since a road can have a composite spatial extent consisting of disconnected segments, this means that there must be some way to travel between every two segments of road in the network. In order to evaluate topological relationships between the set of road segments (rather than roads) in the network, the *decompose* operator is used to refer to individual road segments. The *connected* binary topological operator discussed earlier is used to compare pairs of road segments. The *linked* relation is then used to specify that it must be possible to find a finite sequence of connected pairs linking any two road segments. Assuming that we have the set of roads $r_1,\ldots,r_n$ in the road network, this constraint would be formally specified as *linked*(*connected*, *decompose*($\{r_1,\ldots,r_n\}$)).

4 Topological Constraints on Spatial PW Relationships

In [9], we classify spatial PW relationships based on whether the spatial extent of the whole object is *derived* from or *constraining* those of its parts, termed *spatial derivation* and *spatial constraint* and illustrated respectively by a supranational organization and a building site. Topological constraints between parts are listed as a secondary characteristic leading to further variants beyond the basic classification. The binary and n-ary topological relationships defined in Sections 2 and 3 respectively can now be used respectively to refine the spatial constraint category based on whole-part topology and to illustrate the definition of additional variants based on part-part topology. Formal definitions for spatial PW relationships are given in [9].

Only the inclusion relationship (part union *inside* or *equals* whole) was considered and defined as a sub-category (*spatial inclusion*) of spatial constraint in [9]. We can use the classification of binary topological relationships proposed in Section 2 to provide a more general method of defining topological relationships between a whole and its parts and to further refine the spatial constraint category. As in [9], the goal is to identify useful types of spatial PW relationships. Therefore, refinement is pragmatic (i.e. where we were aware of clear examples) rather than exhaustive. Following this rationale, three more spatial constraint types are identified: *spatial interior*, *spatial equal,* and *spatial cover;* where the relationship of the part union with the whole is respectively *inside*, *equals*, and *contains* or *equals.* The spatial inclusion and spatial interior constraints are transitive, since the topological constraint between the part union and the whole can be equivalently expressed as a constraint between each part and the whole individually. Therefore, any sub-components of a structure located on a building site are also located on that building site. The same is not true of spatial cover or equal, and so these categories are not transitive.

Spatial constraint sub-categories are illustrated in the top portion of Fig. 4, with the specific binary topological constraint between the *part union* and the *whole* indicated in bold type for each sub-category. *Spatial cover* is exemplified by a guaranteed phone service coverage area that must be completely covered by (i.e. inside or equal to) the GU of the phone service cells' spatial extents. A building site and the structures on that building site represent an example of *spatial inclusion*, since no structure can extend outside the building site. The stricter constraint of *spatial interior* applies to house furnishings (referring here to appliances and furniture), since the furnishings must be inside but cannot completely cover the area of the house in order to ensure walking room. Finally, the GU of taxi dispatch zones (the area over which a given taxi driver ranges) must be exactly equal to the metropolitan area covered by the taxi company, i.e. *spatial equal*. This ensures complete coverage of the metropolitan area without risking cases where the company insurance policy may not be applicable.

Variants of the basic spatial constraint sub-categories can be defined based on additional topological constraints between the parts using the n-ary topological relationships from Section 3, as illustrated by the examples in the bottom portion of Fig. 4. The n-ary topological constraint applicable to a specific example is indicated in bold type. The second argument of the n-ary topological constraint (the set of spatial extents) is omitted in the figure and following discussion for readability.

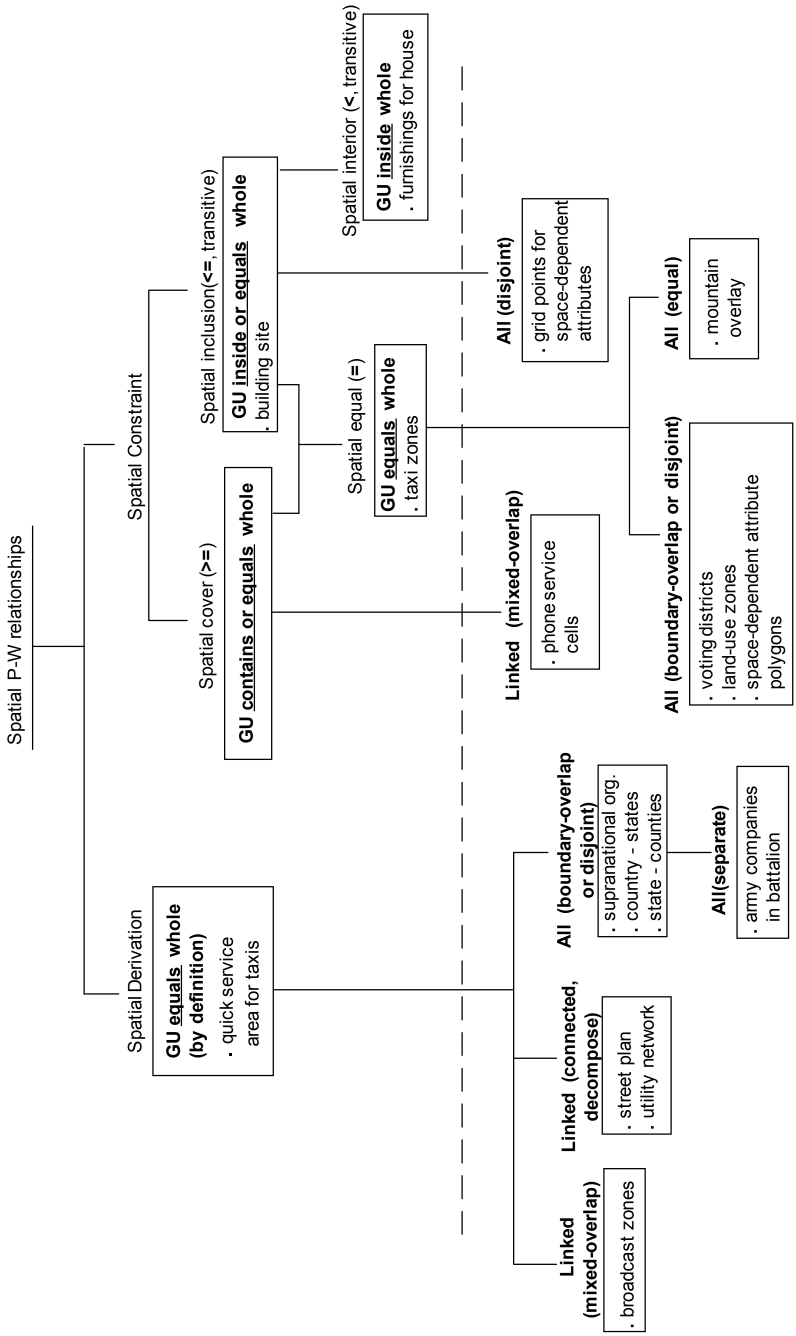


Fig. 4. Topological Constraints in Spatial Part-Whole Relationships

To ensure that the placement of a set of broadcasting transmitters results in continuous broadcasting coverage across the set of transmitters, the *linked(mixed-overlap)* constraint is used to specify that there is a sequence of overlapping broadcast zones—each a simple, bounded spatial extent. Constraints on a set of phone service cells are similar to those on broadcast zones, except with the additional constraint that every point in a given area has phone service—that is, is contained in some phone service cell. This is reflected in the use of a spatial constraint (*spatial cover*) modeling construct for phone cells instead of spatial derivation.

In the transport and utility application examples where a given street or utility component can have a composite spatial extent, *decompose* is used with the *linked(connected)* constraint to ensure a continuous transport and utility network (i.e. the *whole*) with no isolated components (i.e. the *parts*).

Since boundaries are used to uniquely partition administrative responsibility, there cannot be cases of overlapping interior points for member countries of a supranational organization, states in a country, counties in a state, or voting districts. Similarly, shared interior points would be inconsistent in the case of land-use zones and space-dependent attribute polygons, used to represent differences in permissible land usage or an observed attribute values respectively. The predicate *all(boundary-overlap or disjoint)* is used to specify the constraint that interior points cannot be shared between parts in a spatial PW relationship.

In the case of army companies, they must be disjoint and spread out (i.e. not interpenetrating so *all(separate)*) for strategic reasons and to reduce the risk of friendly fire. Sample points used to measure space-dependent attributes should be spread out to improve the sampling accuracy. In this case, disjoint points are necessarily separate, so there is no need to specify this constraint explicitly.

Finally, we have the case of overlays for different thematic attributes over a given region (e.g. mountain vegetation, hydrography, and elevation) in a geographic application. In this case, the constraint *all(equal)* is used to specify that the spatial extents of overlays must be equal.

5 Conclusions

In this paper, we discuss techniques for specifying topological constraints on spatial PW relationships during the analysis and conceptual design phases of spatial application development. A two-level classification scheme for describing binary topological relationships is proposed that is general enough to be suitable for a range of different applications yet is simple to use and understand. The final set of seven relationships—separate, interpenetrating, equal, contain, inside, boundary-overlap, mixed-overlap, and interior-overlap—is complete and mutually exclusive. The defined relationships include those between single and mixed-dimension composites, irregular (with cuts, punctures, holes, self-crossings, extra end-points, loops, etc.), partially bounded or unbounded, and 3D spatial objects. We then define modeling constructs for the specification of n-ary topological relationships. Finally, we show how the proposed techniques can be used to specify topological constraints—both between parts and between the whole and parts—in spatial PW relationships.

Existing conceptual modeling languages such as UML or spatiotemporal extensions based on UML [1,10] contain provisions for general constraint specification but no specific support for describing topological constraints on spatial composites. The techniques proposed here can be used in conjunction with such languages to add the necessary support. For example, to specify that the set of land-use zones associated with an administrative region must be overlapping, the n-ary topological constraint *all(boundary-overlap or disjoint)* can be included in curly braces (used to indicate a constraint in UML) on the association link between the administrative regions and land-use zone classes. The efficient implementation of the specified constraints in later development phases requires the use of representation-dependent algorithms to verify intersection and difference of spatial extents and their components (for composite spatial extents). An overview of the types of representations used for 0D-3D spatial objects and associated algorithms used for these operations are described in [12].

Future work includes the extension of set-based constraint specification to other spatial characteristics such as orientation or metrics and the incorporation of time restrictions in the topological modeling techniques proposed here to support spatiotemporal applications.

References

1. Brodeur, J., Y. Bedard, and M. Proulx. Modelling Geospatial Application Databases using UML-based Repositories Aligned with International Standards in Geomatics, Proc. 8th ACM GIS, 2000, 39-46.
2. Claramunt, C. Extending Ladkin's Algebra on Non-convex Intervals towards an Algebra on Union-of Regions, Proc. 8th ACM GIS, 2000, 9-14.
3. Clementini, E., P. Di Felice, and G. Califano. Composite Regions in Topological Queries, Information Systems, 20(7), 1995, 579-594.
4. Clementini, E. and P. Di Felice. A Model for Representing Topological Relationships between Complex Geometric Features in Spatial Databases, Info. Sciences, 1994, 1-17.
5. Egenhofer, M, E. Clementini, and P. Di Felice. Topological Relations between Regions with Holes, Intl. Journal of GIS, 8(2), 1994, 129-142.
6. Egenhofer, M. and J.R. Herring. Categorizing Binary Topological Relationships Between Regions, Lines, and Points in Geographic Databases, Technical Report, Dept. of Surveying Engineering, Univ. of Maine, Orono, ME, 1991, 1-33.
7. Hadzilacos, T. and N. Tryfona. A Model for Expressing Topological Integrity Constraints in Geographic Databases, Proc. of Intl. Conf. GIS—From Space to Territory: Theories and Methods of Spatiotemporal Reasoning, 1992, 252-268.
8. Little, T.D.C. and A. Ghafoor. Interval-Based Conceptual Models for Time-Dependent Multimedia Data, IEEE TKDE, 5(4), 1993, 551-563.
9. Price, R., N. Tryfona, and C.S. Jensen. Modeling Part-Whole Relationships for Spatial Data, Proc. 8th ACM GIS, 2000, 1-8.
10. Price, R., N. Tryfona, and C.S. Jensen. Extended Spatiotemporal UML: Motivations, Requirements, and Constructs, Journal of Database Management, 11(4), 2000, 14-27.
11. Tryfona, N. and M.J. Egenhofer. Consistency Among Parts and Aggregates: A Computational Model, Transactions in GIS, 1 (3), 1997, 189-206.
12. Worboys, M.F. GIS, A Computing Perspective, Taylor & Francis, London, 1995.