## National Center for Geographic Information and Analysis

# Multiple Topological Representations

By

Bud P. Bruegger and Werner Kuhn
University of Maine, Orono

Technical Report 91-17

July 1991

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

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

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

## Preface and Acknowledgernents

This report contains intermediate results of ongoing research on topological data structures for GIS, representing the state as of summer 1990. The work described was conducted as part of Research Initiative 3 "Multiple Representations" of the NCGIA. Support from the National Science Foundation (NCGIA grant No. SES 88-10917) is gratefully acknowledged. Many thanks go to Kate Beard for editing previous versions of this paper and to Andrew Frank for his comments and advice. Discussions with Renato Barrera and Max Egenhofer contributed to this paper and are greatfully acknowledged. Thanks also to Bob Franzosa for an interesting topology course and comments on this paper.

**Abstract**

Humans use multiple mental models of the world to reason efficiently at different levels of abstraction. Current geographic information systems (GIS) normally use only a single model or representation of the world. If only a single level of abstraction is supported, the utility of the system is limited. Future GIS should support multiple representations at different levels of abstraction, so that adequate levels of abstraction can be found for a large range of scales. Queries that require little detail can then be processed as efficiently as those which require much detail.

This paper proposes an approach to supporting multiple levels of abstraction in a GIS through multiple topological representations (MTR). The representations are linked by hierarchical relations between elementary spatial building blocks called cells. We introduce largest homogeneous cells (LHC's) to guarantee the hierarchical structure of these links. LHC's can easily be modelled in topological representations that use different kinds of cells. They therefore form an interface between an MTR and the underlying implementations of (single) topological representations. MTR's can represent arbitrary configurations of objects in any dimension.

## 1. INTRODUCTION

Human beings use mental models to reason about the world. Models are abstractions of the world, representing only a few aspects compared to the huge amount of detail disregarded. Humans use a multitude of mental models emphasizing different aspects of the world (Mark 1989). These models may represent the same real world object at many different levels of detail. Mental reasoning processes utilize mental models that offer the most adequate level of detail.

Current geographic information systems (GIS) normally deal only with a single model or representation of the world. Their support for spatial reasoning is therefore limited compared to human reasoning. In particular, a lack of representations containing an appropriate amount of detail can cause reasoning processes to be very slow. To improve reasoning capabilities, GIS have to include multiple representations of the world. Research initiative 3 of the National Center for Geographic Information and Analysis (Abler 1987, NCGIA 1988) studies problems sternming from this requirement (Buttenfield 1989, Bruegger 1989a).

A GIS with a single topological representation can take a long time to answer topological queries which seem easy to humans. As a solution, this paper proposes an approach to *multiple topological representations* in GIS that allow access to topological information at different levels of detail. Multiple topological representations integrate several (single) topological representations resulting from different abstractions into a consistent structure. The proposed concept of *largest homogeneous cells* enables to interface multiple topological representations with various known topological representations (Corbett 1979, Frank 1986, Herring 1987) employing different kinds of cells.

The paper is organized as follows. Section two reviews previous and related work. Section three describes how the use of a single topological representation can lead to performance problems unexpected by the system user. We then sketch the basic concept of the proposed solution. In preparation of this solution, section four introduces the concept of largest homogeneous cells (LHC's) and demonstrates how LHC's can easily be modelled in representations using arbitrary kinds of cells. In section five we describe an approach to multiple topological representations based on special properties of largest homogeneous cells. We conclude in section six that the proposed topological structure featuring multiple levels of detail allows quick extraction of information at multiple scales. The outlook in section seven sketches an extension of multiple topological representations to a generalization of a strip tree organizing metrical information at different levels of detail.

## 2. PREVIOUS AND RELATED WORK

This section reviews previous work which is important for the understanding of this paper. In a first subsection, we explain the terms *topological representation*, *object*, and *domain*. The next three subsections prepare for the definition of largest homogeneous cells which will be defined in terms of point set topology. The summary of a formal framework for topological properties and relations proposed by Egenhofer and Franzosa (Egenhofer 1989b, 1990) describes the information that a topological representation models. Largest homogeneous cells will be defined in section four as the minimal cell structure which models this topological information. We restrict our studies of topology to an n-dimensional Euclidian (metric) space. In a last subsection we reference related work on multi-resolution structures.

### 2.1 Topological Representations

The foundation for multiple topological representations is a topological representation. Topological representations describe topological information. For example, the fact that two objects intersect is topological information. Arbitrary rotations, translations, scalings, or map projections do not affect it. Generally, topological information is invariant under so called homeomorphic transformations which preserve neighborhoods (Croom 1989).

The basic building blocks to represent objects in topological representations are cells (NCDCDS 1988). For example, a 0-cell is a node, a 1-cell is a chain, and a 2-cell is a region. Geographic objects can be represented by collections of these cells. Topological relations are captured by boundary and co-boundary relations which connect adjacent cells.

Topological representations have become an accepted method for organizing spatial objects in GIS. Topological models have been described by Corbett (1979), White (1979), Frank and Kuhn (1986), and implementation aspects by Jackson (1989) and Egenhofer et al. (1989a). Single topological representations are being used in modem GIS such as TIGER (Broome 1990) and TIGRIS (Herring 1987). Different authors propose various cell concepts for their representations where cells have specific properties. For example, simplicial complexes (Frank 1986) use triangular 2-cells whereas TIGER and TIGRIS allow a more general shape of 2-cells. We use the term *topological representation* to denote a single layer which integrates multiple themes. The reason for the exclusion of layered representations with topology is given in section three.

The purpose of topological representations is to model the topological properties of a selected part of the world. The *world* is described by a *set of objects and relations*. Humans assume the world to be a continuous infinitely detailed space but perceive distinct objects (Gentner 1983). Examples are trees, houses, and gardens, but also leaves as parts of a tree, tiles on the roof of a house and a grain of sand in the garden. Some objects are not even directly perceivable because of their large size, for example continents.

Of all the infinitely many objects of the world, a GIS only contains a small *set of objects* of interest, defined here as the domain. Since topological representations model topological properties and relations of objects, different domains must lead to different kinds of topological representations. The *domain* can contain an *arbitrary collection of objects in arbitrary configurations*.

For example, objects can partially take up the same space such as a house and a parcel, they can form a hierarchy such as counties, states, and countries, or they can be separate. In case of a hierarchy, some objects consist of others. For example, states are aggregations of counties.

### 2.2 Notions from Point Set Topology

This subsection reviews the relevant definitions of topological properties for readers who are not familiar with point set topology. We define the following terms: *boundary, interior, intersection*, and *connectivity*.

Point set topology is one branch of mathematical topology (Munkres 1966, Croom 1989) which describes properties of spaces. An n-dimensional Euclidian space consists of an infinite number of points. The definition of all properties is based on the notion of a *neighborhood*. In metrical spaces, the distance function defines neighborhoods. An example for a neighborhood is as an open disk in the plane of radius r, centered on a point a, i.e., [x I distance (a, x) < r).

Intuitively, a set of points (e.g., a region in the plane) has boundary points and points in the interior. Formally, a point x belongs to the *boundary* if every (arbitrarily small) neighborhood of x contains a point outside the set and a point of the set. Figure 1 shows a point set S with two distinct points x and y. Every neighborhood around x contains at least a point of the outside and one of S; x is therefore a boundary point. The neighborhood shown around y does not contain any points outside S; y is therefore not a boundary point.

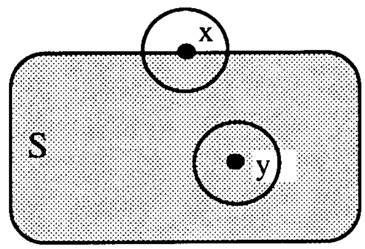
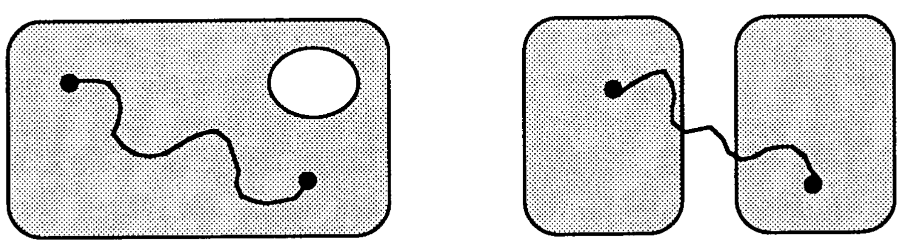



Figure 1: A point set S with a boundary point x and an interior point y.

Points which are elements of a set but not boundary points are called *interior* points. Points that do not belong to the set and are not a boundary point are called *exterior* points. Every (non-empty and bounded) set has a boundary, but not every set has an interior. For example, the set containing a single point is its own boundary and does not have an interior. Note that the definition of boundary does not require the boundary points to be part of the set.

*Intersection* is defined in set theory. Two point sets are said to intersect if they have at least one point in common. If they do not intersect, they are said to be *disjoint*.

One has an intuitive understanding of what a connected point set is. Formally, a set is defined to be *connected* (path connected) if a path can be found between any two element points and the path is totally contained in the set (see figure 2).



(a) (b)

Figure 2: A connected set (a) and a set consisting of two separate components (b).

### 2.3 Objects as Point Sets

Let us look at how the definition of boundary and interior applies to geographic objects. *Objects* are modelled such that they contain their own boundary. The *interior* and *boundary* of an object is given by the point set theoretic definition. Note that this must be done in the minimal embedding space for an object. For example, the boundary of a curve must always be derived in a one-dimensional space, no matter whether the embedding space is one-, two-, or three-dimensional. Figure 3 shows a region and a curve with their boundaries and interiors.

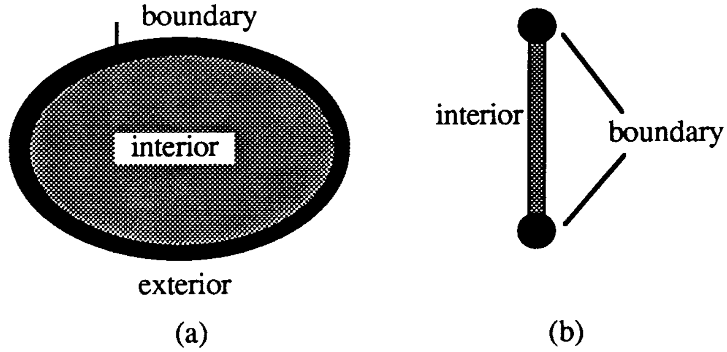


Figure 3: Objects consisting of an interior and a boundary as derived in the minimal embedding space.

### 2.4 Topological Properties and Relations

This subsection reviews a framework for topological properties and relations proposed by Egenhofer and Franzosa (Egenhofer 1989b, 1990). Topological properties we are interested in can be classified into two categories: (1) *topological properties of individual objects* (for example, whether an object has holes) and (2) *topological relations between pairs of objects* (for example, whether two objects overlap).

*The topological properties of individual objects* that we are interested in can be described in terms of *connectivity of their boundary and their interior*. Figure 4 shows some examples (the boundary of the objects are shown with thick lines, the interiors are dotted): The object shown in (a) has two holes. Its interior is connected, i.e., in one piece. Its boundary is not connected but consists of three separate pieces. The object shown in (b) consists of two separate parts. Its interior is not connected but consists of two separate pieces. Its boundary also consists of two pieces.

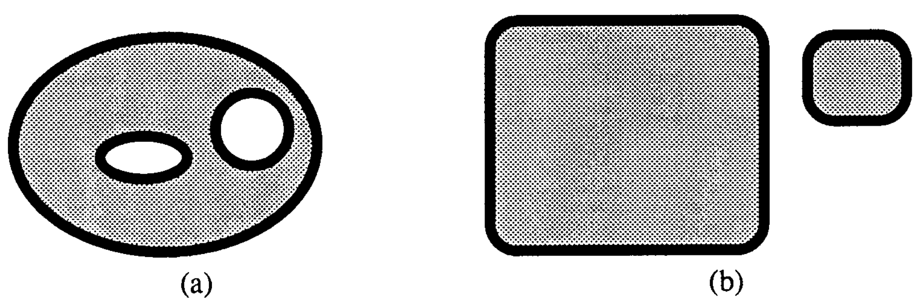


Figure 4: Topological properties of individual objects.

*All topological relations between objects* can be expressed in terms of *intersections of their boundaries and interiors* (Egenhofer 1989b, 1990). Figure 5 shows some examples: In situation (a), the interior of A intersects the interior of B, both boundaries intersect, the boundary of A intersects with the interior of B, but the boundary of B does not intersect with the interior of A. These intersection relations are obviously different in (b).

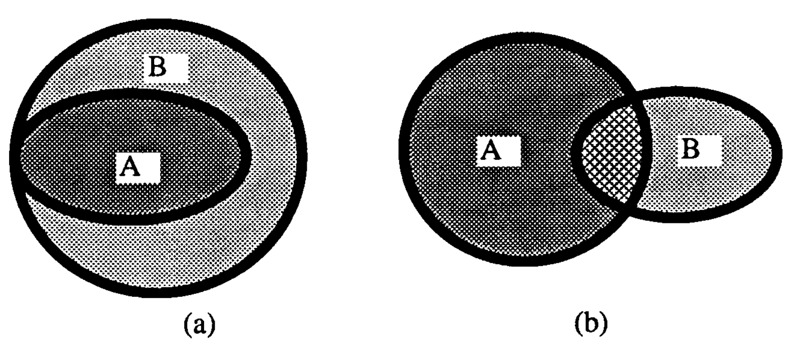
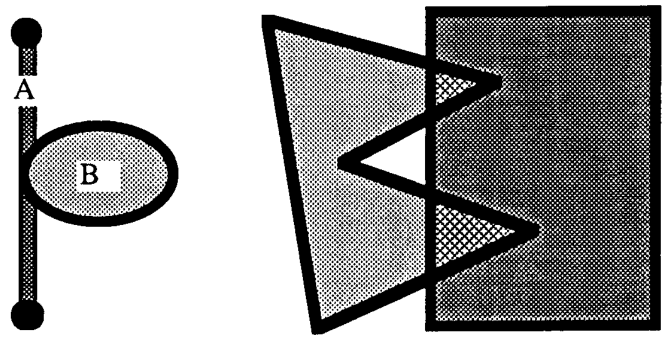


Figure 5: Topological relations between areal objects.

Figure 6 shows two other topological relations between objects. In situation (a) object A is a line which demonstrates that the object dimension has to be taken into account. Situation (b) shows that the intersections of the boundaries and interiors also have to be checked for connectivity (Egenhofer 1989b, 1990): The intersection of the interiors is not connected.




(a) (b)

Figure 6: Topological relations between objects of different dimension (a) and with separated intersections (b).

### 2.5 Related Multi-Resolution Structures

Multiple topological representations can be seen as hierarchical representations of topological properties. In the field of spatial data handling, hierarchical structures have often been proposed as a means to relate multiple representations of different spatial resolution. Strip trees (Ballard 1981), Quad- and Octrees (Samet 1984) are typical examples. A follow-up paper is planned that will describe an extension of the proposed design to a multi-resolution geometrical representation generalizing strip trees. Jones and Abraham propose a similar generalization of strip trees which is not based on topological data structures (Jones 1986). ETAK uses a topology based multi-resolution geometrical representation in its car navigation systems (ETAK 1988). The conception is similar to the design proposed in this paper. However, it does not use largest homogeneous cells and is therefore restricted to one specific implementation of cells.

## 3. PROBLEMS WITH SINGLE TOPOLOGICAL REPRESENTATIONS AND THE CONCEPT OF MULTIPLE TOPOLOGICAL REPRESENTATIONS

Our discussion of performance problems with single topological representations and the proposed solution is based on the *assumption* that the *extraction of information is much more frequent than insertions and updates*. Consequently, we are interested in *efficient information extraction* while insertions and updates are not considered time critical.

This assumption explains why we consider only single topological representations which consist of a single layer with multiple themes: In layered representations, queries involving topological relations between objects of different themes require overlay processing. This prevents fast extraction of such information. In single layer representations, on the other hand, since the overlay processing has been pre-computed during insertion (Frank 1987), the extraction of such information can be efficient. (Actually, multiple topological representations have the potential to separate themes that will never be topologically related into separated representations).

### 3.1 Problems with Single Topological Representations

Single topological representations model a single level of abstraction. This limits the information that can be extracted efficiently to a relatively small range of detailed topological properties. When the representation contains a lot of information irrelevant to a GIS query, the resulting performance hinders interactive use. For example, let a GIS organize hypothetical utility lines, houses, parcels, towns, districts, counties, states, countries, and highways modelled as sets of cells. Figure 7a shows parcels composed of triangular cells. In (b) the same situation is shown in a "zoomed out" view. The parcels shown in (a) now only cover a small part of a town. Further zooming out yields the situations shown in (c) where the town is a small part of a county, and in (d) where the county is part of a state. This sequence illustrates how the number of component cells increases drastically with increasing relative size of the objects. Furthermore, it becomes clear that when we are dealing with large area objects, most of the information represented by their component cells is irrelevant.

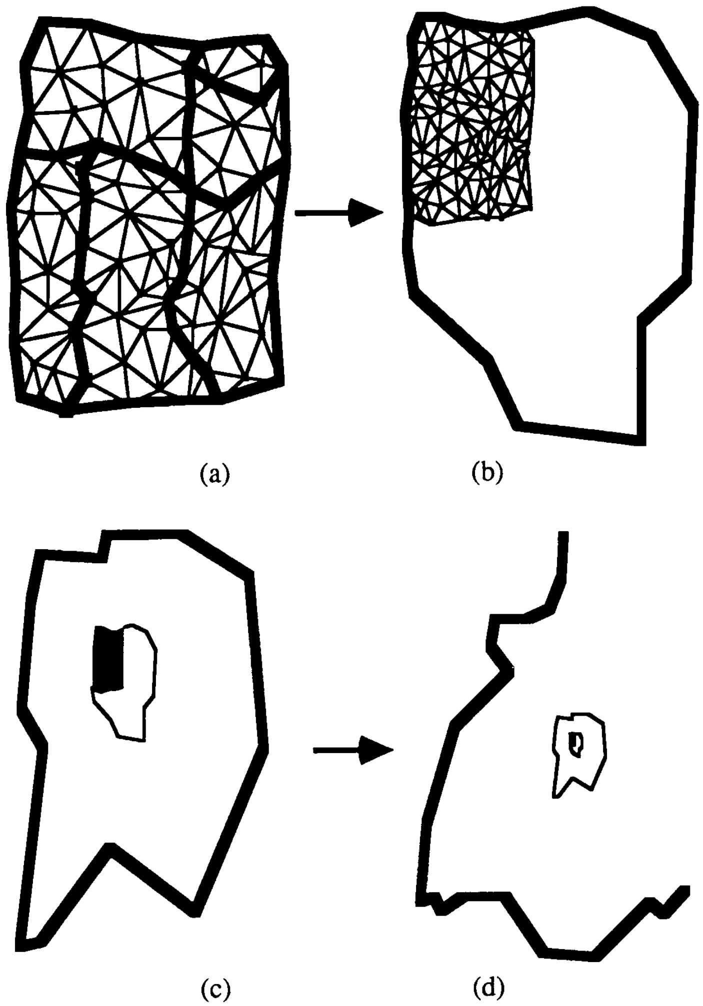


Figure 7: Parcels, towns, counties, and states consisting of an increasing number of topological cells.

Let us now look at two possible queries. For the first query, the level of abstraction of the representation is adequate. The second query looks at the world at a higher level of abstraction. We will see how in this latter case, query processing is considerably slowed down by irrelevant information.

The first query "Through which parcels does power line 235 run?" deals with some of the smallest objects of the example GIS. Each parcel is composed of only a small number of cells. The query asks for the parcels which have component cells in common with power line 235. Independent of which search strategy is chosen, the search time depends on the number of cells that have to be inspected. As the number of cells per parcel is small, the query can be processed efficiently.

The second query demonstrates how performance problems arise when the topological representation contains a lot of irrelevant information. The query "Through which states does 1-95 run?" deals with some of the largest objects of the example GIS. They are represented by a very large number of component cells because all topological relations between the smaller objects are included in the representation. For every state of interest, a very large number of cells has to be inspected which makes the query process very slow. This demonstrates how performance for queries degrades in case of a single topological representation at small scales. Response times rapidly get unacceptable for interactive GIS use (Meixler 1985).

The following example demonstrates that the described efficiency problem cannot easily be solved with faster hardware. Let us assume that the United States were homogeneously covered by small objects like parcels of an approximate size of 100 x 300 ft and that they are each represented by only a single 2-cell. An average state would then consist of approximately 60 million 2-cells and the whole United States of about 3 billion 2-cells. Let the representation further include townships and climatic zones; these zones being larger than states. In a first approximation, the processing time for topological queries can be assumed to be a linear function of the number of cells inspected. For humans, the questions "which of 10 given parcels lie in a given township?" and "which of 10 states lie in a given climatic zone?" are approximately equivalent in complexity. However, it takes a GIS about 60 million times longer to answer the second question since states in the average contain 60 million times the number of 2-cells of parcels and all 2-cells of parcels and states, respectively, have to be inspected in the two queries. This difference in response is unexpected to the user. If we assume that it takes about 10 milli-seconds to answer the first question, it would take about 170 hours for the second question. Even with considerably faster hardware there would always be a very noticeable difference between the response times of the two queries.

In a single topological representation which includes a large number of object classes, the number of cells per object is generally high. This is normally the case in a GIS where data is shared among several user groups since this will generally require information on more themes. Therefore, performance problems occur particularly when information about large objects is extracted. *GIS based on a single topological representation are therefore currently limited in their range of scales and the variety of user groups that can be supported*.

### 3.2 Basic Concept of the Proposed Solution

Multiple topological representations (Bruegger 1989b, 1989c, 1990) as proposed in this paper consist of several single topological representations featuring different levels of abstraction. The component representations are linked by hierarchies over cells of different levels. Multiple topological representations allow for an efficient extraction of information at different levels of abstraction. Figure 8 illustrates the basic design of a multiple topological representation.

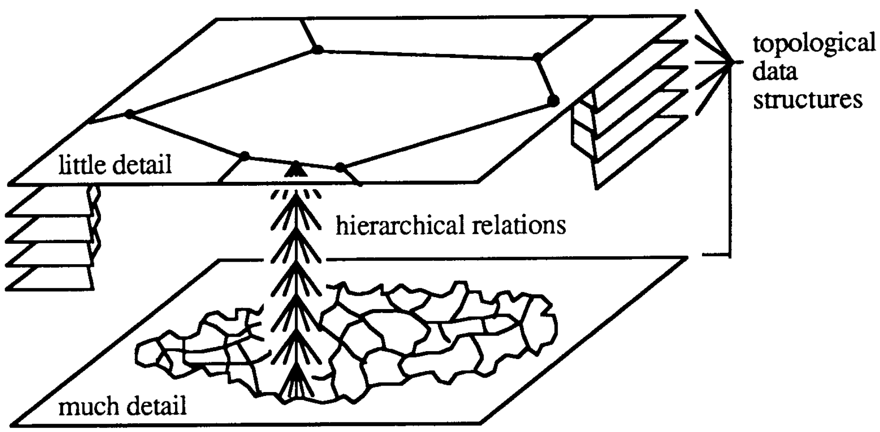



Figure 8: Structure of a multiple topological representation.

The *extraction* of information from multiple topological representations can be efficient at any level of abstraction. However, the *insertion* of data takes more time than in a single topological representation since in the redundant structure, data about objects has to be inserted in several representations instead of just one. Under the assumption that the extraction of information is much more frequent than insertions and updates, multiple topological representations promise to solve the problem described above.

## 4. LARGEST HOMOGENEOUS CELLS (LHC's)

LHC's are the foundation of multiple topological representations. In particular, a special property of LHC's guarantees hierarchical links between cells of different representations. With different cell concepts, these links would not be hierarchical.

LHC's form the minimal complete topological representation. A first subsection defines the term complete representation. On this basis, we define LHC's and describe how objects can be modelled as sets of LHC's. LHC's are not yet another kind of cell. Indeed, we show that they are the common denominator of all possible kinds of cells and can therefore easily be modelled in topological representations that use different cell concepts. In this way, LHC's interface between the integrating structure of a multiple topological representation and its component representations which use cell concepts different from LHC's.

### 4.1 Complete Topological Representations

We call a topological representation complete, if it models all topological properties and relations captured by the framework (Egenhofer 1990) reviewed in section 2.4. We can express completeness by the following set of criteria:

(1) *Distinction of interior and boundary of objects:*
A complete representation must distinguish the interior, the boundary, and the exterior of objects.

(2) *Derivation of Intersections:*
It must be possible to derive intersections of objects and of their parts.

(3) *Notion of Connectivity:*
The connectivity of objects and their parts must be determinable. In case of disconnected objects, the separated parts must be countable.

(4) *Dimension of Objects:*
The dimension of objects must be accessible.

### 4.2 LHCs as Minimal Complete Representation

In point set topology, n-dimensional Euclidian space is modelled by an infinite number of points. This is not appropriate for implementations on finite computers. Combinatorial topology (Spanier 1966) uses a finite number of cells as basic entities to model Euclidian space. These cells are point sets. Of all the infinitely many possible sets of points that can be formed, only a small number are chosen as cells; namely a set of cells that satisfies the criteria for a complete representation. This means that the point set equivalents of cells must distinguish between the interior, the boundary, and the exterior of objects (distinction requirement). This is only the case if cells representing parts of the interior are distinct from cells representing parts of the boundary or the exterior.

We define *largest homogeneous cells* (LHCs) as the largest possible cells satisfying the *distinction requirement*. Representations based on LHC's use the minimal number of cells necessary. LHC's are therefore the basis *of minimal topological representations.*

Figure 9 shows the LHC's of a representation for a single object. Cl, for example, corresponds to the set of all points belonging to the interior. In other words, all points contained in Cl are "owned" by the interior of the object shown.

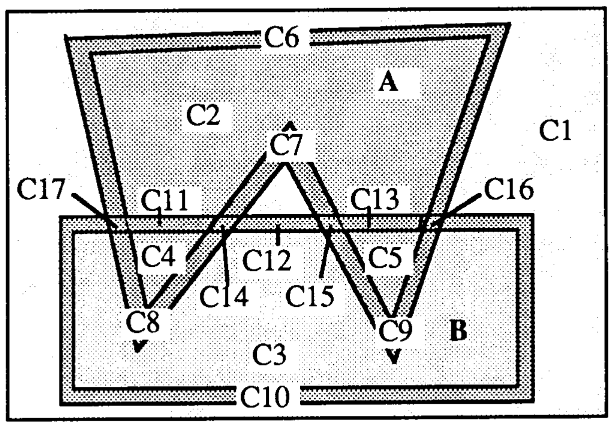


Figure 10: LHC's created by two objects.

If more than one object is involved, a point can have several owners. LHC's are constructed by grouping all points with the *same owners* into *connected* sets. They are called homogeneous because all component points have the same owners. Figure 10 demonstrates this for the case of two objects A and B in a two-dimensional space. The interior, the boundary, and the exterior of both A and B are potential owners of points in the space. For example, all points in Cl are owned by the exteriors of A and B. The points of C7 are owned by the boundary of A and the exterior of B. The points of both C8 and C9 are owned by the inside of B and the boundary of A. Although all points of these two cells have the same owners, they form two distinct cells to satisfy the requirement of being connected.

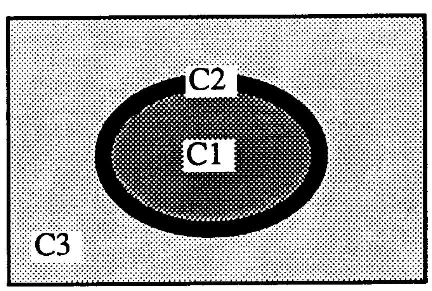


Figure 9: LHC's created by a single object.

The concept of LHC's is an extension of the concepts of *greatest common geographical units* proposed in (Peucker 1975) and *Geographical Tabulation Unit Base* (GTUB) in (Meixler 1985). The latter two concepts are restricted to polygons in the plane. In contrast, LHC's are defined for *any dimension of space* and for *any dimension of objects* (points, curves, polygons, volumes, etc.).

Further, LHC's also structure the boundary of objects. For example, in the plane, a piece of the intersection of two polygons can be modelled by a 1 -LHC and the intersection of three polygons by a 0-LHC. In contrast, the greatest common geographical units and GTUB's correspond to only 2-LHC's and leave the boundaries of objects unstructured.

It follows from the definition that LHC's *are always disjoint*. Further, there are no points which do not belong to a cell. All points of the space are distributed into disjoint point sets (cells) which thus form a *partition* (Gill 1976).

### 4.3 Objects as Sets of LHC's

Here we study how *objects* are represented as sets of LHC's. For a complete representation, the whole object, its boundary and its interior must be available as separate entities. This can be achieved, for example, by representing the interior and the boundary as a set of LHC's (see figure 11). The *object representation* then consists of its dimension and the *relations between the object and its LHCs.* Since all possibilities are conceptually equivalent, we do not limit generality by using this representation in the paper.

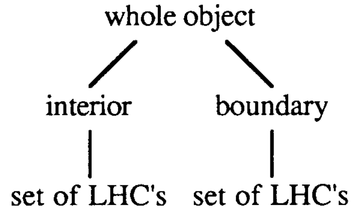



Figure 11: Object representation consisting of a separate boundary and interior.

Since LHC's satisfy the distinction requirement, *intersections* of parts of objects can always be found by set intersection yielding the common cells. The *connectivity* of sets of cells follows from boundary and co-boundary relations between cells (Egenhofer 1989b). Thus, a *topological representation* consists of a *collection of objects* with their *dimensions*, *relations between each object and its composing cells*, and *relations between adjacent cells.*

### 4.4 LHC's based on Arbitrary Cell Structures

Arbitrary subdivisions of LHC's, called *implementation cells*, are also valid cells as they satisfy the criteria for a complete representation. Implementation cells have *special properties* which make them better suited for implementation in the computer. For example, certain implementation 2-cells are not allowed to have holes, or 2-cells of simplicial. complexes (Frank 86) must be triangular. Although the number of implementation cells is larger than that of LHC's, they can lead to more efficient implementations as simpler and more efficient algorithms can be used and less special cases have to be considered. LHC's can be formed of any kind of implementation cell. We therefore call LHC's the *common denominator* of all cell concepts.

Since LHC's are the foundation of *multiple topological representations*, they have to be modelled explicitly. However, multiple topological representations can still use *any kind of implementation cells* to take advantage of their special properties. LHC's are modelled as sets of implementation cells, i.e., they are treated like objects in the implementation cell environment (see figure 12). This is always possible, since all implementation cells are subdivisions of LHC's. LHC's can be seen as an interface between an arbitrary implementation cell structure and the MTR-specific software.

Objects in an MTR are sets of LHC's. *Intersections* of the interior and the boundary of objects with other objects can be found by checking for LHC's belonging to both object parts. In order to answer questions about connectivity of object parts, they have to be mapped to sets of implementation cells since single topological representations support operations to answer these questions for sets of implementation cells. Figure 12 visualizes the object representation in the case where implementation cells are used.

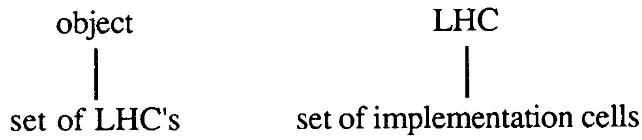



Figure 12: Representation of an object using implementation cells.

## 5. THE STRUCTURE OF MULTIPLE TOPOLOGICAL REPRESENTATIONS

This section develops the Multiple Topological Representations (MTR) concept. Each representation in an MTR has a different notion of relevant and irrelevant information. This allows selection of an appropriate representation for the extraction of topological information, i.e. a representation that does not contain a lot of irrelevant information. In this way, reasoning processes can be performed more efficiently as they are not slowed down by irrelevant information.

### 5.1 Multiple Topological Representations

*Multiple Topological Representations* (MTR) integrate several (single) topological representations into a consistent structure. Each of the topological representations is the result of a specific abstraction process which filters out irrelevant aspects of the world (Gentner 1983).

We saw in section four that a topological representation is determined by its domain, i.e., the set of objects represented. In the abstraction process, the objects which are not relevant are filtered out, such that the domain contains only relevant objects. The topological representation then organizes the relevant objects and their topological relations.

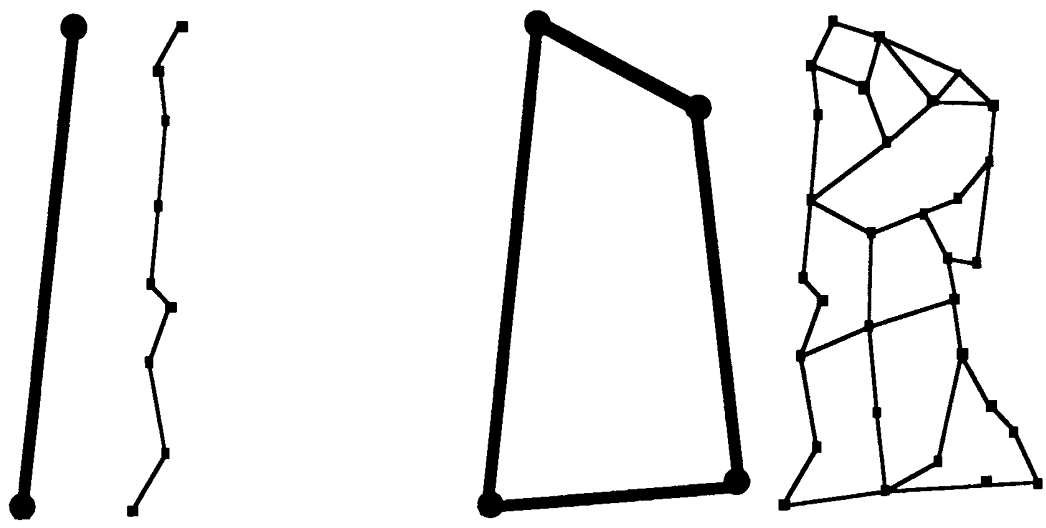
The decision on the relevance of an object is made according to both the scale of the geometrical description and the purpose of the representation. At small scales, small objects are normally considered irrelevant. For example, a small scale map generally does not show houses or parcels. However, considering the purpose of the representation, small objects can still be relevant. For example, a political map of the world can show some major cities in spite of their relatively small size.

### 5.2 Linking Representations of an MTR

To keep an MTR consistent, its component representations have to be linked with each other. A *link consists of relations connecting different representations of the same point set*. The relations connect a set of LHC's of one topological representation with a set of LHC's of another one. This guarantees consistency in the MTR and allows propagation of changes from one representation to another.

In the general case, relations of a link connect m LHC's of one representation with n LHC's of another one (*m:n -relations*). Whenever two representations have an object in common, such m:n-relations exist. The numbers m and n are integers and depend on the cell partitions.

In this paper we are particularly interested in the special case of 1:m-relations. We will show that *1:m-relations* are sufficient for many applications of MTR. Figure 13 shows examples of such relations between LHC's. Figure 13a shows a higher level 1 -cell and two 0-cells together with their lower level equivalents. The equivalent of the 0-cells are again 0-cells, i.e., they are related by 1: 1 -relations. The higher level I -cell corresponds to a set containing I- and 0-cells. Figure l3b shows a higher level 2-cell with its boundary consisting of 1 - and 0-cells. The left I -cell is the same as in figure 13a. The higher level 2-cell corresponds to a set of 2-, 1-, and 0-cells on the lower level.



(a) (b)

Figure 13: Higher level cells correspond to sets of lower level cells.

The special case of 1:m-relations occurs when one of the domains is a subset of the other. Then a single LHC of the higher level representation corresponds to a set of LHC's on the lower level. Figure 14 shows how the LHC's A, B, and C of the higher level representation (a) split up into several LHC's on the lower level (b) which contains more objects.

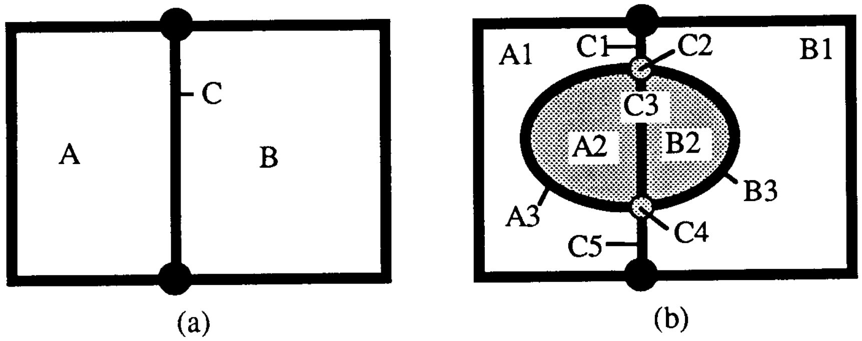



Figure 14: The LHC's of the lower level representation (b) form a refinement of those of the higher level representation (a)

If the link consists only of 1:m-relations between LHC's, then we say that the cell partition of the lower level forms a *refinement* (Gill 1976) of that of the higher level, i.e., each of the lower level LHC's is totally included in exactly one higher level one. To require a refinement is actually sufficient to guarantee 1:m-relations.

Implementation cells in contrast to LHC's do not necessarily form a refinement. Figure 15 demonstrates this for the case of simplicial complexes (Frank 86). The thick lines and points denote LHC's, while the thin lines split them up into simplicial implementation cells. The lower level (b) does not refine the higher level (a). For example it is impossible to find the lower level equivalent of the 1-cell A-C in (a): If we try to map it to A-E + E + E-C in (b), then E and E-C are included in the two higher level cells A-C and B-C. The same goes for any other possibility of connecting A and C.

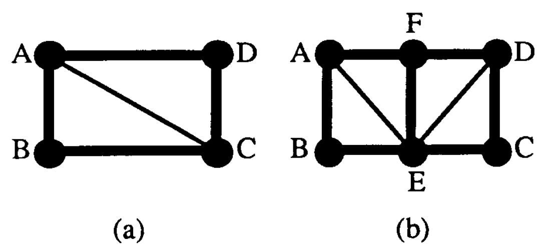



Figure 15: Implementation cells in a higher (a) and lower level representation (b) generally do not form a refinement.

The link between a higher and a lower level representation, in case of a refinement, consists of a set of only 1:m-relations between LHC's. Figure 16 shows such a link between two 1-dimensional topological representations. 0-cells are shown as disks and ellipses. The I:m-relations are visualized by thin lines.

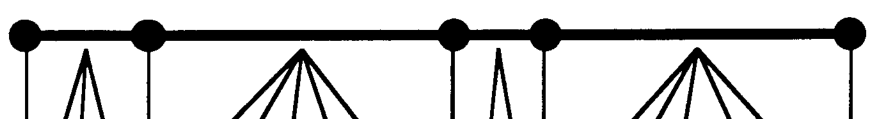


Figure 16: Link of two 1-dimensional topological representations.

### 5.3 Links as Order Relations

Since the refinement relation of a link orders the related representations into a higher and a lower level one, it can be seen as an order relation (Gill 1976). To express that the lower level *a* is linked with a higher level *b*, we will write *a £ b*. We will say that *a* refines *b* and that *a* is more refined than *b*. The formal requirements for an order relation, i.e., reflexivity, antisymmetry, and transitivity (Gill 1976), are satisfied by the refinement relation.

In an MTR, all 1:m-relations have to be modelled. Relying on the *transitivity* of the refinement relation, many links follow implicitly from the explicitly modelled links. The transitivity rule states that if *a £ b* and *b £ c* then also *a £ c* must hold. Figure 17 illustrates how to derive implicit links: Representation r2 is linked with both representations below.

However, only the links connecting r2 with rl and rl with r0 are modelled explicitly. The link connecting r2 with r0 can be derived. For example, the single 1 -cell in r2 is equivalent to a set of cells in r0. (Relations between the 1-cell and nodes are omitted in the drawing). It is straight-forward to derive this 1:m-relation from the 1:m-relations between the explicitly linked representations.

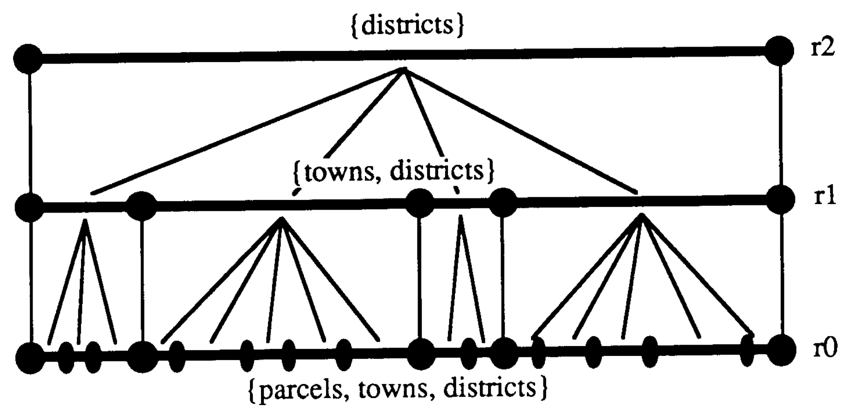


Figure 17: Hierarchical relations between cells over two levels of abstraction.

### 5.4 Strictly and Partially Ordered MTR

Order structures are important for many applications of spatial data handling (Kainz 1988). For example, partially ordered sets have been proposed for the organization of spatial inclusion (Kainz 1988, Greasley 1988, Saalfeld 1985).

Based on the structure of order relations in an MTR, we can distinguish between two major classes of MTR*: strictly ordered MTR and partially ordered MTR*. This paper concentrates on strictly ordered MTR as they are simple and allow interesting applications. In a strictly ordered MTR, a refinement relation exist between any pair of component representations. In a partially ordered MTR there exist some pairs of component representations which cannot be compared by an order relation, i.e., whose cell partitions are not related by refinement.

Figure 18 shows examples for both kinds of MTR. The strictly ordered MTR in figure 18a contains the complete set of 1:m-relations between equivalent point sets of different representations. This holds for strictly ordered MTR in general. Strictly ordered MTR do not contain any m:n-relations between their levels. The 1:m-relations form a *hierarchy over LHC's.* This hierarchical structure allows views of the same spatial unit at different levels of abstraction. Strictly ordered MTR share this property with other hierarchical data structures such as striptrees (Balard 1981) or quadtrees (Samet 1984).

The partially ordered MTR in figure l8b consists of two strict orders d, e, f and d, g, h. However, there are no order relations connecting e and g, e and h, f and g, f and h.

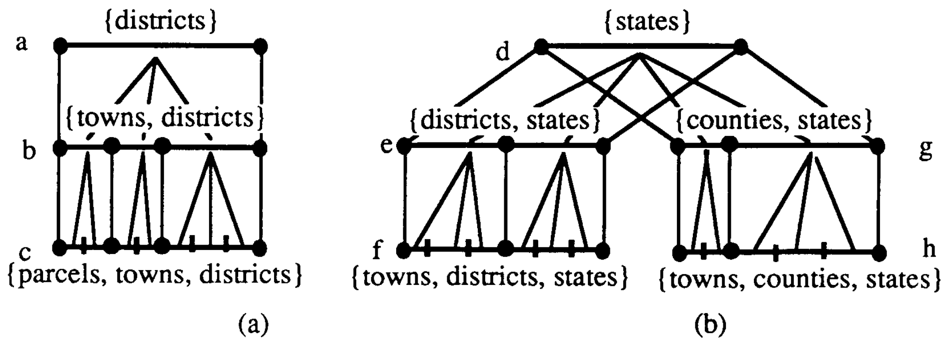



Figure 18: A strictly ordered MTR (a) and a partially ordered MTR (b).

Let us look at the 1:m- and m:n-relations of this partially ordered MTR. The representations f, e, d and h, g, d form strictly ordered MTR's, respectively. Therefore, all 1:m-relations within these strictly ordered MTR's are modelled. The following two examples demonstrate that, in general, some m:n-relations are implicitly modeled while others are not. The m:n-relation between the representations of a state in h and e can be derived from the 1:m-relations connecting h with g, g with d, and d with e. The m:n-relations between the representations of a town in h and f cannot be derived since towns do not exist on the higher levels.

### 5.5 Objects in a Strictly Ordered MTR

The person administrating an MTR has to determine in which representation objects shall exist and of which cells the objects are composed. In the case of a strictly ordered MTR, it is straight forward to describe a given order in the form of a *schema* that determines which objects appear on which level. Because of the strict order on the topological representations, we can enumerate them and speak of levels of abstraction. If we assume that the refinement relations originate from a subset relation between the domains, then all objects must exist in the lowest level representation of the MTR. They further exist on all higher levels up to some maximum level.

Let our strictly ordered MTR contain the representations r0, rl, r2, ..., rn, where r0 denotes the most refined representation and rn the one with least detail. To guarantee refinement relations, let the domains di of ri be related by subset relations, i.e., di is a subset of di - 1. For every object class, the schema specifies the maximum level on which the object appears. We call this the *level of importance* of the object as the more important objects generally appear on higher levels. The level of importance in the schema is given by an integer that is greater or equal to 0. Figure 19 gives an example schema for our hypothetical MTR.

**4:** countries
**3:** states, interstates
**2:** districts, counties
**1:** towns, primary roads
**0:** utility lines, houses, parcels

Figure 19: Schema for a strictly ordered MTR.

Figure 20 illustrates the strictly ordered MTR that is described by the schema of figure 19.

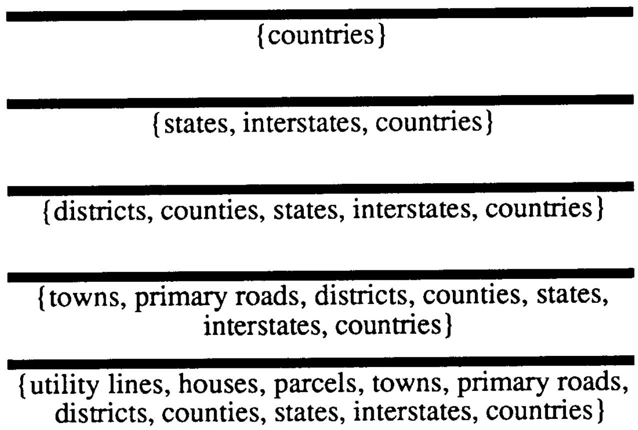



Figure 20: MTR that results from the schema of figure 19.

In a strictly ordered MTR, objects are represented on several levels as sets of LHC's but it is not efficient to define the same object multiple times. Here, we show how the 1:m-relations allow to get by with a single object definition.

If we define an object as a set of LHC's on a certain level of an MTR, then the links imply a representation of the object in all representations below. This is shown in figure 21. Penobscot County is defined as a set of LHC's on level 3. Level 3 is the level of importance for counties. The 1:m-relations between LHC's allow us to directly map this set of LHC's to any lower level, i.e. Penobscot County is represented on all these levels implicitly.

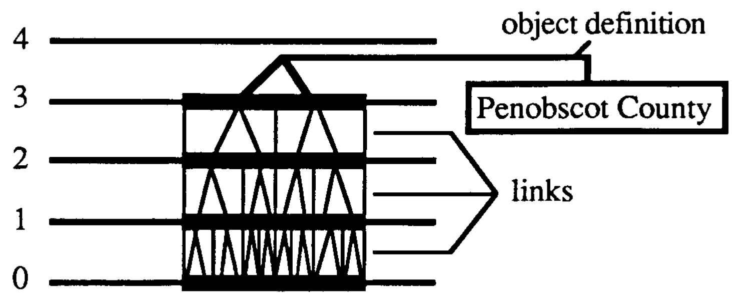



Figure 21: Object definition in a strictly ordered MTR.

We showed how we can get from an object to a set of LHC's in any representation where the object exists. In addition, the 1:m-relations and the object definitions allow discovery of objects containing a certain LHC.

**6. CONCLUSIONS**

A strictly ordered MTR with the schema of figure 19 solves the performance problems described in section three. The query "Through which states does interstate I-95 run?" can be executed on level 3--the highest level where interstates and states exist. It organizes only topological relations between states, interstates, and countries with countries being the only objects irrelevant to the query. As they are much larger than states, they do not increase the number of cells per object significantly compared to a representation containing states and interstates only. This holds independent of whether the objects nest inside each other or not. As in the given case, all states are totally included in the United States, the number of cells per object is not increased at all in this case. The level of abstraction of representation r3 is therefore optimal for the query.

This application example of a strictly ordered MTR demonstrates how the extraction of topological information is more efficient in MTR's due to the different abstraction levels supported. The effort for finding an adequate representation in an MTR is negligible compared to the cost of query execution.

Multiple Topological Representations (MTR) integrate several topological representations into a consistent structure. Different representations result from different abstractions reflecting which properties of the world are relevant. In single representations, queries are slowed down by a large amount of irrelevant data. An MTR offers representations at adequate levels of abstraction which allow efficient reasoning over a large range of scales.

MTR are based on the concept of largest homogeneous cells (LHC's). The concept of MTR is *dimension independent*, i.e., objects and embedding spaces of arbitrary dimension can be modelled. This follows from the definition of LHC's as point sets and the fact that point sets are dimension independent. MTR's can organize *arbitrary configurations of objects*. Although LHC's in strictly ordered MTR are related hierarchically, this does not require a hierarchical order on objects. *Implementations of MTR* can be *based on any sound structure of implementation cells*. This is achieved by modelling LHC's as sets of implementation cells in an MTR.

One of the major goals of GIS technology is to save costs for data acquisition and maintenance by sharing data among several user groups. Different user groups have different requirements for the scale and the level of abstraction of their data. Typically, a GIS supporting multiple user groups also organizes a large number of object classes. Single topological representations neither handle a large range of scales nor a large number of object classes. Multiple topological representations on the other hand are designed to handle this variety. The flexibility of multiple views facilitates administrative and political problems of data sharing and promises increased cost savings in data acquisition and maintenance.

The application of the proposed multiple topological representation scheme is not limited to GIS. Together with the extension to multiple geometrical representations it is of interest for any branch of science and engineering that models spatial properties. In particular, its ability to model multiple levels of detail will be important in fields like CAD/CAM and robotics.

## 7. OUTLOOK

Extracting information from an MTR involves the selection of an adequate level of abstraction. Future research will have to address this problem to find efficient ways of query execution.

We are currently working on an extension of a *partially ordered MTR* called *Multiple Geometrical Representations* (MGR). These are hierarchical representations of both metrical and topological properties, based on a complete topology organizing objects of arbitrary dimension.

Metrical aspects of geometry describing positions and shapes are not independent of topological aspects. Consequently, geometrical representations use a *topological structure to support the organization of metrical data*. Experiences with geometrical representations using incomplete topological information show that a complete modelling of topological properties is crucial. "Spaghetti" structures (Peuker 1975), for example, imply the simple topological notion of 1- and 0-cells. This does not allow a complete representation of topology and causes consistency problems. Modem GIS use complete topological representations based on cells to organize the metrical data (Herring 1987).

Hierarchical representations of metrical properties offer multiple levels of abstraction for the shape and position of objects. Strip trees (Ballard 198 1) are the best known example for such hierarchical representations in the vector world. They are based on an equally simple notion of topology as spaghetti representations. They are further restricted to modelling curves and do not support the hierarchical representation of objects of different dimensions. Multiple geometrical representations combine the multi-resolution approach with a complete topological representation.

A second planned extension of MTR will allow to model *dimension changes*. It is based on a definition of relevant topological information. With this extension, the same representation offers views of objects in different dimensions. For example, at a lower level of abstraction, cities and roads can be modelled as two dimensional objects. On a higher level, cities can be represented as points and roads as lines. High level questions about routing can then be answered very efficiently.

## References


Abler, R., 1987, The National Science Foundation National Center for Geographic Information and Analysis. International Journal of Geographical Information Systems, 1, 303.

Ballard, D. H., 1981, Strip-Trees: A Hierarchical Representation for Curves. Communications of the ACM, 24, 3 10.

Behnke, H. et al., (editors), 1974, Fundamentals of Mathematics (Cambridge: MIT Press).

Boudriault, G., 1987, Topology in the TIGER File. In: N. R. Chrisman (ed.), Auto Carto 8 Proceedings, Baltimore, MD.

Broome, F. R., and Meixler, D. B., 1990, The TIGER Data Base Structure. Cartography and Geographic Information Systems, 17, 39.

Bruegger, B. P. et al, 1989a, Research Topics on Multiple Representations. In: Barbara P. Buttenfield and Joseph S. DeLotto (eds.), Multiple Representations - Scientific Report for the Specialist Meeting. National Center for Geographic Information and Analysis, Department of Geography, SUNY at Buffalo. Buffalo, NY.

Bruegger, B. P., and Frank, A. U., 1989b, Hierarchies over Topological Data Structures. In: ACSM-ASPRS Annual Convention, Baltimore, MD.

Bruegger, B. P., and Frank, A. U., 1989c, Hierarchical Cell Complexes. 12th Annual Applied Geography Conference, Binghamton, NY.

Bruegger, B. P., and Frank, A. U., 1990, Hierarchical Extensions of Topological Data Structures. FIG, XIX International Congress, Helsinki, Finland.

Buttenfield, B. P., and DeLotto, J. S. (editors), 1989, Multiple Representations -Scientific Report for the Specialist Meeting. National Center for Geographic Information and Analysis, Department of Geography, SUNY at Buffalo, Buffalo, NY.

Corbett, J. P., 1979, Topological Principles of Cartography. Technical Report 48. Bureau of the Census, Department of Commerce.

Croom, F. H., 1989, Principles of Topology (Philadelphia: Saunders College Publishing).

Egenhofer, M. J., et al., 1989a, A Topological Data Model for Spatial Databases. In: A. Buchmann and 0. Gunther (eds.), Symposium on the Design and Implementation of Large Spatial Databases, Santa Barbara, CA.

Egenhofer, M. J., 1989b, A Formal Definition of Binary Topological Relationships. In: Litwin, W., and Schek, H.-J., (editors), Third International Conference on Foundations of Data Organization and Algorithms (FODO), Paris, France, (New York: Springer-Verlag).

Egenhofer, M. J., and Franzosa, R. D., 1990, Point-Set Topological Spatial Relations. Submitted for publication.

ETAK, 1988, Etak MapEngine: Programmer's Guide. Etak, Inc., 1430 O'Brien Drive, Menlo Park, CA 94025.

Frank, A. U., and Kuhn, W., 1986, Cell Graph: A Provable Correct Method for the Storage of Geometry. In: D. Marble (ed.), Second International Symposium on Spatial Data Handling, Seattle, WA.

Frank, A. U., 1987, Overlay Processing in Spatial Information Systems. In: N. R. Chrisman (ed.), Auto Carto 8 Proceedings, Baltimore, MD.

Frank, A. U., Egenhofer, M. J., and Kuhn, W., 1990, Computational Topology: Data Structures and Algorithms. Submitted for publication.

Gentner, D., and Stevens, A. L., (editors), 1983, Mental Models (Hillsdale: Erlbaum Associates).

Gill, A., 1976, Applied Algebra for Computer Sciences (Englewood Cliffs: Prentice-Hall).

Greasley, 1., 1988, Data Structures to Organize Spatial Subdivisions. In: ACSM-ASPRS Annual Convention, St. Louis, Missouri.

Herring, J. R., 1987, TIGRIS: Topologically Integrated Geographic Information System. In: N. R. Chrisman (ed.), Auto Carto 8 Proceedings, Baltimore, MD.

Jackson, J. P., 1989, Algorithms for Triangular Irregular Networks Based on Simplicial Complex Theory. In: ACSM-ASPRS Annual Convention, Baltimore, MD, March 1989.

Jones, C. B., and Abraham, 1. M., 1986, Design Considerations for a Scale-Independent Cartographic Database. In: Proceedings of the 2nd International Symposium on Spatial Data Handling, Seattle, WA.

Kainz, W., 1988, Order Structures in Geographic Information Systems, Ph.D. thesis (in German). Technical University of Graz, Austria.

Kinnear, C., 1987, The TIGER Structure. In: N. R. Chrisman, (ed.), Auto Carto 8 Proceedings, Baltimore, MD.

Mark, D. M., 1989, Multiple Views of Multiple Representations. In: Barbara P. Buttenfield and Joseph S. DeLotto (eds.), Multiple Representations - Scientific Report for the Specialist Meeting. National Center for Geographic Information and Analysis, Department of Geography, SUNY at Buffalo, Buffalo, NY.

Meixler, D., and Saalfeld, A., 1985, Storing, Retrieving and Maintaining Information on Geographic Structures: A Geographic Tabulation Unit Base (GTUB) Approach. In: Auto Carto 7 Proceedings, Washington, DC.

Munkres, J., 1966, Elementary Differential Topology (Princeton: Princeton University Press).

National Center for Geographic Information and Analysis, 1988, Summary of Research Plan. University of California - Santa Barbara, State University of New York -Buffalo, University of Maine.

NCDCDS, 1988, The Proposed Standard for Digital Cartographic Data. The American Cartographer, 15, 23.

Peckham, J., and Maryanski, F., 1988, Semantic Data Models. ACM Computing Surveys, 20, 153.

Peucker, T. K., and Chrisman, N., 1975, Cartographic Data Structures. The American Cartographer, 2, 55.

Saalfeld, A., 1985, Lattice Structures in Geography. Auto Carto 7 Proceedings, Washington, DC.

Samet, H., 1984, The Quadtree and Related Hierarchical Data Structures. ACM Computing Surveys, 16, 187.

Spanier, E. 1966, Algebraic Topology (New York: McGraw-Hill Book Company).
White, M., 1979, A Survey of the Mathematics of Maps. Auto Carto IV Proceedings, Washington, DC.