ONSRUD

Bruegger, B. P., and A. U. Frank. "Hierachical Extensions of Topological Datastructures." Paper presented at the 19th Congress of International Federation of Surveyors, Helsinki, Finland, 10.-19. June 1990 1990.

FIG

XIX CONGRESS

INTERNATIONAL FEDERATION OF SURVEYORS
FÉDÉRATION INTERNATIONALE DES GÉOMÈTRES
INTERNATIONALE VEREINIGUNG DER
VERMESSUNGSINGENIEURE

10.–19.6.1990
HELSINKI FINLAND

PAPER. EXPOSÉ. BERICHT.

Land Information Systems
Systèmes d'Information du Territoire
Landinformationssysteme

3
COMMISSION

FIG XIX INTERNATIONAL CONGRESS
HELSINKI, FINLAND, 1990

HIERARCHICAL EXTENSIONS OF TOPOLOGICAL DATASTRUCTURES*

EXTENSION HIERARCHIQUE DES STRUCTURES DE DONNEES TOPOLOGIQUES

HIERACHISCHE ERWEITERUNG VON TOPOLOGISCHEN DATENSTRUKTUREN

Bud P. Bruegger & Andrew U. Frank - U.S.A.

SUMMARY

Humans maintain multiple mental models to reason about the world. They use models appropriate to given situations in order to avoid getting lost in detail. GIS normally feature only a single representation for their reasoning process in a given domain. This paper presents multiple topological representations designed to improve reasoning efficiency of GIS.

RESUME

Les humanins employent multiple mentaux de la rélité pour leur raisonnement. Ils utilisent des modèles adaptés aux situations pour ne pas se perdre dans trop de détail. Les SIT normalement n'ont qu'une seule représentation d'une certaine domaine. Cette communication présente des multiples représentations topologiques qui rendent les déductions dans un SIT plus efficace.

ZUSAMMENFASSUNG

Menschen benützen mehrere Gedankenmodelle um Aussagen über die Welt zu machen. Sie verwenden den Situationen angepasste Modelle, um sich nicht in Details zu verlieren. GIS verfügen normalerweise nur über ein einziges Modell für einen bestimmten Zweck. In diesem Aufsatz wird eine mehrfache topologische Representation vorgestellt, die die Effizienz von automatischem Schliessen in GIS verbessert.

---

* This research is part of initiative 3 "Multiple Representations" of the NCGIA. The support from NSF for the NCGIA under grant No. SES 88-10917 is gratefully acknowledged.

1. INTRODUCTION

Humans use mental models to reason about the world. Models are abstractions of the world, representing only a few aspects in comparison to the huge amount of detail disregarded. We have a multitude of mental models emphasizing different aspects of the world [Mark 1989]. Such different models may represent the same real world object at many different levels of detail. In our reasoning process we select those 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 abilities to reason, i.e., to derive information from the stored data, is significantly inferior to human reasoning. One reason for this shortcoming is the lack of representations containing less detail.

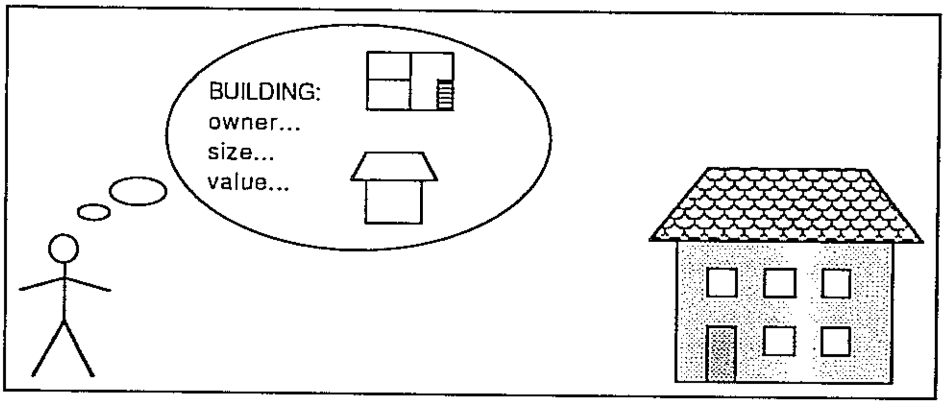


Figure 1: Humans use a multitude of mental models to reason about the infinitely detailed world.

In order to improve the reasoning capabilities of GIS we have to include multiple representations of the world in our systems. These representations have to be kept consistent and the systems have to know which representation is most adequate for a required reasoning process. The Research Initiative 3 of the National Center for Geographic Information and Analysis [Abler 1987][NCGIA 1988] studies these problems in more detail [Buttenfield 1989].

This paper describes multiple topological representations (MTR), a special case of multiple representations. The theory of single topological representations that are parts of an MTR has been described by Corbett [Corbett 1979], White [White 1984], Frank and Kuhn [Frank 1986], and implementation aspects by Jackson [Jackson 1989] and Egenhofer et al. [Egenhofer 1989b]. Single topological representations have found application in some models of modern GIS [Kinnear 1988][Boudriault 1988][Herring 1988]. Multiple topological representations are based on hierarchies over topological cells [Bruegger 1989b] and are a special case of multiple representations [Bruegger 1989a].

The second chapter gives an overview of single and multiple topological representations. The third chapter demonstrates how spatial objects are organized in a multiple topological representation. Chapter 4 describes the process of inserting multiple representations of an object and how to find an adequate representation for object retrieval.

2. TOPOLOGICAL REPRESENTATIONS

This chapter shows which aspects of spatial objects are organized in topological representations and some properties of single topological representations are examined. They are important components of multiple topological representations (MTR). In a third paragraph, the structure of multiple topological representations is summarized from [Bruegger 1989b].

2.1 Topology of Spatial Objects

Today's tools of modelling and software engineering [Brodie 1984][Peckham 1988] offer sophisticated ways of expressing the structure and behavior of spatial objects in GIS. A model of a spatial object can be separated into a geometrical part and a semantic part.

spatial object = geometry + semantics

The geometrical part can again be separated into a topological, a metrical, and an order aspect [Egenhofer 1989a] [Kainz 1988].

object geometry = topology + metrics

Of course, the topological and the geometrical aspect are closely related.

Topological representations organize only the topological aspects of the spatial object. The topological aspects of a spatial object describe those properties that are invariant under topological transformations, such as translation, rotation, and scaling. All other aspects of a spatial object can be accessed from its topology.

2.2 Single Topological Representations

Single topological representations are the components of MTR. They organize the topology of the spatial objects in the form of cells. The cells of such a representation can be constructed by intersecting every object with every other object [Bruegger 1989b]. Point objects and intersection points, or *nodes*, are called *0-cells*. The parts of linear objects between intersection nodes and the boundaries of areal objects between nodes are called *edges* or *1-cells*. The areas between the edges are called *faces* or *2-cells*.

The topological aspect of spatial objects can now be described as a set of cells.

topological object = set of cells

Oriented topological objects are chains of cells where the cells have an orientation. They are used for example, to express the boundary of an areal object separating an *inside* on the one side of the chain from an *outside* on the other.

oriented topological object = chain of cells

Topological representations provide a complete set of operations to deal with cells, chains and sets of cells.

2.3 Multiple Topological Representations (MTR)

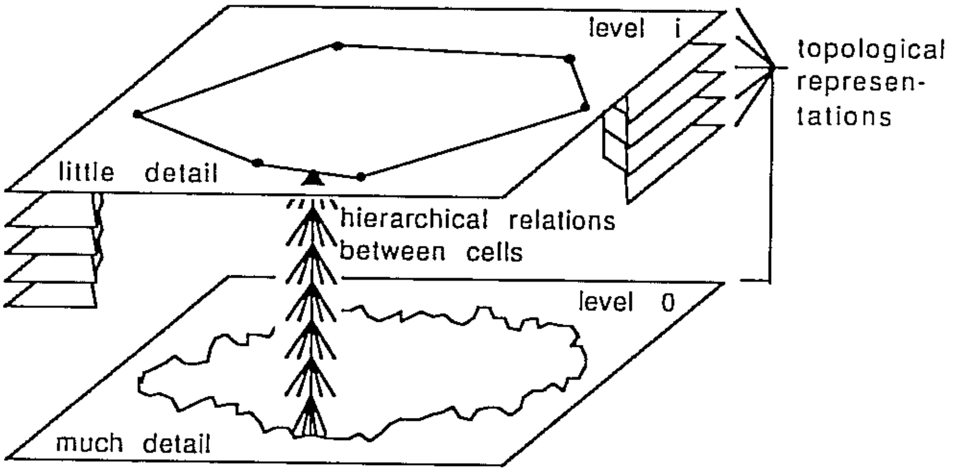


Figure 2: The multiple topological representation, a stack of topological representations with hierarchical relations between cells interconnecting levels.

Figure 2 demonstrates that an MTR consists of a stack of topological representations. Each representation is called a representation level. The cells of adjacent levels are related hierarchically, i.e., there is an antisymmetrical and transitive relation R between cells of adjacent levels. At every level, the MTR provides all the operations of a single topological representation. In addition, the operations *consists_of* and *contained_in* exist which map cells, chains, or sets of cells onto their corresponding structure on the level below and above, respectively. Higher level cells are sets of cells on lower levels. They can, therefore, be treated with the same operations provided for topological objects.

The base level organizes the topology of all classes of spatial objects. Higher levels organize only a subset of the object classes of the level below. These subsets are given by a schema that assigns an *importance level*, i.e. the maximal level in the MTR where they still exist, to every object class. Typically, classes of small objects, e.g., parcels, are less important than classes of large objects, such as states.

On higher levels of the MTR, objects of small importance level are considered detail. Therefore, topological relations of the small objects among themselves and with other objects are left out on higher levels. The decrease of the number of cells demonstrates the decrease of detail.

3. SPATIAL OBJECTS IN MTR

In a previous paper, hierarchies over topological cells were introduced [Bruegger 1989b]. The remainder of this paper looks more closely at how to organize the topological aspects of spatial objects using this hierarchy. Firstly, a description of how users define the object topology and how it can be translated to an internal, cell-oriented definition. Secondly, a detailed description of the internal definition is presented.

3.1 User Definition of the Object Topology

Users of GIS think of objects as a whole and not just of their topological aspects. Accordingly, they communicate with the GIS in the language of semantic modelling or in a language dealing with all the aspects of geometry. Users normally do not know the meaning of cells.

The user has three possibilities to describe an object: (1) Describing the semantics separately and specifying the geometry by giving the metrics of the object; (2) Describing the semantics separately and specifying the geometry in terms of the geometry of other objects using set operators (e.g. union, intersection, difference); (3) Describing both the geometry and the semantics using abstraction mechanisms known from semantic modelling; i.e., *generalization*, *association*, and *aggregation* [Brodie 1984][Peckham 1988]. Geometrically these abstractions are equivalent to the union operator of sets; however, the semantic relationships between the objects are different. An example for such an object description is the association of the member Counties to a State. The geometrical part of the definition is obvious, but it has also semantic implications; for example, in the legal and political field.

These descriptions of the object geometry have to be translated to a cell-oriented definition of the topological aspects of the object. In case (1), all topological incidences between the new and all previously defined objects have to be found using the metric data. The incidences directly imply the cells. A detailed description of this process can be found in [Jackson 1989][Egenhofer 1989b]. In cases (2) and (3), the cells of the new object follow directly from the sets of cells describing the previously defined objects and the set operators connecting them.

3.2 Internal Definition of the Object Topology

The definition of the topology of a spatial object in the MTR is a set of component cells. Although an object of dimension d contains cells of the dimensions d down to 0, referencing only d-cells is sufficient. The component cells of lower dimension are implicitly contained. For example, knowing the component 2-cells of an areal object implies all the component 1- and 0-cells.

i:d-object denotes a d-dimensional object that exists up to level i. *Component* is a super class of *cells* and *objects*. An i:d-component is of dimension d and is defined on level i. Using this notation, an object can be defined as follows:

i:d-object = set of i:d-components

Note that parts can always be reduced to a set of cells. If objects are used in the definition, the updating of component sets after splitting component cells becomes easier. The component set can contain only elements of the same dimension. Because of the way the hierarchy of cells is built, higher level cells--and thus higher level objects--will always have parts falling outside of the object to define. Lower level cells on the other hand would lead to inefficient definitions which do not take advantage of the multiple topological representation. For consistency reasons the elements of the set are not allowed to overlap.

The internal and the user definition are only closely related in the case where the user describes the object in terms of objects of the same importance level. This will only rarely be the case as typically these definitions are objects of lower importance; e.g., a districts in terms of parcels. In case (3) where the object was defined using an abstraction mechanism, the user definition and the internal definition will be redundant. This redundancy is necessary in order to take advantage of the MTR.

All component cells in the set reside in the *level of importance* assigned to the object class. Using the *consist_of* operation this definition can be mapped down to any lower level. In the same way the *owner* operation returns all objects that contain the argument cell. The *maxLevel* operation returns the index of the highest level on which an owner object of the cell exists. These operations allow the user to access an object from every level below the definition level.

4. INSERTION AND RETRIEVAL OF OBJECTS IN MTR

The user definition of the geometrical aspects of an object can be translated to a set of base level cells. The internal definition refers to cells on the *level of importance* assigned to the object class. The definition has to be propagated from the base level up to the level of importance without introducing inconsistencies in the MTR. The first paragraph of this chapter describes this process. The second paragraph describes the possibilities of retrieving objects in an MTR.

4.1 Insertion

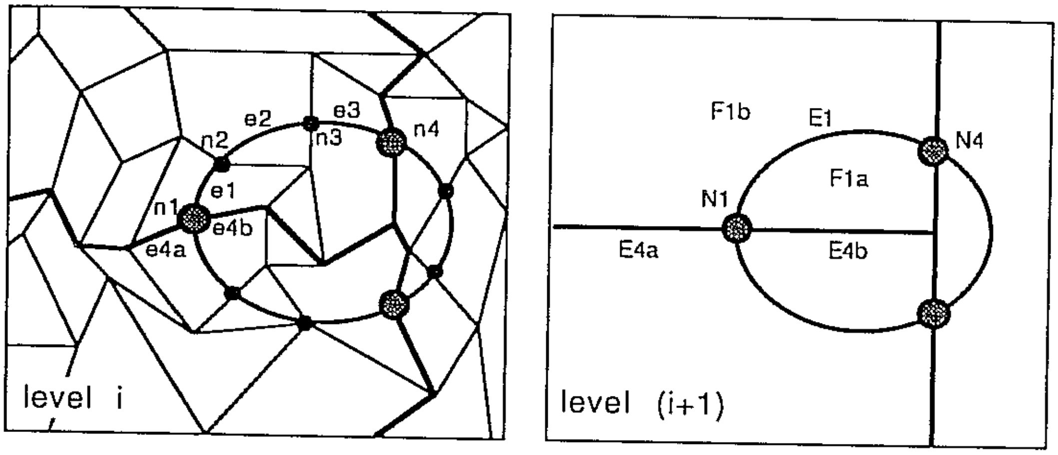
The process described here takes the definition of the object topology and propagates it up one level. If more than a one-level step is required, this procedure can be repeated accordingly.

Propagation is equivalent to an insertion on the higher level. The topological relations on the lower level provide enough information for the insertion process; no metric data has to be inspected. The insertion operations for these cells are provided by the single topological representations that build the MTR.

The higher level representations have to know where to insert a new cell. This location is expressed in terms of incidences with cells already existing on the higher level. The incidence cells can easily be received on the lower level using the boundary and coboundary operation. By using the *contained_in* operation they can be mapped up to the higher level.

Of course, not all cells of the new object are inserted on the higher level directly. Certain cells are aggregated to higher level cells before the mapping. Aggregations take place where cells describe topological relations with objects of an importance level less than the higher level.

In the process of inserting cells, existing cells are split into several cells. Both the definitions of higher level cells and objects keep a reference to their component cells. These references have to be updated after splitting. Note that the concept of hierarchically related cells considerably limits the amount of updates compared to a single topological representation.



(a) (b)

Figure 3: Two cell structures of adjacent levels after insertion of a new object depicted as an ellipse. Lower case labels stand for cells at level i, uppercase labels for cells at level (i+1). Labels n, N, e, E, f, and F stand for nodes, edges, and faces, respectively.

The following example illustrates the propagation process: Figure 3 shows two cell structures of adjacent levels i and (i+1) after insertion of a new object depicted as an ellipse. Thick lines show cells of objects of an importance level of (i+1) or higher, thin lines show cells of a lower importance level.

Node n1 is necessary to express the topological relation between the objects shown in thick lines; therefore, it is to be inserted in level (i+1). To determine the location of the insertion, first, a boundary operation on n1 is performed which yields the edges e4a and e4b. Then the *contained_in* operation maps both edges up to the edge E4 that will later be split into E4a and E4b. Knowing that N1 lies on E4 determines the location.

After mapping up nodes, higher level edges can be inserted. An analysis of the importance level of the remaining nodes and edges indicates that the set of cells {e1, n2, e2, n3, e3} map onto the higher level edge E1 connecting N1 and N4. Its position in the higher level can be determined by finding its incidence cells: A boundary operation applied to the set of cells returns n1 and n4. These nodes can be mapped onto N1 and N4. A coboundary operation applied to the set returns the faces on both sides of the aggregated edge. Mapping them up using *contained_in* yields the face F1. Knowing the incidence cells N1, N4, and F1, E1 can be inserted.

This last insertion split F1 into F1a and F1b. The references from these faces to faces on level i have to be updated. Finding the boundary of both faces in level (i+1), mapping them to level i using *consists_of* and applying the *inside* operation yields the updated set of component cells.

4.2 Retrieval

MTR offer a set of operations dealing with objects. They either return objects or a Boolean value. We will look only at the former type of operation, i.e., operations used to retrieve objects. They are directly inherited from the single topological representations which make up the MTR. The retrieval of an object often consists of a series of operations rather than just one. For example, to retrieve all parcels lying at the state boundary, a boundary and a coboundary operation have to be performed.

Some spatial data base management systems (DBMS) such as PANDA [Egenhofer 1989c] support spatial access methods based on metric criteria (e.g. a window). Both single and multiple topological representations can extend a DBMS, adding a spatial access methods based on topological criteria. The objects to be retrieved are specified in terms of the operations of the topological representation. The following examples illustrate possible retrieval criteria: "all neighboring objects of"; "all objects that are contained in"; "all objects that are connected by objects of a given set". A combination of a topological access criteria with others (e.g. an object class based method) would yield a very powerful spatial DBMS.

Retrieval of objects is different in single topological representations as compared to MTR, because in the latter case objects are represented on several levels. Retrieval operations can be performed on different levels. In addition to the operations inherited from the single representation, an MTR provides operations to map the intermediate results of the retrieval to other levels.

The base level contains all objects and is equivalent to a single topological representation. Operations performed on this level are equivalent to those of a single topological representation. Higher levels organize only subsets of all objects. Therefore, operations performed on these levels have restricted range as only objects of a certain importance level can be returned. The operations can be performed much faster, as the amount of detail (i.e. cells) is less on higher levels.

Users do not normally know that the system keeps multiple representations of objects. Therefore, the system has to decide by itself on which level operations shall be performed. This decision is based on the object classes involved in the query as well as the context of the query (e.g. objects currently displayed, display window and scale).

The strategy of selecting a level for an operation maximizes performance. The operation is always performed on the highest level possible. In the case of only a single object class involved, the operation is processed on the level of importance. If several objects are involved, the minimum of the levels of importance is chosen. The mapping of intermediate results onto higher levels is done as early as possible; the mapping down onto lower levels as late as possible.

The following example query demonstrates the selection of a level for its operations: "Retrieve all parcels in the state of Maine that lie at the border with Canada." Assume that parcels are of importance level 0, states of level 2 and nations of level 3.
Intersect (Maine, Canada) returns the common boundary. It is performed on level 2--the highest level where both states and nations exist. A coboundary operation on this part of the boundary yields the parcels of interest (together with other objects that have to be eliminated). As parcels exist only on level 0 the common boundary has to be mapped onto level 0 prior to execution of the coboundary operation.

Actually, there is an even better way to process the example query. Instead of intersecting Maine and Canada on level 2, the boundary of Maine and Canada can be derived first on level 2 and 3, respectively. The boundary of Canada is then mapped onto level 2 where the intersection operation is performed. The cost for this way of processing is considerably less because more of the processing can be done on higher levels where a smaller number of cells has to be inspected. This example demonstrates that a query optimizer is necessary in order to take full advantage of the reasoning capabilities of an MTR.

5. CONCLUSIONS

Multiple Representations are a necessary framework for reasoning about geographic space in GIS. Multiple Topological Representations (MTR) are a special case of Multiple Representations.

The paper described how spatial objects are organized in a MTR and how insert and retrieve operations work. Object retrieval based on topological criteria is a special case of spatial reasoning. It can be processed much more efficiently in MTR as compared to single topological structures. Retrievals that lead to unacceptable response times in single representations can be processed in MTR. This extends the application domain of topological representations. For example, multipurpose systems with a very large number of object classes and the coexistence of very small and very large objects (e.g. parcels and nations) can be implemented with MTR.

The description of object retrieval demonstrated that query optimization is necessary in MTR to exploit the whole potential of efficient reasoning.

ACKNOWLEDGEMENTS

Max Egenhofer's valuable comments on this paper are highly appreciated. Thanks to Werner Kuhn for his stimulating discussions on the topic.

REFERENCES

[Abler 1987] Ronald Abler. The National Science Foundation National Center for Geographic Information and Analysis. International Journal of Geographical Information Systems, Vol. 1, No. 4, 1987.

[Boudriault 1988] G. Boudriault. Topology in the TIGER File. In: N. R. Chrisman (ed.), Auto Carto 8 Proceedings, Baltimore, MD, 1987.

[Brodie 1984] Michael L. Brodie and Dzenan Ridjanovic. On the Design and Specification of Database Transactions. In: Michael L. Brodie, John Mylopoulos and Joachim W. Schmidt (eds.), On Conceptual Modelling. Springer-Verlag, New York, NY, 1984.

[Bruegger 1989a] Bud P. Bruegger et al. 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, 1989.

[Bruegger 1989b] Bud P. Bruegger and Andrew U. Frank. Hierarchies over Topological Data Structures. In: ACSM-ASPRS Annual Convention, Baltimore, MD, March 1989.

[Buttenfield 1989] 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, 1989.

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

[Egenhofer 1989a] Max J. Egenhofer. Spatial Query Languages, Ph.D. thesis. University of Maine, Orono, ME, 1989.

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

[Egenhofer 1989c] Max J. Egenhofer and Andrew U. Frank. PANDA: An Extensible DBMS Supporing Object Oriented Software Techniques. In: T. Hauder (ed.), Informatik-Fachberichte, Vol. 204, 1989.

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

[Herring 1988] J. R. Herring. TIGRIS: Topologically Integrated Geographic Information System. In: N. R. Chrisman (ed.), Auto Carto 8 Proceedings, Baltimore, MD, 1987.

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

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

[Kinnear 1988] C. Kinnear. The TIGER Structure. In: N. R. Chrisman, (ed.), Auto Carto 8 Proceedings, Baltimore (MD), 1987.

[Mark 1989] David M. Mark. 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, 1989.

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

[Peckham 1988] Joan Peckham and Fred Maryanski. Semantic Data Models. ACM Computing Surveys, Vol. 20, No. 3, September 1988.

[White 1985] M. White. A Survey of the Mathematics of Maps. Auto Carto 7 Proceedings, Washington (D.C.), 1985.

301.2/15

Authors' address:

Bud P. Bruegger and Andrew U. Frank
National Center for Geographic Information and Analysis
University of Maine
107 Boardman Hall
Orono, ME 04469
U.S.A.

e-mail:
BUD@MECAN1.bitnet
FRANK@MECAN1.bitnet