Properties of Geographic Data: Requirements for Spatial Access Methods[1]

Dr. Andrew U. Frank
Associate Professor, Department of Surveying Engineering
Associate Director
National Center for Geographic Information and Analysis
University of Maine
Orono ME 04469
(207) 581-2147 (Fax: (207) 581-2206)
FRANK@MECAN1.bitnet

Abstract

Spatial access methods and the corresponding data structures are necessary to acheive the expected overall performance of Geographic Information System (GIS) and are currently a prominent topic for research in spatial databases. Their performance is influenced by the properties of spatial data and they are optimized for different tasks. This paper details the specific properties of spatial data in a Geographic Information system. It concentrates on Geographic Information Systems that store data describing objects with a distinct identity and does exclude image and remote sensing databases from consideration - given that the characteristics of the data are very different.

We estimate values for the parameters of some of the properties described and point to characteristics where GIS data differ from VLSI or CAD data. Estimates for the size and structure of GIS data are also given. The results can be used for selection of a spatial access method for a GIS. They are also useful for the optimization of spatial access methods to respond to the specific requirements of a GIS. Last, the parameters given will be useful for the construction of benchmarks to test realistically performance of implemented systems.

1. Introduction

Spatial access methods and the corresponding data structures are necessary to acheive the interactive performance of Geographic Information System (GIS) and are currently a prominent topic for research in spatial databases [Buchmann, et al. 1990, Guenther and Buchman 1990, Samet 1989]. Several data structures have been proposed. They are based on the observation that geometric data contains some structure which must be exploited to yield better performance [Nievergelt 1990].

Comparing spatial access methods and selecting the one most appropriate for a GIS is hindered by the absence of a description of properties of spatial data and of requirements of applications. Kriegel used artificially constructed data sets and his results show that the data structures compared were sensitive to differences in the data [Kriegel, et al. 1990]. For the practical purposes of designing GIS software, one might conclude that the small differences observed between the spatial access methods do not matter [Smith and Frank 1990]. This may be wrong for a specific application area, where characteristics of the data or of the dominant spatial access favors one over the other of the proposed methods. GIS, VLSI and CAD/CAM, not to speak of other usages of spatial access methods [Wang and Shasha 1990], have apparently different requirements.

A group of experts recently agreed that the description of realistic requirements for specific applications of spatial databases should be a high priority research issue [Smith and Frank 1989].

[1] Funding from NSF for the NCGIA under grant SES 88-10917, from Intergraph Corp. and Digital Equipment Corp. is gratefully acknowledged.

Frank, Andrew U. "Properties of Geographic Data: Requirements for Spatial Access Methods." In *Advances in Spatial Databases*, edited by Oliver Günter and Hans-Jörg Schek, 225-34. Heidelberg-Berlin: Springer Verlag, 1991.

Systematic descriptions of requirements of applications are seldom done in terms relevant for the computer science researchers (an exception is [Oosterom 1990]). I previously discussed the database requirements for a GIS, from the perspective of an application designer [Frank 1981, Frank 1988]. Here we will use the point of view of the designer of a data structure for spatial access or of a spatial database and list the properties of geographic data and the requirements of a GIS as they are relevant for the design of spatial access methods.

The properties of spatial data in a GIS we discuss are influenced by our perception of how these systems operate and the problems encountered in designing data structures for spatial access. However, we will not discuss spatial access methods and only make cursory references to specific influences on the performance of spatial access methods to justify the choice of parameters where this seems necessary.

Meaningful parameters describing the data and applications will help to construct standard benchmarks, which are inexpensive to implement, and allow one to predict performance for specific uses [Cattell 1988, DeWitt, et al. 1990, Kriegel, et al. 1990]. This is not only useful for the design and selection of data structures but will contribute to the estimation of execution time and thus be useful for query optimization [Guenther and Buchman 1990, Ooi 1991].

The investigation considers only geographic information systems which are 'object' centered, i.e. have a vector or topological data structure and deal with spatial objects that have a distinct identity and are searched for in 2 dimensional space. It is not concerned with 'image databases' as their requirements are very different [Guenther and Buchman 1990].

The paper starts with a short description of Geographic Information Systems and a specific example of one application of such a system. In the following sections, we describe the data in a GIS, starting with requirements from geometric modelling, then detailing the functionality needed from the data model and last describe the spatial distribution of data. Section 8 gives estimates of the typical data volumes. We conclude with indications how to proceed to measure the listed properties and mention the need for a similar 'requirements description' for the dominant uses of GIS data.

2. Definition of GIS for This Study

For the purpose of building spatial access methods, and without attempting a general definition, Geographic Information Systems store large amounts of data that describes spatial (geographic) objects, like roads, political boundaries and rivers. The data contains both spatial (geometric) description of the objects and non-spatial attribute values that characterize the objects. This is essentially the group 'Geographic/Cartographic applications' in [Guenther and Buchman 1990]. We follow Guenther and Buchman's differentiation between spatial databases and image databases, produced mostly by remote sensing. GISs which store data values for a regular spatial subdivision (grid) and do not have a concept of a spatial object with identity, have very different characteristics, due to very different implementation strategies, and due to a very different underlying spatial concept [Frank 1990]. These so called 'raster' systems are not discussed here.

We are primarily considering those GISs that are used interactively by end users and do not focus on those used for the maintenance of maps, which are then printed and distributed as paper copies to the users [Frank 1991] (thus our object of study is a subset of the 'Geographic/Cartographic applications of [Guenther and Buchman 1990]).

A typical example of such a GIS is a municipal information system as used by a town or county for the management of all the spatial data necessary for the town's administration. Such systems are widely used, with slightly different contents and missions, for example for

- tax rolls to maintain lists of all taxable properties in a town,
- zoning data used for approving building permits,
- Automated Mapping and Facilities Management (AM/FM) systems as used by nearly all public utilities,

- systems for the management of spatial resources, e.g., forest or agricultural land, infrastructure for recreation.

Typical data types in a municipal GIS are:

- land ownership (parcels), with boundary, name of owner etc.,
- current and permitted land use (zoning),
- road network, with details about traffic density, road class, maintenance etc.
- location of public buildings (school, administration, police stations, etc.),
- rivers, sewer lines.

For each data element, we have geometric and descriptive attributes. Geometry data is often based on a precise survey of the area.

3. Characteristics of Geographic Data

In the following sections we describe the data as it is typically found in a GIS. It is important to separate parameters that describe the data independent of a specific spatial access method and others, which are only applicable and meaningful for a specific solution (e.g., 'overlap' in R and R+ trees). Clearly the parameters selected depend on a specific spatial concept [Frank 1990], which determines, how we model reality in a spatial database. We use - as indicated above - a concept of 'identified object in space', closely related to a 'topological' or 'vector' data model. A different spatial concept (e.g., the 'raster' data model) would lead to a different spatial data model and then to different properties and different parameter values. From this point of view, this section describes the dominant, most adequate methods to model geographic data for the kind of applications listed in the previous section.

The list is structured in four parts, namely parameters related to

- modelling geometry
- requirements for the data model
- spatial distribution of geometric object, and
- changes during use and maintenance of the data set.

It is followed by an estimate of data volume in a GIS. It is clear, that this list is not complete, but a first attempt to be expanded in the future. We make an effort to give a quantitative measure for most of the parameters, based on values found in the litterature and our own work. They are first estimates, indicating ranges or typical values.

4. Geometric Modelling

A set of base geometric data types must include point, lines and areas; this set must be extensible to adapt to application needs.

The base geometric types can be points, lines and areas; some systems allow arbitrarily formed lines, modelled as sequences of points, while others use splines, arcs of circles, etc. Early systems - primarily for map maintenace - dealt only with points and lines (so called 'spaghetti' data structures). The requirements of applications within a GIS differ and now known set of geometric primitives fulfills all requirements. A general purpose system must therefore allow user to extend it.

The higher dimensional spatial objects must be built from the lower dimensional ones.

Lines are constructed from points, areas from lines. Geographers speak of a toplogical or vector data structure, comparable to the boundary representation in CAD [Requicha 1980]. In database terms, treatment of complex (non atomic) objects is necessary [Wolf 1990, Zdonik and Maier 1990]. This is a natural break down of objects into smaller ones, so no spatial object is very large, but includes references to other spatial objects that it shares with others to form the geometry and is different from VLSI, where spatial objects seem not to share boundaries, and also different from some spatial models used in CAD.

We have advocated the construction of all geometry from simplicial complexes, using combinatorial topology as a base [Egenhofer, et al. 1990, Frank and Kuhn 1986]. This approach is somewhat similar to 'constructive solid modelling' in CAD. The geometry of real objects is formed as an aggregate of simplices to a simplicial complex. The more general, cell based methods are widely used [Corbett 1979, Herring 1989].

The topological relations do not form a hierarchical structure.

Smaller spatial objects are part of more than one complex object: the boundary line is the border of two areas, a point can be a corner of several areas. These relations represent topological relations: a boundary line limits a parcel, a county is adjacent to another[Alexandroff 1961]. It is important for many applications that one can follow such relations quickly, for example to find a connected path along a street network. It has not been possible to implement these access methods using relational operators efficiently [Haerder, et al. 1987].

The spatial access methods must be based on the most general abstraction of spatial objects.

The variation between the different object types is large and extensibility required. Thus the spatial access method must rely on only the most general solutions such as minimal bounding rectangle for each object [Wolf 1990] (the cell tree [Günther 1988] uses another general method, see also [Hinrichs 1985]). This is necessary to separate the complexity of the search operation from the complexity of the object geometry [Nievergelt 1990].

5. Object Modelling

In order to adequately model the data and their relations, a powerful data model is necessary. The next paragraphs describe some of the specific requirements for the data model. Attempts to structure data using the relational or variants of the entity-relationship model clearly show the limitations [Frank and Tamminen 1982, Meyer 1989] and forces the designer into artificial constructions to make the data structure as perceived by the user fit the data model. We expect that the data models based on object-oriented concepts provide most of the tools necessary [Kim 1990, Zdonik and Maier 1990].

Objects in this type of GIS have an identity and cannot arbitrarily be divided for the purpose of the spatial access method.

Objects have an identity and geometric and attibute data which described their properties. It is not possible to subdivide objects along arbitrary boundaries imposed by the spatial access methods (field boundaries or similar [Chrisman 1990]) where each part has its own set of attributes. If a spatial access method requires spatial subdivision (clipping [Wolf 1990]), then it must be possible to find an object based on spatial location and any condition on the attribute values of the whole object. We say that the objects are non-atomic (i.e., complex in the object-oriented sense) but not spatially decomposable (in spatial pieces that still have the same meaning).

The spatial access method must be able to accomodate many different spatial data types, not just the base geometric types.

The number of different data types in a GIS is in the order of 100 .. 1000 and most of them are spatial (we estimate that only about 1/4 of the data types in a GIS are not spatial). It is necessary to be able to search spatialy for objects based on their type. For example queries like 'find all school buildings within <<area>>' are customary, and finding first all areal featuers and then select from them the school buildings seems wasteful.

A complex object may consist of sub-object (components) which should be stored with the object and not accessed using a reference.

A complex spatial object may need to have subobjects, which are directly stored and are not just references to objects stored elsewhere. For example an irregular line can be modelled as a sequence of points (objects), but these are not shared with other lines and it is desirable to directly include them with the line object and not just include references which then lead to additional database accesses. Some of the object-oriented data model include such a feature, for example making a difference betweeen 'objects' and 'values' (e.g. O2 [Deux 1989]).

Spatial objects are often formed as aggregates of other spatial objects and do not have their own spatial description.

A census block is an aggregate of parcels, a town an aggregation of school districts. The data model used must adequately model these structures (similar to "bill of material"). A spatial object can participate in more than one aggregation; a parcel is part of a school district and a voter's precinct. There exist multiple hierarchies (3 .. 10) each several levels deep (3 .. 5).

Objects in a GIS are often classified in groups and subgroups, and an object can be seen, depending on the point of view, as being in more than one classification hierarchy.

Different application specialist see different kinds of generalizations and a single object class can be seen as a special case of multiple general ones. For example natural streams are a generalization of brooks and rivers; waterways are a generalization of navigable rivers, channels, etc. Clearly navigable river inherits properties both from waterways and streams [Egenhofer and Frank 1987]. We find inheritance chains that are 2 .. 10 classes deep.

The same geometric form can define more than one object.

A single geometry can serve to describe multiple spatial objects depending on the point of view. A line may be a boundary for a parcel, a county and a river. A piece of land may be a parcel and a unit of land use. Geometry of an object is then expressed by reference to the geometric object and multiple spatial objects share the same geometry. Nevertheless, it must be possible to search for the spatial objects based on their spatial location even if they do not contain a detailed geometric description.

The same object can be represented by more than one geometry (multiple representation).

Spatial objects may be modelled with a few, quite different geometries, depending on the level of detail considered: the same road may be a (center) line or an area, depending on the scale and the geometric model used. This may be used to support cartographic generalization for improved output [Oosterom 1990], but has much wider applicability [Buttenfield and DeLotto 1989]. Databases may contain multiple geometries, which are substantially different for a single object [Frank 1991]. The data model must be able to deal with such associations [Oosterom 1990]. Together with the 'multiple meanings per geometric object', we find n:m relations between geometry and spatial objects.

The data types shown on a map can be ordered by 'importance', indicating the scales of maps they may appear on.

Map output can only be understood if sufficient spatial context is provided; isolated spatial objects do not 'speak for themselves' [Frank 1982]. Certain types of objects are more likely to be displayed on a map for context, because they are landmarks and stand out in the user's perception of space. Landmark size is somewhat linked to importance, thus large spatial objects tend to be important [Frank 1981]. These objects are retrieved for small scale maps of large areas, whereas objects of lesser importance are retrieved primarily for large scale maps of smaller areas.

6. Distribution of Objects in Space

Some objects may be completely inside another one.

A spatial object may completely include another spatial object, often if the smaller one is a constituent part of the larger one (say counties forming a state), but the same condition is encoutered even if no other relation exists between the two (an area of uniform soil type within a town). As Freestone [1990] observed, this can cause problems for some spatial access methods . Generally, areas are only partially ordered by inclusion [Kainz 1988].

Generally, objects overlap each other spatially.

For most points in a GIS, several geometric objects overlap: for example, at any location we have:
- an ownership parcel,
- a land use zone,
- a census block,
- a town.

The overlap of objects vary, but in general 5 to 10 objects will be found at any given location. This is the overlap of spatial objects, not the overlap of 'fields' in an R tree, as used in Faloutsos performance analysis - which is an internal parameter dependent on the construction of the data structure [1987].

Spatial data can be organized in thematic layers.

Some systems organize the objects in thematic layers, where all objects of one type with their geometric descriptions form a layer (and are often stored in a separate file). A very important class of themes form 'categorical coverage' [Beard 1988], i.e., a partition, which lend themselves to this organization.

It is then necessary to 'overlay' one such thematic layer with another one and determine their intersection, i.e., areas with uniform properties in more than one theme [Tomlin 1989]. During the overlay operation, difficult problems of the finite precision in computational geometry have to be dealt with. We have therefore advocated that the geometry of all objects is integrated when data are collected, essentially precomputing and storing all geometric intersections [Frank 1987]. Saalfeld has given formulae for the number of regions in the result of an overlay [Saalfeld 1991] and a commercial system has shown that acceptable performance can be achieved [Herring 1990].

The size (spatial extent) of the objects in a GIS vary enormously.

The largest objects are usually as large as the total area covered and the smallest ones are much smaller, spanning relations of $1:10^6$ or more (in diameter, not in area). For example a GIS for a county will deal with the county as the largest object, of say a diameter of 100 km; the smallest objects in the same system will perhaps measure 1 m (and resolution of the location of the object may again be smaller, up to .001 m).

The variation in size of objects within a type are large, but limited.

The size of the object and the kind of object are related: for example, the size of a county is in a certain range, and this range does not overlap with the range of the size of buildings (i.e. the largest building is much smaller than the smallest county) - somewhat similar to the two populations in [Faloutsos, et al. 1987]. Such ranges may span 1:100. As an example consider buildings: the smallest have dimensions of 3m, the largest are in the several 100 m. This seems to be a point where GIS data is different from VLDB data, as Faloutsos reports [1987].

The number of spatial objects per unit of space varies greatly.

The distribution of objects in space is irregular and objects are clustered (one of the parameters in Kriegel's data sets [1990]). Ooi calls this parameter 'locational skew' and reports that it helps differentiate the performance of several spatial access methods[Ooi 1991]. The differences in density may be very substantial and surpass 1:1000 [Frank 1981]. The area of a GIS usually covers an irregularly bounded area and no data is available for the space outside (this may be half the area, depending on the exact form of the area of interest).

7. Change of Data Over Time

Both, object geometry and the spatial distribution of objects change through updates.

Objects change in time: objects are created and deleted, change their geometry, etc. Even the distribution of objects in space may change - areas may change from agricultural use to suburban, and the number of objects to be recorded for a new subdivision is much higher than before. A data structure must be able to adapt to such changes dynamically, either immediately balancing or tolerate unbalanced storage first and include a balancing operation which is carried out later, when resources are available.

Many GIS have a need to preserve previous values (historic data).

If a GIS is used for data that have legal significance, e.g., property maps, not only the current state is needed but it is necessary to be able to reconstruct a previous situation (the situation 'as of <<data>>').

Versions are necessary to support design work in a GIS.

If a GIS is used to do design work, designers must be able to work relatively independently and draft possible solutions in parallel, working in small groups and storing their changed data separately from the currently valid data, usually called versions, etc. [Katz 1990].

8. Data Volume in a GIS

GISs are very large collections of data. The data volume depends on the special purpose a system is built for and the area covered. Nevertheless, the quite general rule of 'constance of data volume' for an organization applies. If it either collects much detail for a small area or it collects generalized data for a large area, the product of data density times area is relatively constant.

If the data in a GIS is structured with minimal redundancy, the records will be quite small and contain a large amount of references to other records. In a specific example, we found more than 12 references to other records stored in a single record. We estimate that the total storage space per record amounts to 100 .. 500 bytes. Not all implementation strategies for geometric objects need records of varying length, but storage size between different object types varies substantially.

In order to improve performance it is sometimes recommended to store the complete geometry of an object with the object attribute data. This can be in addition to a minimal redundancy storage, but introduces redundancy and requires attention during updates. It also more than doubles storage requirements. Records that contain a full description of object geometry can be substantially larger, with storage size varying from object to object and storage of records of varying length being necessary.

The number of records is dependent on many aspects and depends primarily on area, density of population and the comprehensiveness of the data collection, especially the number of topics included. A simple estimate can be made using the number of inhabitants in an area. Sizeable towns are most in need of a GIS, thus we may assume 100,000 to 10 million inhabitants as typical. On average, the number of parcels is about equal to the number of inhabitants. A parcel results in 20 .. 100 records. This yields a total of 2 .. 1,000 * 10^6 records. If we assume that 50% of the storage is used for overhead or is empty, the database uses 0.4 .. 1,000 Gbytes.

Existing cartographic databases fall in the same order of magnitude of size: Goodchild reports 25 Mbytes per map sheet and the coterminous U.S. is covered by 100,000 sheets, which will lead to a total of 2,500 Gbytes. For a specific large town of the U.S. (Los Angeles county) the data volume is estimated at 300 Gbytes [Goodchild 1990].

9. Conclusions

The paramters listed here show some evident differences between data in a GIS and data in other spatial information systems, e.g. VLSI or CAD, but the parameters listed here are certainly not complete and describe some but not all, of the factors influencing performance of spatial access methods. There is considerable work left to complete the list of relevant parameters. As this list depends on spatial concept and data model, efforts to assess data conceptualized or structured using a different model, are warranted. Of primary interest are raster data as used in remote sensing systems.

In order to determine accurate values for the parameters, operational definitions and methods to measure them in a given data set are necessary. This is perhaps only worthwile for parameter for which we can optimize data structures specifically and not for others which help only to decide between two different spatial access methods. It would be worthwile, to test the known data structures for their reaction to the above parameters. Many of the existing tests do report results either only globally for a number of not well characterized data sets or the parameters varied and reported for, are a combination of the ones listed here. For example, variations in distribution and overlap are not separated in [Ooi 1991].

A special problem with the developed data structures for spatial access is their construction in isolation. A GIS, or any other similar information system that deals with spatial data, e.g., a CAD system, is a complex collection of programs, of which the spatial access methods are only a part. It becomes, therefore, important to consider the interaction between spatial access and other required components of a GIS, specifically the database management system. For this reason we have not only discussed the purely spatial aspects, e.g. the spatial distribution, but also other characteristics of spatial data that affect their organization in a database management system.

In order to properly select and optimize spatial access methods, we need not only know the properties of the data but also understand the loads. What are the dominant access methods that users of a GIS ask for? Which ones are crucial to optimize to reach acceptable performance? We assume, that access to data to produce a small map on the screen is crucial, because it requires retrieval of a large number of objects and the user is waiting for the answer. A detailed description of this and other typical accesses to spatial data in a GIS is necessary.

REFERENCES

P. Alexandroff. 1961. *Elementary Concepts of Topology*. New York, NY: Dover Publications, Inc.

K. Beard. 1988. Multiple representations from a detailed database: a scheme for automated generalization. Ph.D. Thesis thesis, University of Wisconsin - Madison.

A. Buchmann, O. Günther, T. Smith and Y.-F. Wang ed. 1990. *Design and Implementation of Large Spatial Databases - Proceedings of the First Symposium SSD '89, Santa Barbara CA, July 1989,*. Lecture Notes in Computer Science. New York: Springer Verlag.

B. P. Buttenfield and J. S. DeLotto. 1989. *Multiple Representations: Initiative 3 Specialist Meeting Report.* Report 89-3. : NCGIA.

R. G. G. Cattell. 1988. "Object-Oriented DBMS Performance Measurement". In *Advances in Object-Oriented Database Systems.* Edited by K. R. Dittrich. 364 - 367. Bad Muenster: Springer Verlag.

N. R. Chrisman. 1990. Deficiencies of sheets and tiles: building sheetless databases. *International Journal of Geographical Information Systems* 4 (2) : 157-167.

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

O. Deux. 1989. "The Story of O2". In *Fifth Conference on Data and Knowledge Engineering.* :

D. J. DeWitt, P. Futtersack, D. Maier and F. Velez. 1990. "A Study of Three Alternative Workstation Server Architecture for Object-Oriented Database Systems". In *Very Large Databases.* Edited by D. McLeod, R. Sacks-Davis and H. Schek. 107 - 121. Brisbane: Morgan Kaufman.

M. J. Egenhofer and A. U. Frank. 1987. "Object-Oriented Databases: Database Requirements for GIS". In *International Geographic Information Systems (IGIS) Symposium (IGIS'87).* Edited by R. T. Aangeenbrug and Y. M. Schiffman. 189-212. Arlington, Virginia:

M. J. Egenhofer, A. U. Frank and J. P. Jackson. 1990. "A topological data model for spatial databases". In *Symposium on the Design and Implementation of Large Spatial Databases.* Edited by A. Buchmann, O. Günther, T. R. Smith and Y.-F. Wang. 271-286. New York, NY: Springer-Verlag.

C. Faloutsos, T. Sellis and N. Roussopoulos. 1987. "Analysis of Object-Oriented Spatial Access Methods". In *SIGMOD, Proceedings of ACM Special Interest Group on Management of*

Data, 1987 Annual Conference. Edited by U. Dayal and I. Traiger. 426 - 439. San Francisco, CA: ACM Press.

A. U. Frank. 1981. "Applications of DBMS to Land Information Systems". In *Seventh International Conference on Very Large Data Bases VLDB*. Edited by C. Zaniolo and C. Delobel. 448-453. Cannes, France:

A. U. Frank. 1982. MAPQUERY: Database Query Language for Retrieval of Geometric Data and its Graphical Representation. *ACM SIG Computer Graphics* 16 (3) : 199 - 207.

A. U. Frank. 1987. "Overlay Processing in Spatial Information Systems". In *AUTO-CARTO 8, Eighth International Symposium on Computer-Assisted Cartography*. Edited by N. R. Chrisman. 16-31. Baltimore, MD: ASPRS & ACSM.

A. U. Frank. 1988. Requirements for a Database Management System for a GIS. *Photogrammetric Engineering & Remote Sensing* 54 (11) : 1557-1564.

A. U. Frank. 1990. "Spatial Concepts, Geometric Data Models and Data Structures". In *GIS Design Models and Functionality*. Edited by D. Maguire. Leicester, UK: Midlands Regional Research Laboratory, University of Leicester.

A. U. Frank. 1991. "Design of Cartographic Databases". In *ICA Handbook Cartography*. Edited by J.-C. Muller. Forthcomming.

A. U. Frank and W. Kuhn. 1986. "Cell Graph: A Provable Correct Method for the Storage of Geometry". In *Second International Symposium on Spatial Data Handling*. Edited by D. Marble. 411-436. Seattle, WA:

A. U. Frank and M. Tamminen. 1982. "Management of spatially referenced data". In *Land Information at the Local Level*. Edited by A. Leick. 330-353. Orono, ME: Surveying Engineering, University of Maine at Orono.

M. W. Freeston. 1990. "A Well Behaved File Structure for the Storage of Spatial Objects". In *Symposium on the Design and Implementation of Large Spatial Databases*. Edited by A. Buchmann, O. Günther, T. R. Smith and Y.-F. Wang. 287 - 300. New York, NY: Springer-Verlag.

M. F. Goodchild. 1990. "A geographical perspective on spatial data models". In *GIS Design Models and Functionality*. Edited by D. Maguire. Leicester: Midlands Regional Research Laboratory.

O. Guenther and A. Buchman. 1990. Research Issues in Spatial Databases. *SIGMOD Record* 19 (4) : 64-68.

O. Günther. 1988. *Efficient Structures for Geometric Data Management*. Lecture Notes in Computer Science. New York: Springer Verlag.

T. Haerder, K. Meyer-Wegner, B. Mitschang and A. Sikeler. 1987. "PRIMA---A DBMS Prototype Supporting Engineering Applications". In *13th International Conference on Very Large Data Bases*. 433-442. Brighton, England:

J. R. Herring. 1989. "A fully integrated geographic information system". In *Auto-Carto 9*. Edited by E. Anderson. 828-837. Baltimore, MA: ASPRS & ACSM.

J. R. Herring. 1990. "TIGRIS: A Data Model for an Object Oriented Geographic Information System". In *GIS Design Models and Functionality*. Edited by D. Maguire. Leicester: Midlands Regional Research Laboratory.

K. H. Hinrichs. 1985. The Grid File System: Implementation and Case Study of Applications. Doctoral dissertation thesis, ETH.

W. Kainz. 1988. "Application of Lattice theory to geography". In *Third International Symposium on Spatial Data Handling*. Edited by D. F. Marble. 135-142. Sydney, Australia: International Geographical Union, Commission on Geographical Data Sensing and Processing.

R. H. Katz. 1990. Toward a Unified Framework for Version Modeling in Engineering Databases. *ACM Computing Surveys* 22 (4)

W. Kim. 1990. Object-oriented Databases: Definitions and Research Directions. *IEEE Transactions on Knowledge and Data Engineering* 2 (3) : 327-341.

H.-P. Kriegel, M. Schiwietz, R. Schneider and B. Seeger. 1990. "Performance Comparison of Point and Spatial Access Methods". In *Symposium on the Design and Implementation of*

Large Spatial Databases. Edited by A. Buchmann, O. Günther, T. R. Smith and Y.-F. Wang. 89-114. New York, NY: Springer-Verlag.

N. v. Meyer. 1989. A Conceptual Model for Spatial Cadastral Data in a Land Information System. Ph. D. thesis thesis, University of Wisconsin.

J. Nievergelt. 1990. "7 ± 2 Criteria for Assessing and Comparing Spatial Data Structures". In *Symposium on the Design and Implementation of Large Spatial Databases*. Edited by A. Buchmann, O. Günther, T. R. Smith and Y.-F. Wang. 3-28. New York, NY: Springer-Verlag.

B. C. Ooi. 1991. *Efficient Query Processing in Geographic Information Systems*. Lecture Notes in Computer Science. Berlin: Springer Verlag.

P. v. Oosterom. 1990. Reactive Data Structures for Geographic Information Systems. Ph.D. thesis, Rijksuniversiteit te Leiden.

A. Requicha. 1980. Representation for rigid solids: theory, methods and Systems. *ACM Computing Surveys* 12 (4) : 437-464.

A. Saalfeld. 1991. An Application of Algebraic Topology to an Overlay Problem of Analytical Cartography. *Cartography and Geographic Information Systems* 18 (1) : 23 - 36.

H. Samet. 1989. *Applications of Spatial Data Structures. Computer Graphics, Image Processing and GIS*. Reading, MA: Addison-Wesley Publishing Co.

T. R. Smith and A. U. Frank. 1989. *Very Large Spatial Databases - Report from the Specialist Meeting*. Technical Repot 89-13. Santa Barbara, CA: NCGIA.

T. R. Smith and A. U. Frank. 1990. Very Large Spatial Databases - Report from the Specialist Meeting. *Journal of Visual Languages and Computing* 1 (3) : 291 - 309.

C. D. Tomlin. 1989. *Geographic Information System and Cartographic Modeling*. New York: Prentice Hall.

T.-L. Wang and D. Shasha. 1990. "Query Processing for Distance Metrics". In *Very Large Databases*. Edited by D. McLeod, R. Sacks-Davis and H. Schek. 602 - 613. Brisbane: Morgan Kaufman.

A. Wolf. 1990. "The DASDBS GEO-Kernel: Concepts, Experiences, and the Second Step". In *Symposium on the Design and Implementation of Large Spatial Databases*. Edited by A. Buchmann, O. Günther, T. R. Smith and Y.-F. Wang. 67 - 88. New York, NY: Springer-Verlag.

S. Zdonik and D. Maier. 1990. "Fundamentals of Object-Oriented Databases". In *Readings in Object-Oriented Database Systems*. Edited by S. Zdonik and D. Maier. 1 - 32. Palo Alto, CA: Morgan Kaufmann Publisher, Inc.