

What's Special about Spatial? Database Requirements for Vehicle Navigation in Geographic Space*

(Extended Abstract)

Max J. Egenhofer
Department of Surveying Engineering
Department of Computer Science
and
National Center for Geographic Information and Analysis
Boardman Hall
University of Maine
Orono, ME 04469-5711
max@mecan1.maine.edu

Abstract

Geographic databases are becoming a popular subject for research projects. It is acknowledged that database requirements for such applications as Geographic Information Systems (GISs), computer cartography, remote-sensing image databases, and emergency routing/dispatching are distinct from those for traditional database applications; however, the specific database needs due to the properties of geographic data are frequently overlooked. Here, we describe a specific problem domain, a *vehicle navigation system*, and analyze the properties of data to be modeled appropriately and managed efficiently by database management systems. Vehicle navigation systems need continuous, fast access to very large databases. Through the combination of *real-time access* and *geographic data*, we identify particularly challenging database requirements of this application with respect to data models, query languages, and query processing.

1. Introduction

Geographic databases are integral parts of geographic information systems (GISs), an annual multi-billion dollar market, or serve as inventories for such scientific investigations as monitoring global change. These applications display, manipulate, and analyze large sets of geographic data, to which they must have fast access. Current commercial GIS products use database management systems (DBMSs) only rarely, because their performance and modeling capabilities are insufficient. More typically, DBMSs are used for the storage of

alphanumeric information, called *attribute data*, and the geometric components are stored in dedicated files.

In the past, the database community considered the particular needs of geographic databases only occasionally, e.g., in query languages and query processing (Go *et al.* 1975; Frank 1981; Lohman *et al.* 1983; Roussopoulos and Leifker 1985; Egenhofer and Frank 1988; Güting 1988; 1989; Ooi and Sacks-Davis 1989; Ooi *et al.* 1989; Aref and Samet 1991; Becker and Güting 1992). An exception was the focus among database researchers on storage and access methods for spatial data (see Samet (1989) for a comprehensive review of what has been published up to the late 80's). Only recently has the design of database management systems for geographic applications received increasing attention by the database research community (French *et al.* 1990; Günther and Buchmann 1990; Smith and Frank 1990). Prominent database research programs for handling geographic data include the NCGIA Research Initiative on Large Spatial Databases (Smith and Frank 1990), the European activity Basic GOODS (Ausiello 1991), and Sequoia 2000 (Guptill and Stonebraker 1992). Additional exposure to the database needs of geographic applications has come through the bi-annual symposium series on the Design of Large Spatial Databases (Buchmann *et al.* 1989; Günther and Schek 1991).

Past discussions of database needs combined *geographic* properties with *spatial* properties (Neuhold and Stonebraker 1988; Günther and Buchmann 1990). The only distinction made was based on different characteristics in the underlying spatial concepts and implementation strategies: *Object-centered spatial databases* deal with spatial objects that have a distinct identity and represent their relationships in some embedding space. Such objects are bounded and spatially referenced, primarily through neighborhood relationships, but also with respect to one or more global reference (coordinate) systems. *Spatial image databases*, on the other hand, store data values for a regular spatial subdivision and do not use the concept of a spatial object with identity.

This paper takes a different approach as it specifies the properties of data with which a specific application of geographic databases—a vehicle navigation system—has to deal. Vehicle navigation systems assist drivers in planning trips, selecting routes, giving driving instructions, and guiding them through geographic space, i.e., space that is large-scale, beyond the human body, and that refers to human scale. Prominent commercial car navigation systems are

* This work was partially supported by grants from Intergraph Corporation and the University of Maine Summer Faculty Research Fund. Additional support from NSF for the NCGIA under grant No. SES 88-10917 is gratefully acknowledged.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

SIGMOD /5/93/Washington, DC, USA

© 1993 ACM 0-89791-592-5/93/0005/0398...\$1.50

Bosch's Travelpilot and Etak's Navigator (see White (1991) for a survey). Typically, cars are equipped with an on-board computer and a dash-mounted CRT. The CRT is used primarily to display simple maps with the location of the car, and to convey driving instructions. Street network data sets are distributed on CDs or cassette tape so that most data necessary to navigate the vehicle is locally available. Additional storage devices (hard disk, floppy) will be necessary to record incoming instructions about updates of the availability of certain roads. Links to customer address databases, stored on carry-on laptops, may be necessary. The car is also equipped with sensors that record constantly its position. The most sophisticated sensors use satellite measurements from the Global Positioning System (GPS). Other cars use sensors attached to the wheels to determine the speed and direction difference of the moving car. Particular requirements are placed on the ergonomics of the system—drivers must be able to operate them while driving. Information displayed on the screen must represent the essentials only so that the driver can grasp and understand them within the blink of an eye. For managing the storage of their data, these systems use dedicated file systems rather than DBMSs.

We challenge the database community to develop new database management systems that are tailored for handling the requirements of geographic data. We do not offer any particular solution as there are certainly numerous ways of meeting the challenge; however, we will provide a detailed description of some of the *properties* of the data to be managed and the kinds of *operations* (and queries) users will want to perform with a vehicle navigation system. The example selected is rich enough to cover many aspects of geographic databases, and some of their combinations are so unique that they warrant special attention from the database designers. This is an approach quite different from the usual trends in computer science where one tries to find domain-neutral solutions, though it follows the general ideas laid out recently about the involvement of domain specialists in database research for scientific applications (French *et al.* 1990). We claim that it is necessary to study such concrete and specific examples of application domains, beyond the generalized discussion of geographic or even spatial databases, in order to identify their specific database needs, because within an application domain such as geographic databases, the database requirements may be vastly different. For example, a remote sensing image database deals with the retrieval of very large sets of unstructured data, while a system maintaining records about ownership of land parcels contains highly-structured geometry and attribute data that changes frequently and underlie rules prescribed by legal authorities.

2. Modeling Issues in Vehicle Navigation Systems

Navigation is a real-world application, in which much of the complexity of dealing with geographic data can be observed (Kuipers 1978; Kuipers and Levitt 1988). The ontology of a vehicle navigation system includes such spatial objects as roads, their intersections, landmarks along roads, and cities to which users want to be guided or through which they are driving. These objects are all geographic in nature, that is, they have some properties with respect to location, extent, shape, and neighborhood. The spatial objects are also concurrently part of several reference systems through which users want to access them—geodetic coordinates such as

lat/long or UTM, postal addresses, city blocks, or mileage along the road.

At a first glance, one might be tempted to consider the geometry of these objects as they are represented on maps—roads are lines, their intersections are points, cities are points that will be displayed as a particular cartographic symbol. This view may correspond to what one thinks of when planning a trip (Timpf *et al.* 1992). It becomes immediately clear that this geometric model would have only limited use in navigation tasks in which more detailed geometric knowledge is necessary: some roads may be one-way only, therefore, the “lines” must have an orientation; other roads may have access restrictions due to their width—most often at bridges—or their height, for instance in tunnels. At a different level of resolution, necessary for the task of driving and giving driving instructions, a completely different “view” may occur as roads have an extent with separate lanes going into the same, or opposite, directions; intersections may be areas as in the case of two country roads crossing, or complex “clover leaves” if several Interstate highways “cross;” and cities that were thought of as simple points, are areas separated by roads subdividing them into city blocks. Sometimes, it may be appropriate to consider geographic space at a planning level as a flat, 2-dimensional surface, however, additional spatial dimensional information is necessary in order to model correctly overpasses vs. intersections.

3. Typical Queries

In order to understand the database requirements it is necessary to know the kinds of operations performed in a vehicle navigation system. While the operations may appear to be mainly geometric manipulations, there is an equally strong need for operations involving non-geometric manipulations. These operations use non-geometric properties of geographic data, e.g., roads have names (I95, Route 1a), they are classified (Interstate, secondary road, dirt road), and their segments may have a toll. Users ask queries with spatial as well as non-spatial constraints (“Which route to take around New York City?”), and they request the query results in both graphical and alphanumeric form (e.g., the instruction “turn left at the next intersection” vs. a sketch showing the location of the car and the route it is taking). The following queries are typical requests users ask:

- Display a map with the vehicle's current location and in the context of an appropriate geographic environment. The map—showing the position and displaying some crucial travel information such as landmarks, nearby cities, and roads—must be constantly updated as the vehicle moves along the road. These updates include (1) panning, i.e., moving ahead as the car moves along the road, (2) changing the orientation of the map when the vehicle follows the curved road in order to match the driver's view with the map (“ahead is up”), and (3) zooming in or out, i.e., changing the scale and the level of detail. Such a map may also serve as the background and display medium for further types of queries such as
 - Display the vehicle's route while driving along a road.
 - Find a particular location based on a postal address.
 - Find the shortest path among two or more locations (known as the traveling salesperson problem). The operation of routing cars may require elaborate spatial analyses such as finding the shortest path where

“shortest” may have multiple interpretations, e.g., in terms of distance, time, gas consumption, or toll.

- Generate driving instructions (either textual, verbal, or graphical) to guide the driver along a certain route. It is crucial that these instructions come at the right time—too late is obviously unacceptable, and may even be dangerous, while too early might be confusing for the driver.
- Get information with respect to the current location such as next exit, closest gas station, expected travel time to an Italian restaurant, or the distance to some city. This may require the link with a “Yellow Page” or customer database.

It is important to recognize that the car’s location, the reference for most queries, changes constantly. Its current position may be recorded in coordinates observed with GPS or—if satellites are temporarily unavailable—derived from a combination of the last known position and data collected by sensors attached to the car’s wheels, which is then matched with the information about the road network in the database.

4. Properties of Geographic Data

Although geographic data are frequently associated with maps, geographic databases are more than just repositories of maps or map maintenance systems. Geographic databases manage geographic data, while maps are one of many possible graphical forms in which geographic data may be presented. These issues are relevant for the storage and retrieval of geographic data, as well as for the interaction with geographic databases (i.e., query languages) and for reasoning processes to derive information (e.g., query processors).

4.1 Geometry

Geometry is a major components of any kind of geographic data. Usually, the geometric models for geographic objects are built from primitive geometric objects—points, lines, and areas. These primitives are interrelated, e.g., a simple line is bounded by two points, and an area has several lines as its boundary. Complex geometric objects are then aggregates of the primitives. It is important to recognize that the composition of primitive objects is most often not strictly hierarchical, but partially ordered, i.e., a primitive may belong to more than one complex geometric object such as the intersection between two highways (at the planning level) being part of four lines.

Among the spatial objects, there are a number of complex geometric relationships that must be preserved and are used extensively when querying the system:

- Connectivity, e.g., from one intersection to another.
- Adjacency, e.g., counties to the left and to the right of the roads.
- Order, e.g., the linear sequence of landmarks along the road, and also cyclic orderings such as the sequence of incoming and outgoing roads in a roundabout.
- Metric relations and evaluates such as the lengths of segments, cardinal directions such as north and east, or approximate distances as such near and far.

It is important to distinguish here between *geometry* and *graphics*. Graphics is concerned with the display of data, considering such visualization issues as color and shading.

Geometry, on the other hand, deals with mathematical properties such as topology, metric, and order. Within the realm of geographic information systems, both aspects are important, however, with respect to storage and retrieval, the geometric properties are the crucial factors. Geometry, not graphics, has to be stored in order to maintain consistent geographic databases and to process complex spatial queries at the necessary level of detail.

4.2 Distribution of Geographic Objects in Space

Geographic space is irregular and does not follow any preconceived patterns. With respect to storing and retrieving spatial objects in geographic databases, two aspects are relevant:

- The density of objects varies across geographic space. For example, in urban areas highway exits and entrances are every 1-2 miles, whereas in rural, less populated areas there may be 20-30 miles between two exits.
- The spatial extent of geographic objects may vary dramatically. Compare the sizes of I95—which extends from Maine to Florida—with the size of the town of Orono, both being stored in the same geographic database.

4.3 Temporal Changes

Databases for car navigation should be dynamic. Updates may involve the change of attribute values (a new name for a road) as well as geometry (construction of a new road). Some of these changes may be temporary (a tunnel may be flooded; a road may be blocked due to some accident; a bridge under repair such that one lane is closed) and will have to be reset to their previous value after some time. Others are permanent, e.g., after construction has been completed, there will be a third lane. Very rarely does it occur that the geographic objects in a vehicle navigation system “move” or “rotate.” On the other hand, there are objects such as cars moving through geographic space.

Two types of updates of geographic data occur that have significantly different distribution cycles: instantaneous distribution of traffic congestion, accidents, roads under construction, etc. (received over satellite or cellular phone) and periodic updates of the base data (e.g., in the form of a new CD) with newly constructed roads, updated road names, and most recent heuristics for routing.

4.4 Data Volume

By most standards, car navigation databases are very large, although certainly not as large as the terabyte databases generated from remote sensing. Their data volume is primarily governed by the tremendous number of individual (geometric) parts stored. In object-centered geographic databases, individual records are usually quite small—a point may have an integer pair to describe its x- and y-coordinates and some number (or string) to identify it. Besides this attribute information, geometric objects usually have many references (pointers) to other objects. A typical data source for vehicle navigation in the U.S. are the Census’ TIGER files, which provide sufficient detail for the driver level. The entire set

comprises some 600 MB of topologically structured vector data about the geometry of roads, geographic attribute data, and additional features such as county boundaries. The street network for a vehicle navigation system with its three different levels is generally too complex to be derived "on the fly" from a single representation; therefore, additional data volumes are created.

5. Challenging Database Issues

Current database technology falls short of the demanding needs of vehicle navigation systems. The requirements may be summarized as, "fast (real-time) retrieval of geographic data at different levels of detail." This combination of geographic data and temporal (processing) constraints is unique for navigation applications.

5.1 Real-Time Access to Geographic Data

Database management systems for vehicle navigation must cope with the ever-changing position of the car, which is the primary reference for continuous querying. It is crucial that certain types of queries are answered *on time* (e.g., before the driver reaches the exit). The amount of data to be retrieved while driving along the road is a factor of (1) the data density, (2) the scale at which the data must be displayed, and (3) the travel speed. Fortunately, drivers travel at a slower speed when they need more detail.

It is also necessary that the database management system and the application are closely integrated as they have to share resources (storage and CPU). The application may set priorities for the use of resources, e.g., if the connection to the satellites that provide positional information gets lost, more resources will be necessary to calculate the current position.

5.2 Query Languages and Query Processing

Most queries in vehicle navigation have to deal with spatial concepts; therefore, query languages must be extensible to allow for definitions of high-level spatial concepts and relations (Egenhofer 1992). When driving along the road, for example, queries are most naturally expressed as "navigation through space from object to object." Likewise, route planning is primarily based on such spatial relations as connectivity. On the other hand, the "map presentation" as the query result needs the "range search" to retrieve the relevant geographic objects within a certain spatial query window, which is periodically updated.

An exciting opportunity exists as one is able to predict what data will be necessary in the future as the driver moves along a chosen route. Given enough RAM, upcoming road segments may be accessed while the system is idle so that the detailed information will be available once the driver gets close to those road segments.

Multiple representations exist for the same geographic object in a vehicle navigation database, sometimes at different levels of detail. This has implications for query languages and query processing—which representation to select—as well as for database maintenance with complex consistency constraints among redundantly stored data. For example,

approximate query results may be sufficient in some situations (Barrera *et al.* 1992). For a driver in Maine who asks, "How far is it to Portland, Maine?" it is more important to get an answer quickly than to receive a very precise one. "About 30 miles" may be sufficient. On demand, refinements should be possible, if higher precision is required.

6. Conclusions

The goal of this paper was to present a challenge to the database community based on a complex real-world example within the domain of geographic databases. We selected the application of vehicle navigation as it integrates uniquely the demanding tasks of processing very large amounts of geographic data in real time. The scenarios described in this paper may be used as a yardstick for database designers to decide whether their systems can cope with the demanding tasks of managing and maintaining data for a complex geographic application.

7. Acknowledgments

Discussions with Jayant Sharma, Sabine Timpf, Gary Volta, David Pollock, and Langley Willauer provided valuable insight into the peculiarities of wayfinding in geographic space. Kathleen Hornsby helped with the preparation of the manuscript. An earlier version of the motivating example was part of two tutorials on "Geographic Databases: The Issues and Some of the Solutions" held at EDBT '92 and the Fifth International Symposium on Spatial Data Handling with Andrew Frank and Werner Kuhn. I am grateful for their contributions.

8. References

- W. Aref and H. Samet (1991) Optimization Strategies for Spatial Query Processing. *17th International Conference on Very Large Data Bases*, Barcelona, Spain, pp. 81-90.
- G. Ausiello, Ed. (1991) Basic GOODS, International Workshop on Database Management Systems for Geographical Applications, Capri, Italy.
- R. Barrera, M. Egenhofer, and A. Frank (1992) Robust Evaluation of Spatial Queries. *Fifth International Symposium on Spatial Data Handling*, Charleston, SC, pp. 241-248.
- L. Becker and H. Güting (1992) Rule-Based Optimization and Query Processing in an Extensible Geometric Database System. *ACM Transactions on Database Systems* 17(2): 247-303.
- A. Buchmann, O. Günther, T. Smith, and Y. Wang, Eds. (1989) *Symposium on the Design and Implementation of Large Spatial Databases*, Santa Barbara, CA. Lecture Notes in Computer Science Vol. 409. Springer-Verlag, New York, NY.
- M. Egenhofer (1992) Why not SQL! *International Journal for Geographical Information Systems* 6(2): 71-85.
- M. Egenhofer and A. Frank (1988) Towards a Spatial Query Language: User Interface Considerations. *14th*

- International Conference on Very Large Data Bases*, Los Angeles, CA, pp. 124-133.
- A. Frank (1981) Applications of DBMS to Land Information Systems. *Seventh International Conference on Very Large Data Bases*, Cannes, France, pp. 448-453.
- J. French, A. Jones, and J. Pfaltz (1990) Summary of the Final Report of the NSF Workshop on Scientific Database Management. *SIGMOD RECORD* 19(4): 32-40.
- A. Go, M. Stonebraker, and C. Williams (1975) An Approach to Implementing a Geo-Data System. Electronics Research Laboratory, University of California, Technical Report Memo ERL-M529.
- O. Günther and A. Buchmann (1990) Research Issues in Spatial Databases. *SIGMOD RECORD* 19(4): 61-68.
- O. Günther and H.-J. Schek, Eds. (1991) *Advances in Spatial Databases—Second Symposium, SSD '91, Zurich, Switzerland*. Lecture Notes in Computer Science Vol. 525. Springer-Verlag, New York, NY.
- S. Gupta and M. Stonebraker (1992) The Sequoia 2000 Approach to Managing Large Spatial Object Databases. *Fifth International Symposium on Spatial Data Handling*, Charleston, SC, pp. 642-651.
- R. Güting (1988) Geo-Relational Algebra: A Model and Query Language for Geometric Database Systems. in: J. Schmidt, S. Ceri, and M. Missikoff (Eds.), *Advances in Database Technology—EDBT '88 International Conference on Extending Database Technology, Venice, Italy*. Lecture Notes in Computer Science Vol. 303, pp. 506-527, Springer-Verlag, New York, NY.
- R. Güting (1989) Gral: An Extensible Relational Database System for Geometric Applications. *15th International Conference on Very Large Data Bases*, Amsterdam, pp. 33-44.
- B. Kuipers (1978) Modeling Spatial Knowledge. *Cognitive Science* 2: 129-153.
- B. Kuipers and T. Levitt (1988) Navigation and Mapping in Large-Scale Space. *AI Magazine* 9(2): 25-46.
- G. Lohman, J. Stoltzfus, A. Benson, M. Martin, and A. Cardenas (1983) Remotely-Sensed Geophysical Databases: Experience and Implications for Generalized DBMS. *SIGMOD 83 Annual Meeting*, San Jose, CA, pp. 146-160.
- E. Neuhold and M. Stonebraker (1988) Future Directions in DBMS Research. International Computer Science Institute, Berkeley, CA, Technical Report TR-88-001.
- B. Ooi and R. Sacks-Davis (1989) Query Optimization in an Extended DBMS. in: W. Litwin and H.-J. Schek (Eds.), *Third International Conference on Foundations of Data Organization and Algorithms (FODO), Paris, France*. Lecture Notes in Computer Science Vol. 367, pp. 48-63, Springer-Verlag, New York, NY.
- B. Ooi, R. Sacks-Davis, and K. McDonnell (1989) Extending a DBMS for Geographic Applications. *IEEE Fifth International Conference on Data Engineering*, Los Angeles, CA, pp. 590-597.
- N. Roussopoulos and D. Leifker (1985) Direct Spatial Search on Pictorial Databases Using Packed R-Trees. *International Conference on Management of Data, SIGMOD RECORD* 14(4): 17-31.
- H. Samet (1989) *The Design and Analysis of Spatial Data Structures*. Addison-Wesley Publishing Company, Reading, MA.
- T. Smith and A. Frank (1990) Very Large Spatial Databases: Report from the Specialist Meeting. *Journal of Visual Languages and Computing* 1(3): 291-309.
- S. Timpf, G. Volta, D. W. Pollock, and M. Egenhofer (1992) A Conceptual Model of Wayfinding Using Multiple Levels of Abstraction. in: A. Frank, I. Campari, and U. Formentini (Eds.), *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, Pisa, Italy*. Lecture Notes in Computer Science Vol. 639, pp. 348-367, Springer-Verlag, New York, NY.
- M. White (1991) Car Navigation Systems. in: D. Maguire, M. Goodchild, and D. Rhind (Eds.), *Geographical Information Systems*. Volume 2: Applications, pp. 115-125, Longman, London.