STATE AND TIME TOPOLOGIES FOR GEOGRAPHIC INFORMATION

Hazem Raafat, Zhongsen Yang
Department of Computer Science
University of Regina
Regina, Saskatchewan
Canada S4S 0A2
and
David Gauthier
Department of Geography
University of Regina
Regina, Saskatchewan
Canada S4S 0A2

ABSTRACT

Research on time and data models has focused mainly on the identification of extensions to the conventional relational model for non-spatial data. Although these models provide adequate semantic capability to deal with time, they are not suitable for spatial data such as geographical information in which historical information must be spatially referenced. This paper proposes two-level state topologies: a state topology for geographic objects in a GIS database and a state topology for a geographic object. From a temporal perspective, these two-level state topologies may also be viewed as two-level time topologies: a time topology for a GIS database and a time topology for a geographic object. Based on these state and time topologies, the storage approach for geographic historical information are provided.

KEY WORDS: State Topology, Time Topology, Geographic Information Systems

1. INTRODUCTION

More and more it is being realized that the element of time should be introduced into data models in order to represent the dynamically changing world (Snodgrass 1990, SOO 1991, Stam *et al.* 1988). The goal of historical databases is to make the time dimension accessible to users. Snodgrass and Ahn (1985, 1986) have introduced two important aspects of time: *world time* (or valid time) and *system time* (or transaction time). They can be represented by two axes. The world-time axis traces the changes which occur in the real world, and the system-time axis traces the changes that are recorded in the database. A historical database only contains world time. A temporal database contains both world time and system time. In this paper, we focus on historical database for GIS.

There are three possible approaches to include world time into the relations: *relation-based world time stamping*, *tuple-based world time stamping*, and *attribute-based world time stamping*. In the relation-based world time stamping approach (Klopprogge 1981, Mckenzie *et al.* 1987), each relation includes a world time interval during which the data in the relation is effective. The approach creates and stores a new snapshot of a relation when any of its attribute values changes. This approach is simple, but highly data redundant and obscures individual object histories. Tuple-based world time stamping approach (Ariav 1986, Lum *et al.* 1984, Sarda 1990a, Sarda 1990b) maintains a world time interval for each tuple. Whenever any of the attribute values of a tuple changes, its tuple-time stamp is amended and a new tuple may be appended to the relation. Consequently, each relation contains the history for each tuple. This approach is mostly used for representation and implementation of time modelling. One tuple-based time stamping method (Ariav 1986) orders tuples within each relation. Another tuple-based time stamping method (Lum *et al.* 1984) uses two relations to segregate current data from historical data and connect them by *history chains*. Attribute-based world time stamping approach (Gadia 1988) maintains a world-time interval for each attribute value. Thus, each tuple contains a history for each attribute. Although this approach is compact, it requires variable-length fields of a complex domain to hold lists of time-stamped attribute versions and needs an alternate relational algebra to manage them.

The historical database attempts to model an enterprise over time, but it is not suitable for spatial applications which deal not only with thematic and time information, but also with location and topological information. In recent years, more attention has been directed to temporal/historical GIS design related to vector data structures (Langran 1989a, Langran 1989b, Langran *et al.* 1988).

The earliest historical GIS was designed by Basoglu and Morrison (1978). They produced a hierarchical data structure to store and retrieve the historical changes of U.S. county boundaries. Although the system could produce a snapshot of how the particular boundaries appeared on a given time, it did not represent widely-used topological relationships and could not recognize that one line segment might be no longer a particular county boundary, but remain in use as another county boundary through historical subdivision.

Langran and Chrisman (1988) proposed a *space-time composite* data model to treat spatial changes over time. In this conceptual model, each change causes the changed portion of the coverage to break from its parent object and become a discrete object with its own distinct history. Therefore, this method decomposes the object over time into increasingly smaller fragments (objects) and describes them by a variable-length list of attribute sets bracketed by effective dates.

Other researchers are also contributing to temporal GIS. Armstrong (1988) considered time in spatial databases, and developed a framework for incorporating temporality in spatial databases. Worboys (1990a, 1990b) discussed the role of modal logics in a GIS. Hunter and Williamson (1990) proposed a method of storing and processing temporal geographic data by addition of time-encoding attributes to data elements as required and developed a historical digital cadastral database to demonstrate their method. Henrichsen (1986) studied Norwegian Socioeconomic Database which implements time encoding and handles administrative boundaries for the period 1770-1980.

Since there is complexity in the time-semantic representation, most GIS databases now in use do not model time attributes. Therefore, in order to retain historical geographic information, a set of snapshot sequences (data versions) are separately stored in the GIS database. This approach not only causes a tremendous amount of data duplication but also cannot support historical queries over time. Furthermore, most commercial GIS usually use the relational database system to store thematic and topological information, but use file systems to store geographic location information. Thus, semantic and data-granularity mismatches exist between the manipulation of the relational database and the manipulation of the file systems. A relational database system employs a set-oriented, declarative query language, in which the user requests a set of tuples without specifying the detailed steps to obtain this result. In contrast, the file systems usually use a tuple-oriented, procedural programming language, in which the problem solution is expressed as a sequence of detail operations on a global state. The crux of these problems is the lack of a sound temporal/historical spatial data model for GIS.

This paper first reviews temporal/historical database research on both spatial and non-spatial applications. Two-level state topologies for geographic information are proposed. The state topologies are then viewed from the point of time of view, and two-level time topologies are defined. Finally, an approach for representation of the state topologies and the time topologies with historical relations is developed.

2. TWO-LEVEL STATE TOPOLOGIES FOR GEOGRAPHIC INFORMATION

2.1 The State Topology for Geographic Objects

The changes of a geographic object may be viewed as the changes of its states over time. The historical information of a geographic object may be represented by the collection of its states. A mutation (or change) would transform the geographic object from one state to the next (Langran *et al.* 1988). A GIS database may be viewed as the collection of the states of all the geographic objects concerned and transformed from version to version by any of the object mutations.

The duration of a state, which started at time instant Ti, and ended at time instant Tj (not including Tj), is represented as [Ti, Tj). Ti is called *start instant,* and Tj is called *end instant.* The state is called a *historical state.* The duration of a state, which started at time instant Ti, and is now still active, is represented as [Ti, NOW). We consider "NOW" as a moving time variable, and the state as an *active state.* Historical states cannot change, only active states can change into historical states. Figure 2.1 shows the topology of states and mutations for geographic objects in a GIS database.

Each state line in Figure 2.1 is punctuated by the object mutations, and represents the states of a geographic object over time. Two states which share a boundary may be viewed as contiguous neighbors. The states for all geographic objects may be viewed as a topology comprised of these parallel state lines.

2.2 The State Topology for a Geographic Object

A mutation of a geographic object may be caused by spatial change or thematic change. Therefore, the states of a geographic object may be viewed as the composition of the

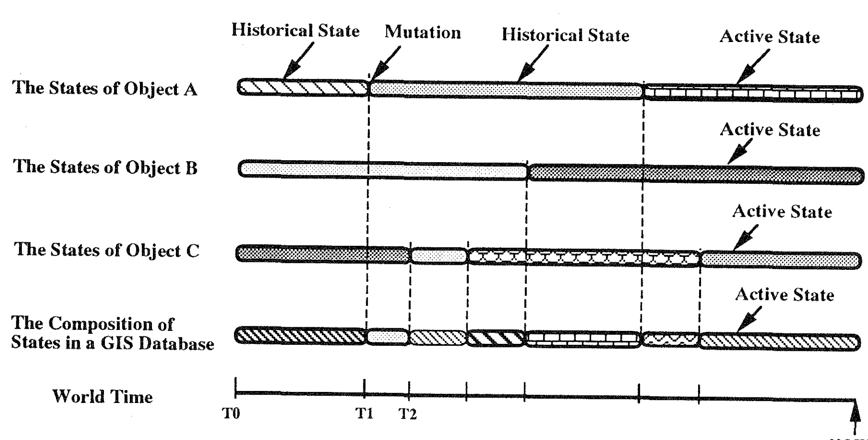

Figure 2.1 The Topology of Geographic Object States and Mutations

object's spatial states and thematic states. A state line in Figure 2.1 may be viewed as the composition of the object's parallel spatial state lines and thematic state lines. For example, suppose the object has a point feature, such as gas station A, shown in Figure 2.2. At time T0, station A was created. At time T1, the station was moved to a new location. At time T2, the ownership of the station was changed, and the new owner moved the station to a new location at time T3. Finally, at time T4, the station was closed. Figure 2.3 shows the state topology of station A.

Another example is shown in Figure 2.4. Polygon A represents grass-land and was created at time T0. At time T1, a part of grass-land A became farming land (arc a1 was changed into arc a11). At time T2, all of farming land B became grass-land (arc a2 was removed, and arc a5 became an arc of polygon A). At time T3, polygon D was split into two polygons, and caused arc a4 to be split into arc a41 and arc a42. At time T4, the ownership of grass-land A was changed. And finally, at time T5, all of grass-land A became a commercial area, and was combined with commercial area C. This means that grass-land A does not exist any more. The state topology of polygon A is shown in Figure 2.5, where we assume that each arc has only one state. That is, any change of the arc is viewed as the death of the arc, and the birth of zero or more new arcs.

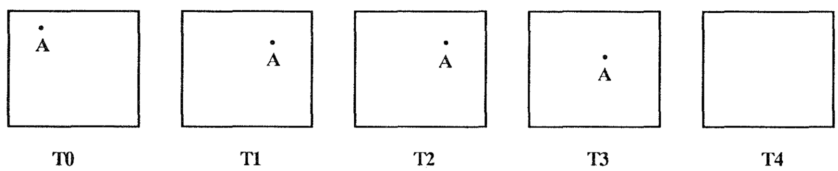

Figure 2.2 The Mutations of Gas Station A

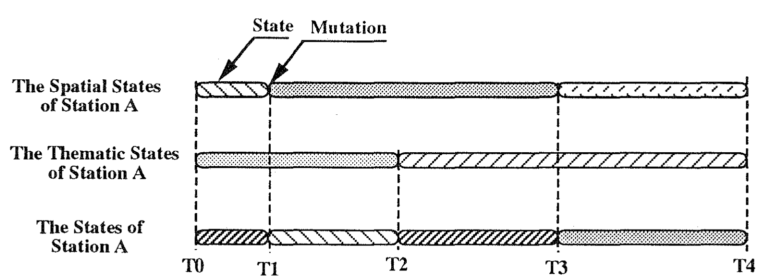

Figure 2.3 The State Topology of Gas Station A

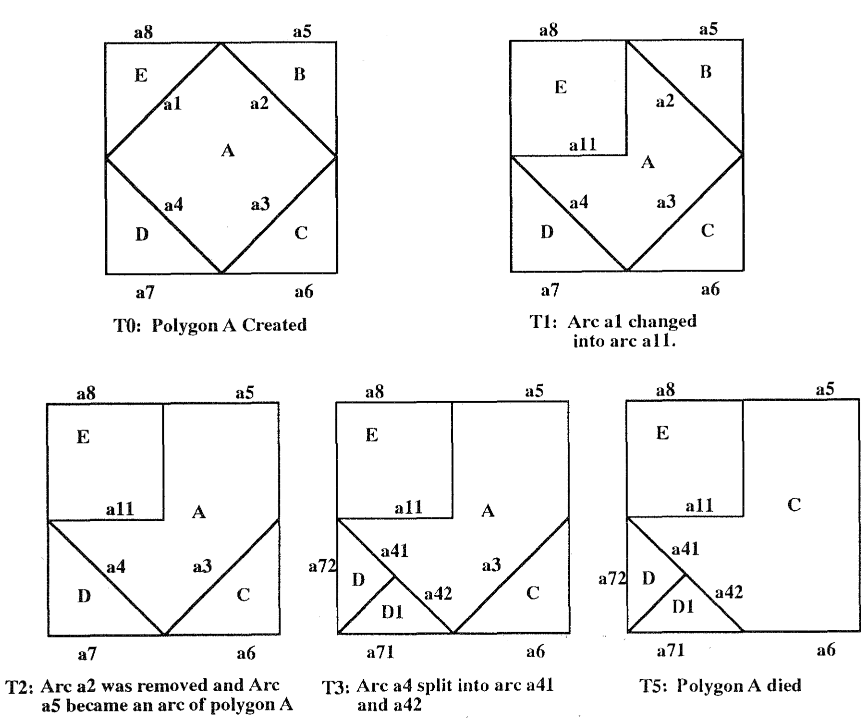

Figure 2.4 The Spatial Changes of Polygon A Over Time

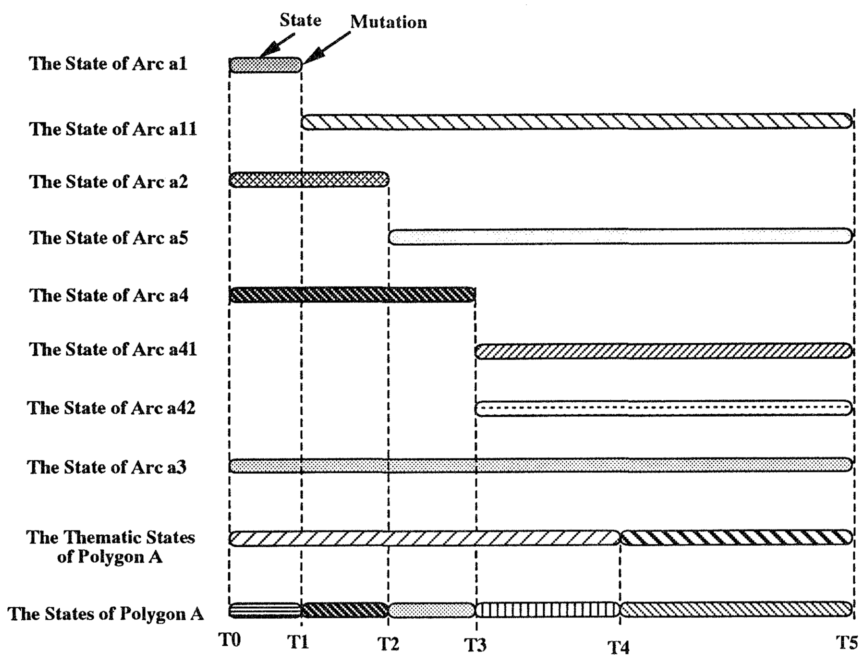

Figure 2.5 The State Topology of Polygon A

3. TWO-LEVEL TIME TOPOLOGIES FOR GEOGRAPHIC INFORMATION

In the real world, we are accustomed to regarding time as a line without endpoints that stretches infinitely into the past and future. In practice, world time represented in databases would begin at the time of the earliest known information and end at the time of the most recent information stored in the database.

3.1 The Time Topology for Geographic Objects

Time is a phenomenon and can only be perceived by its effects. From the point of time of view, every object has a beginning at some point in time. It also has a lifespan during which the object's location or theme may change independently of the others. Finally, it may die in the sense that it does not exist in the real world or has changed into another object. The state topology in Figure 2.1 may be viewed as a time topology for objects. World time for a geographic object may be viewed as a line, and is punctuated by the object mutations. An object state may be viewed as a line segment that represents the duration of a condition, while a mutation is a point that terminates the condition and begins the next. Two line segments which share a boundary may be viewed as contiguous neighbors in time. World time for all geographic objects in a GIS database may be viewed as a topology comprised of these parallel time lines as shown in Figure 3.1.

3.2 The Time Topology for a Geographic Object

From the point of time of view, the state topology for an geographic object, such as the state topology of polygon A in Figure 2.5, may be viewed as time topology for the object. That is, each time line in Figure 3.1 may be viewed as the composition of an object's spatial time durations and thematic time durations. For example, the time line of polygon A may be viewed as the composition of the spatial time durations and the thematic time durations of polygon A, as shown in Figure 3.2.

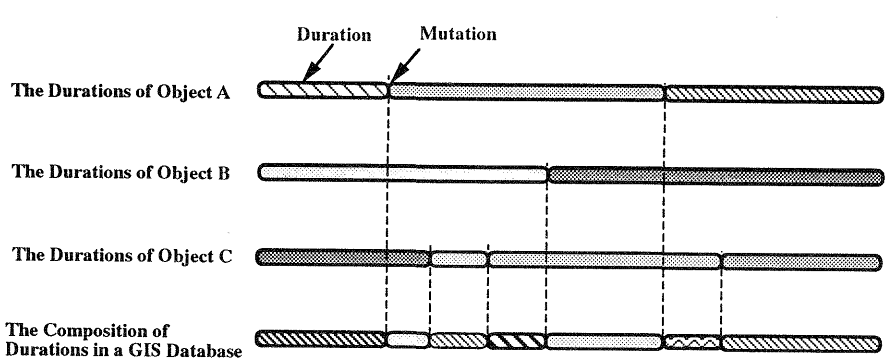

Figure 3.1 The Time Topology of Geographic Objects

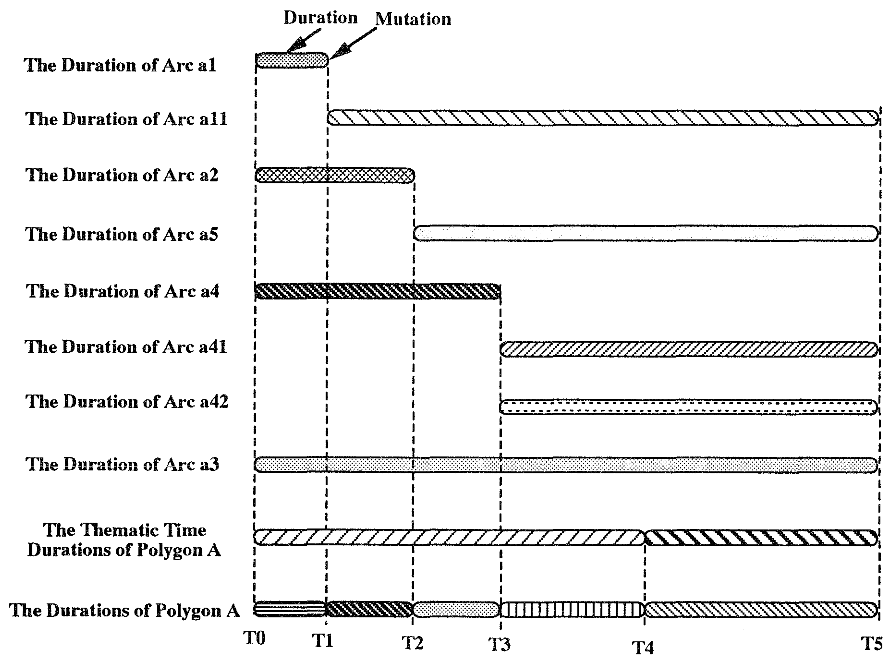

Figure 3.2 The Time Topology of Polygon A

4. THE REPRESENTATION OF OBJECT STATES WITH HISTORICAL RELATIONS

The representation of object states here is based on the tuple-based world time stamping approach. The state and time topologies provide the fundamental for storing geographic information. The states of an object may be represented as the contents of the relations, while the time durations of the object may be represented as the tuples' intervals.

We keep the topological data, location data, and attribute data separately in different relations. Therefore, to record one kind of data does not duplicate the others, and to retrieve one kind of data does not require movement through the others. Furthermore, different topological relationships, such as polygon topology, arc topology, or node topology, may be independently stored, as are the attribute data. Consequently, to record or retrieve one kind of topological relationship does not require reference to the others.

We assume that each arc has only one state. The term a1[Ti, NOW) is used to show that arc a1 was born at time Ti, and is now still alive. A tuple which has time interval [Ti, Now) is called an *active tuple*. If, finally, arc a1 died at time Tj, then only its time duration needs to be amended to a1[Ti, Tj). The corresponding tuple is changed into a *historical tuple*. The historical tuples can only be retrieved; they cannot be changed. The active tuples can be retrieved, and also be amended into historical tuples. Therefore, we represent the evolution of a geographic object over time by recording its changing information. For example, the states of polygon A changes over time can be represented in database as Figure 4.1.

Since we use a set of relations to represent the states of a single data layer. If every tuple of these relations embeds with a world time interval, a large amount of time interval duplication will exist. In order to reduce this time interval redundancy, we classify the set of relations for a data layer into two categories, and embed the world time only on the tuples of the relations in one categary. Then use join to propagate the time attributes to the relations of another categary. For example, in Figure 4.2, although the relation ARC-ATTRIBUTE does not have time attributes, it contains the historical information, such as arc a2 which had changed into a'2 at time T1. The duration of arc a2 can be derived when relation ARC-ATTRIBUTE links with relation HIGHWAY-TOPOLOGY by their common attribute Are-ID. The details about time operations, historical relational algebra for GIS, and extending SQL for historical geographic databases are not discussed here. Researchers interested in these areas can refer to (Sarda 1990a, Sarda 1990b, Yang 1991, Yang *et al.* 1991).

Poly-ID	Arc-ID	From	To
A	a1	T0	NOW
A	a2	T0	NOW
A	a3	T0	NOW
A	a4	T0	NOW

The Topology of Polygon A at Time T0

Poly-ID	Arc-ID	From	To
A	a1	T0	***T1***
A	a2	T0	NOW
A	a3	T0	NOW
A	a4	T0	NOW
A	***a11***	***T1***	***NOW***

The Topology of Polygon A at Time T1

Poly-ID	Arc-ID	From	To
A	a1	T0	T1
A	a2	T0	***T2***
A	a3	T0	NOW
A	a4	T0	NOW
A	a11	T1	NOW
A	***a5***	***T2***	***NOW***

The Topology of Polygon A at Time T2

Poly-ID	Arc-ID	From	To
A	a1	T0	T1
A	a2	T0	T2
A	a3	T0	NOW
A	a4	T0	***T3***
A	a11	T1	NOW
A	a5	T2	NOW
A	***a41***	***T3***	***NOW***
A	***a42***	***T3***	***NOW***

The Topology of Polygon A at Time T3

Poly-ID	Arc-ID	From	To
A	a1	T0	T1
A	a2	T0	T2
A	***a3***	***T0***	***T5***
A	a4	T0	T3
A	***a11***	***T1***	***T5***
A	***a5***	***T2***	***T5***
A	***a41***	***T3***	***T5***
A	***a42***	***T3***	***T5***

The Topology of Polygon A at Time T5

Figure 4.1. The Representation of The Evolution of Polygon A in a Database

HIGHWAY-TOPOLOGY

Highway-ID	Arc-ID	From	To
H1	a1	T0	NOW
H1	**a2**	**T0**	**T1**
H1	a3	T0	NOW
H2	a6	T0	NOW
H3	a4	T0	NOW
H3	a5	T0	NOW
H1	**a'2**	**T1**	**NOW**

ARC-ATTRIBUTE

Arc-ID	Pavement Condition	Lanes
a1	Good	4
a2	**Fair**	**2**
a3	Fair	2
a4	Fair	2
a5	Good	4
a6	Fair	4
a'2	**Good**	**4**

Figure 4.2 The Propagation of Time Attributes From One Relation to Another

5. CONCLUSIONS

The storage of historical geographic information is based on the state and time topologies. The proposed framework is a modified tuple-based time stamping approach. Since a historical database manigement system (HDBMS) can be used to manipulate the historical geographic information, historical queries over time, as well as efficient data sharing, data integrity, and data security can be provided by HDBMS. The stored topological structures do not only simplify and speed the spatial searching, but also trap the data errors. This framework just records the changing information. When a new state emerges, this approach just records the new state, and amends the time duration of the corresponding old state. Therefore, the data redundancy is considerably reduced compared with the data version approach used in the current GIS.

REFERENCES

Ariav, G., 1986, A Temporally Oriented Data Model, *ACM Trans. Database Syst.* Vol. 11, No. 4, pp. 499-527.

Armstrong, M. P., 1988, Temporality in Spatial Databases, *Proceedings of GIS/LIS'88, ACSM, pp. 880-889.*

Basoglu, U., and Morrison, J., 1978, The efficient hierarchical data structure for the U. S. Historical Boundary File, *Harvard Papers for GIS,* Vol. 4, edited by G. Dutton, Reading, Massachusetts: Addison-Wesley.

Gadia, S. K., 1988, A Homogeneous Relational Model and Query Languages for Temporal Databases, *ACM Trans. Database Syst.,* Vol. 13, No. 4, pp. 418-448.

Henrichsen, B., 1986, Directions of Major Archives, *IASSIST Quarterly,* Vol. 3, pp. 3-6.

Hunter, G. J. and Williamson, I. P., 1990, The Development of a Historical Digital Cadastral Database, *International Journal of Geographical Information Systems,* Vol. 4, No. 2, pp. 169-179.

Klopprogge, M. R., 1981, TERM: an approach to include the time dimension in the entity-relationship model. in *Entity-Relationship Approach to Information Modeling and Analysis,* edited by P. P. S. Chen (Sangns, California: ER Institute), pp. 477-512.

Langran, G., and Chrisman, N., 1988, A Framework for Temporal Geographic Information, *Cartographica,* Vol. 25, No. 3, pp. 1-14.

Langran, G., 1989a, A review of temporal database research and its use in GIS Applications, *International Journal of Geographical Information Systems,* Vol. 3, NO. 3, pp. 215-232.

Langran, G., 1989b, Accessing Spatiotemporal Data in a Temporal GIS, *Auto-Carto 9: Ninth International Symposium on Computer-Assisted Cartography,* (Baltimore, Maryland), pp. 191-198.

Lum, V., Dadum, P., Erbe, R., Guenaeur, J., Pistor, P., Walch, G., Werner, H., and Woodfill, J., 1984, Designing DBMS support for the temporal dimension. *Proceedings of the SIGMOD '84 Conference,* held in June 1984 (New York: ACM), pp. 115-126.

Mckenzie, E., and Snodgrass, R., 1987, Extending the relational algebra to support transaction time. *Proceedings of the ACM SIGMOD'87 Conference,* held in San Francisco in May 1987 (New York: ACM), pp. 467-478.

Sarda, N. L., 1990a, Extensions to SQL for Historical Databases, *IEEE Trans. Knowledge and Data Engineering,* Vol. 2, No. 2, pp. 220-230.

Sarda, N. L., 1990b, Algebra and Query Language for A Historical Data Model, *The Computer Journal,* Vol. 33, No. 1, pp. 11-18.

Snodgrass, R., and Ahn, I., 1985, A taxonomy of time in databases. *Proceedings of the SIGMOD '85 Conference,* held in Austin, Texas, in May 1985 (New York: ACM), pp. 236-245.

Snodgrass, R., and Ahn, I., 1986, Temporal databases, *IEEE Computer Mag.,* Sept. 1986, pp. 35-42.

Snodgrass, R., 1990, Temporal Databases Status and Research Directions, *SIGMOD RECORD,* Vol. 19, No. 4, pp. 83-89.

Soo, M., 1991, Bibliography on Temporal Databases, *SIGMOD RECORD,* Vol. 20, No. 1, pp. 14-23.

Stam, R. and Snodgrass, R., 1988, A Bibliography on Temporal Databases, *IEEE Database Engineering,* Vol. 7, No. 4, pp. 231-239.

Worboys, M. F., 1990a, The Role of Modal Logics in the Description of Knowledge in a Geographical Information System, *Proceedings of Cognitive and Linguistic Aspects of Geographic Space,* (NATO ASI, Spain).

Worboys, M. F., 1990b, Reasoning about GIS Using Temporal and Dynamic Logics, *Proceedings of The Temporal GIS Workshop,* University of Maine, Orono, Maine, October 12-13, 1990.

Yang, Z. S., 1991, Historical Spatial Data Model for Geographic Information Systems, Master's Thesis, University of Regina, Regina, Sask., Canada.

Yang, Z. S., Raafat, H. and Gauthier, D., 1991, A Temporal Oriented Spatial Data Model for GIS, *Proceedings of the Canadian Conference on GIS-91,* Ottawa, Canada, March 18-22 pp. 571-583.