HANDLING IDENTITIES IN SPATIO-TEMPORAL DATABASES

A.A. Roshannejad, W. Kainz
Department of Geoinformatics, ITC
P.O. Box 6, 7500 AA Enschede, The Netherlands
Phone: +31 53 874449, Fax: +31 53 874335
Email: {Roshan, Kainz}@itc.nl

ABSTRACT

Rapid advances in the technology of Geographic Information Systems (GIS) have prompted a wide range of spatially referenced applications. However, conventional GISs are dealing with a static view of a continuously changing real world. Taking a snapshot of the world, can be quite useful for a number of applications, however, many other applications are suffering from the lack of incorporation of time in the GIS analysis. Therefore, the increasing demands for temporal capabilities of GISs have started a stream of research in this field. Undoubtedly, handling time is a direct consequence of handling changes. Therefore, change handling, as well as keeping track of tangible data are the most important conceptual and technical problems in spatio-temporal databases.

1. INTRODUCTION

The most desirable way of dealing with geographical phenomena is to model them as they exist in reality. Therefore, the closer a data model represents real-world phenomena, the more comprehensive it is. In other words, all applications are looking for a data model that is able to provide a better and more real perception of the real-world. At the same time, interesting developments in hard and software as well as increasing demands for more updated geo-data from a variety of spatially referenced applications pose a number of questions for current database technology.

Traditionally, maps are produced as an underlying concept which is "A 2 dimensional graphic image which shows the location of things in relation to the Earth surface at a given time" (Keates 1989). What makes a distinction between GIS and traditional line maps is the way of handling geographical data, how to retrieve data and the manner of analyzing data and extracting required information. One of the important questions, in this regard, is to incorporate time in spatial databases. Besides many conceptual inconsistencies between time and space (which makes this combination rather difficult), in order to reach a successful spatio-temporal database, a number of points need special attention and consideration.

One of the basic and fundamental technical problems is handling dynamic data. In fact, time can only be felt by observing changes in real-world phenomena. Therefore, handling time means handling changes. Without the existence of a change, passing time is meaningless. In a static spatial database, by data we mean the most recent snapshot of the world, while in spatio-temporal databases, the most recent data and their histories are not detachable. Therefore, a strong algorithm must be available to keep track of data histories.

At a first glance it seems that handling object identities is a technical problem, however, since we are dealing with concept of change (to be tracked) the problem turns to have a conceptual nature. Particularly, when the database is going to be used for a wide range of applications (for example nation wide), this becomes clearer, because the definition of change and the criterion to create a new object after an event are not unique.

This paper discusses the importance of object identity in spatio-temporal database design. The second section of the paper describes components of spatio-temporal data. Having explained the data and their components we proceed to define change by means of object components. Section 3 gives a general framework to describe change. Section 4 then distinguishes between object as such and the influence of applications to viewing the object. In section 5 a distinction between concept and representation of geographical objects is drawn. Section 6 looks at some research that has been done till now and, while describing the main problems of previous approaches, explains a new conceptual framework for spatio-temporal database design and defines guidelines to handle object identities in spatio-temporal databases. The paper concludes with a discussion of a different perspective to view objects.

2. COMPONENTS OF SPATIO-TEMPORAL DATA

The most noticeable underlying principle of the real world is its dynamic nature. Geo-referenced phenomena are carrying time as an undecomposable component. In fact, each spatial object is known as a composition of three components, called *spatial*, *temporal* and *non spatio-temporal* (or *aspatial*). This is why to point at each object correctly, we have to define a "*W triangle*". This triangle is composed of three W's. Each of the vertices of it points to either "*When*" or "*Where*" or "*What*" (figure 1).

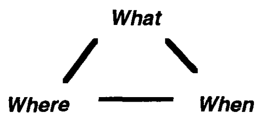

Figure 1. W Triangle

An unambiguous definition of a geographical object has to define these three W's. All the applications (using spatial data) in one way or another, are dealing with geographical object components. Although at the first glance it may be seen that some applications are not interested in all the mentioned components, however, a general and non confusing object identification has to address all three components. In GIS terminology, the geographical object and its components can be illustrated as given in figure 2.

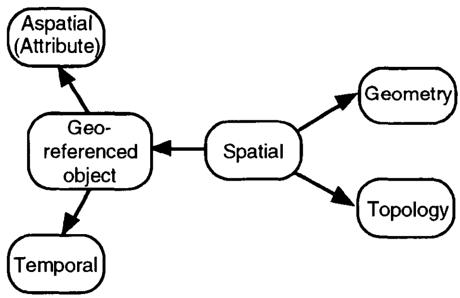

Figure 2. Components of Geo-referenced objects

Among these components, the temporal component is playing a special role. In fact the two others are affected by this one. When we are dealing with changes happening to spatial data, implicitly we are considering the changes in non temporal components. Geometry, topology, as well as attributes of an object are influenced by time. Therefore with a great confidence we can claim that there is no object that is definitely fixed with respect to passage of time. Now, it is clear why time needs to be incorporated in the handling of spatial data, because it is as important as the two other components. As an example, just imagine how the organization of spatial data would look like if we ignore spatial properties and structure of geographic objects, a mass of useless data.

Consideration of time as a component of geographical objects increases the possibility of analyzing objects along the time line and to keep track of their histories. Many applications are referring to time as a valuable criterion. For instance, cadastral applications need to retrieve cadastral data with time stamps. In this application, (for example) time of transferring the ownership of a land can be as important as the spatial attributes of it. Another example is environmental monitoring. Global warming, desertification, deforestation of the rain forest are just few examples of why time is badly required in today's applications. Besides those applications that are dealing with the history of the objects, a considerable number of applications are interested in predicting what will be the next status of the objects, or to determine a trend in the object's development. These applications use the object history to reach the desired goals.

3. WHAT MAKES A CHANGE?

The underlying principle of spatio-temporal databases is the problem of capturing changes in geo-referenced objects. Therefore, at first we have to define what can cause a change in an object. This, in turn, raises the question of what changes are the most desirable ones. This varies from one application to another. A particular application might be interested in aspatial change while the other is dealing with spatial change. Thus the origin of changes has to be identified, first.

As we discussed earlier, a geo-referenced object is a phenomenon that is extended in three directions called *spatial extent*, *temporal extent* and *aspatial extent*. This way of defining an object (as the composition of primitive components) solves many other problems. As long as all the mentioned components are constant in an object, it can be defined or claimed as an unchanged distinguished object with a specific identity. This is the concrete definition of *surrogate* that must, strictly speaking, be unique. As a consequence of this definition, we have to define a new object, with a new object identity, as soon as one of the object's components changes. This is because with respect to the above mentioned definitions, there is a distinction between the object with changed component(s) and the one, that originally holds the object identity. Each of these components is, potentially, subject to change. Some of them, like the temporal component, changes continuously while the others change only irregularly. Since the change in the temporal aspect of a geographical phenomenon is continuous, keeping track of it makes no sense. In fact, considering temporal change leaves us in a paradox, immediately. Therefore, we have to focus on changes of the two other components. Figure 3 gives a view of sources of change.

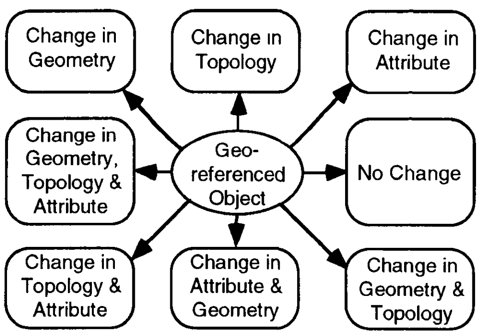

Figure 3. What can cause a change in a geographical object

This composition can help us to decide which component is the most desirable one for a certain application. Therefore, taking the right component for keeping track of the object's history will be the main issue.

In order to solve this problem, one can assign to the object a specific identity number at the time of creation. This number must be unique in all cases. As soon as a change occurs in one of the components, a new address (or identity number) must be assigned to it. The previous identity number will be kept in a history list and can be accessed through that list. In other words, each object will have a current identity number plus a list of previous identities. The list is arranged sequentially. The identity numbers can be regarded as objects themselves. These identity objects have a value (the current identity number) and also contain two addresses, one refers to the previous object and one refers to the next identity number.

Although this method of decomposing geographical objects into their components seems to be a wise solution, however, since different applications view the same object differently, it will be very difficult to decide if the change in a particular component is capable enough to create a new object. This leads to a conceptual problem of viewing geographical objets that needs to be studied more carefully.

4. DEPENDENCIES OF GEOGRAPHICAL OBJECTS FROM APPLICATIONS

As already mentioned, the definition of geographical objects is very dependent on applications. Therefore, the proposed approach in the previous section has a number of shortcomings in the definition of object creation by changes in its components. On the other hand, and with respect to ever changing components of spatial objects, this method will not be a practical solution to the problem of change tracking. This becomes clearer when we ask questions like *"What kinds of changes in the object's components make a new and distinguishable object?"* or *"When can a particular object at a certain time not be called the same object at the other time?"*.

By asking such questions we are likely to receive as many answers as applications are concerned. One application may regard a new object creation when the object moves few centimeters on the surface of the earth (change in spatial component) while the other is not interested in the precise spatial extent of geographical objects but their aspatial components is more noticeable. Even changing the color of a house (with respect to a particular application) may cause creation of a new object in the database.

In the case that the designed data model is going to serve a wide range of applications, we have to handle all the object component changes. For example the spatial component of a growing city is changing very frequently. At the same time its non spatio-temporal components (such as population) are not constant either. This means having a growing database. Of course this is not what we mean by an optimum approach. But in this example, the ever changing city gives the same perception of the city in people's view. Although, for example, Amsterdam today is different from Amsterdam 50 years ago the city is the same in concept. It seems that in order to define an object we have to deal with the concept and not the object's components. Kuhn (1994) intends to define objects by their own operations and properties instead of their linguistic references. This is exactly how we are approaching from a formal (and low-level) object definition and handling to a more proper way of dealing with objects as a self existent real-world phenomenon.

A great attention, nowadays, has also been paid to the way of communication. The stream of data transfers shows a shift from a fully visualized approach to a more comprehensive and complete communication. This must not be limited to the linguistic areas, but has to be extended to the high level conceptual data modeling. In data modeling, applying this approach means making a distinction between what an object is defined by and what represents the object (the mentioned components). This distinction can be the solution to handling changes, and consequently handling time, in a spatio-temporal database. This way of looking at geographical objects, first of all, avoids expansion of the database due to components' changes. The most important advantage of this approach is to get rid of application views to define geographical objects. This is how handling real-world phenomena should be.

5. CONCEPT VERSUS REPRESENTATION

The discussion about geographical objects and their components in a spatio-temporal database leads us to make a distinction between the concept that every object is defined by, and its different representations. In this respect, e.g., Amsterdam as a concept is a city with certain spatial and aspatial properties at each time stamp and also carries certain non changing characteristics which we call 'Amsterdam' as a concept. This concept is what people refer to when talking about the city of Amsterdam. On the other hand, Amsterdam is extended in three directions, as mentioned, spatial, temporal and aspatial. Each of these components can be used for a certain application and therefore, each of them represents the same object from a different perspective.

Thus if one wishes to obtain a more general view of an object, he/she has to use all the representation components at the same time. Of course, every representation can be used for a different purpose. This is how to perceive the object.

People use to classify objects according to their different representations. For example, all the cities in a certain region, all the 2 or more stories buildings, all the rivers with a certain width and so on. This classification helps to retrieve objects in a more convenient manner, however, creates a dangerous confusion that the objects can be replaced by their representations.

As another example, showing how the traditional way of thinking about the geographical objects limits us, we can refer to the importance of the spatial component of geographical objects in traditional analog line maps. Since in a graphical line map, every feature can be represented by its location (spatial extent), it is implicitly assumed that changing the spatial component of a geographical object causes creation of a new object. Therefore it is very important how to define a geographical object.

In order to more emphasize this topic Kuhn (1994) states that "A general assumption in today's thinking about data modeling for spatial data transfers is that spatial objects get their meaning by linking geometric primitives to feature-attribute catalogues". If we can get rid of object representations in the definition of objects, definitely we will have more freedom to handle objects as they are in the real-world. From this perspective, since it is not the spatial component that defines a geographical object, it does not matter if we find that Amsterdam had another spatial component (this spatial change of course is limited to a certain range) some time ago or some time later in the future, different from a certain snapshot of Amsterdam being used now. Its attributes can be different (at a particular time), but still Amsterdam exists independently of the changes in its components. Clearly an object is something independent of its representations.

6. HANDLING OBJECT IDENTITIES

In the spatial domain, an object has geometric as well as thematic components. When dealing with a static spatial database, any unique identifier given to the spatial object would be quite sufficient to distinguish the objects among each other. In the case of handling spatio-temporal data, as spatial data are subject to change over time, and due to having the possibility of accessing the history of the objects, it is very important to indicate and define the criteria which cause creating a new object and consequently, to define which component of the object has to carry that unique identifier.

This problem becomes critical when splitting or unification of object(s) occur. Obviously, in order to keep track of an object along the time line, one has to give it a unique time-invariant identifier. Clifford and Croker (1988) propose a definition of so-called primitive objects that hold a unique and time-invariant identifier with a non ending lifespan. Therefore, any complex object can have a unique identifier that can easily be constructed from a kind of aggregation of some relevant primitive object identifiers.

This can be applied in aspatial databases. Although this seems to be the only wise solution to the problem of object history tracking, however, in spatial databases, defining such primitive objects is not a straightforward approach. When we deal with the raster domain, this approach – to some extent – can still be applicable (where each pixel can be regarded as a primitive object). Although this seems easy to handle in the raster domain, one has always to consider that the next raster data set will not really register exactly the same as the previous one and furthermore, does not necessarily have the same resolution. Since many spatial databases have a vector-based structure, the problem of defining a unique identifier remains unsolved.

Another solution to this problem is proposed by Worboys (1992). His proposal is based on viewing spatio-temporal objects as finite collections of disjoint right prisms (so called ST-atoms). Bases of these prisms are spatial extents and their heights represent their temporal extents. Figure 4 shows changes that occur to spatial objects at time steps indicated by t_1, t_2 and t_3.

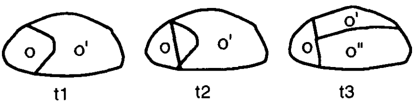

Figure 4. Typical changes occur to spatial objects

The decomposition of spatial objects into ST-atoms is shown in figure 5. Due to Worboy's proposal, the ST-objects O, O' and O'' are decomposed as follows:

- O is represented by the collection of two ST-atoms $< S_1,[t_1,t_2] >$ and $< S_2,[t_2,t_\infty] >$.
- O' is represented by the collection of three ST-atoms $< T_1,[t_1,t_2] >$, $< T_2,[t_2,t_3] >$ and $< T_3,[t_3,t_\infty] >$.
- O'' is represented by the single ST-atom $< U,[t_3,t_\infty] >$.

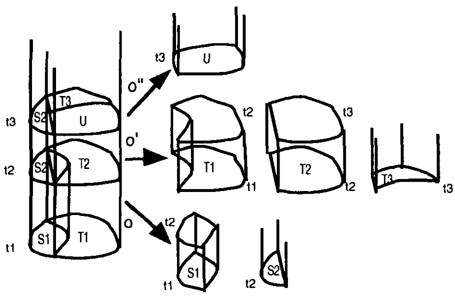

Figure 5. Decomposition of ST-objects to their ST-atoms (after Worboys (1992))

Identification of the objects now becomes confusing. The mentioned decomposition states that T_3 is a new version of T_2 which in turn has been T_1 at time t_1. Considering the spatial extent of T_3, it cannot potentially lead to the spatial extent of T_2 (and the same argument holds for T_2 and T_1). Worse than that is that the history of U is missing. Although U is a completely new object (as far as it is not born from nowhere and always there was something that is modified to make U), ignoring the history of U cannot be accepted. Assume that S_1 and S_2 are bare soil, T_1, T_2 and T_3 are agricultural areas and U is a residential building. Obviously, U is a new object (a building), but as a spatial phenomenon has a history. Simply it has been part of an agricultural area at time t_2.

Another problem can also be observed from the topological point of view. A spatial database must have rigid topological constraints that exactly define, for example, spatial neighbors of an object. This, in turn, creates the problem of maintenance of topological relationships, which is very difficult with the previous approach. These relationships, of course, are not only limited to spatial ones but cover the temporal domain as well. Handling objects as a composition of ST-atoms is not capable enough to keep track of changes in topological relationships. Furthermore, the most important problem in this approach is that ST-atoms are based upon the spatial extent.

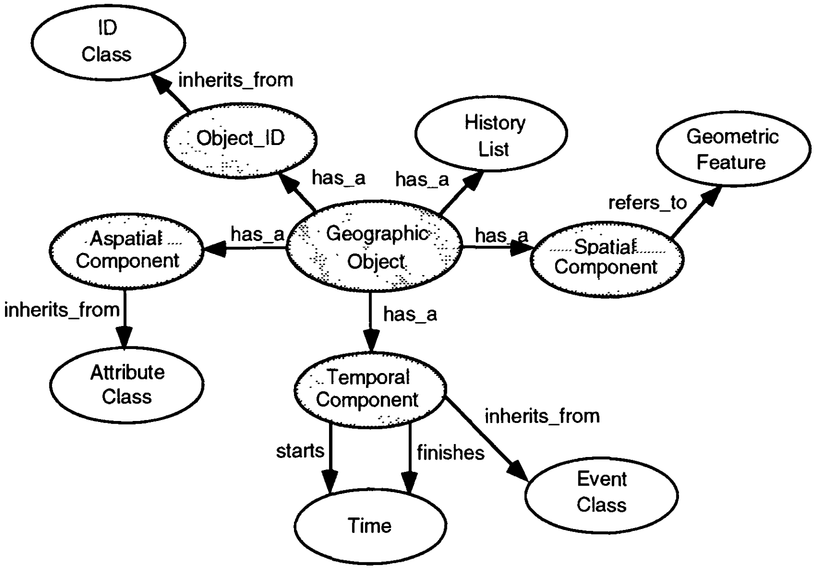

Figure 6. The required structure for defining geographical objects

As it can be seen in figure 6, the identity is attached to the object itself and therefore is independent of object component changes. In terms of object-oriented programming, the following scheme gives a brief description of the structure required to handle geographical objects independently of their components:

```
Geographical_Object
{
Properties:
        Object.Geometry      Refers to geometric features;
         {.....}             Operations on spatial extent of the object
        Object.History       Refers to event class;
         {.....}             Operations on temporal extent of the object
        Object.Attribute     Refers to the list of all possible attributes;
         {.....}             Operations on aspatial extent of the object
}
```

In this structure, the aspatial component inherits from a class called *attribute class*. The attribute class is an enumeration list of all possible attributes that potentially are capable to be attached to the object. The spatial component of the object keeps the geometrical properties of the object and inherits from a more general class that defines *geometric feature*. Geometric feature, in turn, can be defined in a clearer structure which has a proper construction to be able to cover all the required details of handling geometry of the geographic features. It is clear that both aspatial and spatial components are keeping relative histories of the relevant changes in aspatial and spatial properties of the object, respectively. Of course, each status of the object in terms of spatial or aspatial will be recorded in related components.

In this respect, a geographical object has a main body (defined by the concept of the object) and also has two tails, one originating at the spatial and the other at the aspatial component. The temporal component, on the other side, is used for keeping temporal topological relationships. Therefore, object identity, as far as the object's concept is not changed, is unique.

Up to this point, we can define two different types of change. The first type of change is called "*Constructive Change*" and the second type can be named "*Non-constructive Change*". As from the name of these two changes we can understand, constructive change is the change that creates a new object out of an existing one with a new concept different from the previous one. A non-constructive change is one that only changes a certain status of the object in terms of spatial or aspatial components. In non-constructive change, the object identity remains unchanged and only a new part is appended to the end of tail(s) of the object. Obviously, in order to distinguish between every piece of the tail and to be able to keep track of non-constructive changes, all those pieces of the tail(s) are attributed by a relevant time stamp. By this way, if analyzing history of non-constructive changes of the object is required, the mentioned tails can be searched easily.

The object identity of the object inherits from an *ID-class*. In this class all the history of constructive changes is kept. A rough structure of identity can be drawn as:

```
Type identity
value identity_number

Attribute
        identity : Previous_identity_number
        identity : Next_identity_number
End identity;
```

For an implementation of this approach object-oriented tools must be used, because many features such as aggregation, association and multiple inheritance are required. This is exactly what we expected from the definition of the conceptual framework. It is based upon an object-oriented modeling perspective. The base is objects and they are self-existent phenomena. The concept defines the object and not specific representations of the objects.

7. CONCLUSION

The approach described in this paper serves as a basis for a multi purpose data model required for a wide spectrum of applications. The most important advantage of this approach is that it reduces the required change capacity for object histories and that a quick expansion of the database will be avoided.

Viewing geographical objects in the manner described in this paper helps to handle object identities in a more proper and logical way, however, the importance of this approach is not limited to the topic of handling object identities in spatio-temporal databases. In fact, what this approach of perception of geographical objects can do is to get rid of representations of objects as criteria to define the objects. This concept, is a key element in approaching dynamic GIS. Existing GISs are suffering from modeling moving objects, only because the base for definition of geographical objects is the spatial extent of them. This shortcoming of the current GISs can be resolved by this approach. As this paper is part of an ongoing research, the results of applying this approach will be reported in forthcoming international events.

8. REFERENCES

Clifford, J. and Croker, A., 1988, Objects in Time, *IEEE Data Engineering*, Vol. 7, No. 4, pp. 189-196.

Keates, J.S., 1989, *Cartographic Design and Production*, Second edition, Longman.

Kuhn, W., 1994, Defining Semantics for Spatial Data Transfers, *Proceedings of the 6th International Symposium on Spatial Data Handling*, Edinburgh, UK, Vol. 2, pp. 973-987.

Worboys, M.F., 1992, Object-Oriented Models of Spatio-Temporal Information, *Proceedings of GIS/LIS'92*, San Jose, California, USA, pp. 825-834.