History in Spatial Databases

Henri J.G.L. Aalders
Delft University of Technology, Department of Geodesy, the Netherlands
Katholieke University Leuven, Faculty of Applied Sciences, Belgium
Phone: 00-31-15-278 1567, Fax: 00-31-15-278 2745
Email: h.aalders@geo.tudelft.nl

ABSTRACT

Time in spatial databases is an option that is often required by many applications in GIS for analysis of history, trends monitoring changes, etc. However, not many applications are available in the world that registers temporal attributes in their databases.

KEYWORDS

Spatial, time, history, databases, GIS, Open GIS, standards.

INTRODUCTION

The need for temporal data in spatial databases comes from the need for the registration of changes appearing in the world around us. Some registrations require by law the history of the changes; e.g. in some property registrations with a deed registration, the property shows only after researching the deed history for several decades to enable statements about the security of the present property.

Also in the natural environment the resource management the emphasis shifts from inventory and exploitation towards long term monitoring describing interrelation of the natural and human processes on an area.

All these processes require analysis of *change* (where, when related to which objects) and *trends* of change through time with the underlying processes and their spatial distribution as well as the frequency of repeating processes. Central in such analyses is the efficient storage of *spatio-temporal* data in databases, enabling effective study of the data and the processes.

Therefore, the functions of a spatio-temporal database management system should include:

- *inventory* to describe compete databases on theme, object or attribute and instance level;
- *analysis* to explain, exploit data and forecast possible outcomes of certain processes;
- *updating data* by superseding outdated information with new;
- *quality control* to monitor and evaluate new data staying consistent with old data;
- *scheduling* data changes by identifying threshold states, that triggers predefined actions;
- *display* for the generation of maps or tables of a temporal process.

A time line represents the concept of time. Events happening at a specific moment in time are represented as a point on the time line. There is one very specific moment on the time line: now; however, this moment always proceeds. A point in time occupies a position on the time line that can be identified in relation to a time reference system and distance can be measured. However time has a single dimension and distances can be measured in relation to the origin of the temporal reference system. Although time has

an absolute direction (always going forward), by the registration of time in a database time can be measured in two directions (forward, i.e. towards the future or backward, i.e. in the history).
Events from the past occur usually on a single time line. New events may occur on several different time lines related to the present, existing timeline and thus showing their mutual temporal relation (see fig. 1).

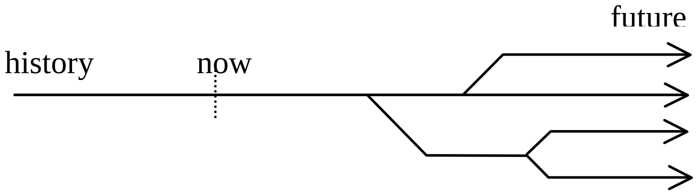


Figure 1: Time line splitting due to different developments.

Time is a dimension analogous to any of the spatial dimensions. Like space, time has a geometry and topology. Temporal geometry gives the position in time of an object, while temporal topology indicates the connectivity between temporal objects to order temporal objects in time using descriptions as after, before or the same (analogous to geometric topological characteristics as overlap, equal, crossing, adjacent inside, outside, etc.).
Time can be measured on two different scales: ordinal and interval. An ordinal scale provides only about relative position in time, while on an interval scale the duration can be indicated. So, time can be indicated by an instant or by duration (of a period). E.g. ISO 8601 [ISO 2000] describes a period of 5 days, 4 hours 30,7 minutes as P5DT4H30.7M, based on [ISO 1992].

DEFINITION OF TIME

Due to relatively short distance travelling, time definition has been a local matter for very long. With the development of railways and airplanes, long distance travelling became possible for many people, and thus requiring a uniform system for time. In 1884 during the International Meridian Conference, 25 countries decided to accept the meridian of Greenwich as 0º meridian and subdivide the world into 24 zones of 15º. Explicitly the conference advised to allow states to decide their own time; this is called ‘local time’.
For temporal databases the metadata specification should refer to the type of calendar used and within the calendar indicated the date (e.g. using the Gregorian calendar – 1582 - with year, month and day). On top, the time in 24 hours local or Co-ordinated Universal Time (UTC) subdivision for the time stamping of temporal object. (Other calendars can be defined in the metadata as the Julian calendar, the Modern Japanese calendar, the Ancient Babylonian calendar, or the GPS calendar, etc.)

TIME IN SPATIAL DATABASES

Time in a spatial database is one of the cornerstones of the description of processes that occur in reality. However, adding time as a concept in a spatial database increases the number of dimensions by one, i.e. a 2-D spatial database becomes three-dimensional, while a 3-D databases becomes four-dimensional, making all analysis and functions dimensionally more difficult in querying.

An important aspect of recording temporal data in a spatial database is the moment to be recorded and in fact several of these may have to be recorded in addition to each other. One can distinguish the moment of:

- designing the object;
- permission to build the object in the reality;
- change in the reality or in the database;
- observation of the realisation or construction in the reality (also photo-time);
- signature, registration, or postmarking the object;
- visualisation on screen or map.

Time registration for spatial phenomena (granularity) may be related to whole spatial datasets, object classes or object groups, object instances, or at attribute level. The application will require each of these or a combination of granularity. Also the smallest temporal unit (i.e. chronos) is defined by the application. Sometimes a part of a second may be important e.g. the position of wildlife animals, while in other situations a daily recording is suffices e.g. the transfer of a parcel's property or an era may be indicated e.g. the forest changes on a continent.

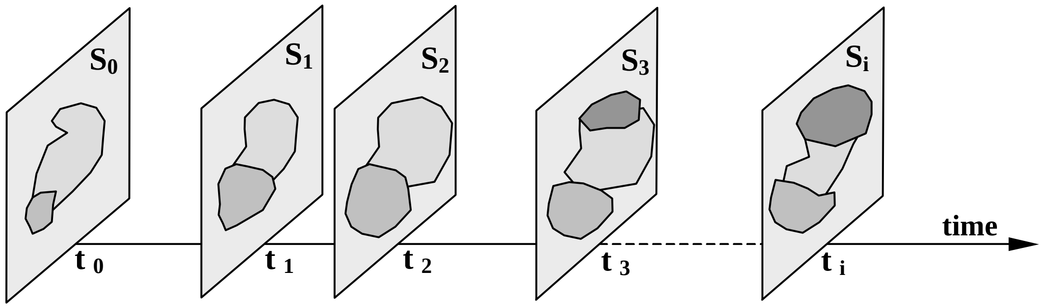


Figure 2: The snapshot approach.
Note: the time between snapshots need not to be the same.
.S_i represent the spatial distribution at a certain time t_i.

TYPE OF CHANGES

Changes in data may be discrete (as changes in ownership of land) or continuous (like natural changes in humidity, foreign exchange rates or stock market quotations. For developing a time-based database, the data model may be similar: in a discrete situation the change value can be the difference between the old and the new value at a certain time t_i; in a continuous situation the change relates often to a time interval and can be described by a formula, indicating the change over the period using t_i as starting value and t_j as the end value.

Specific temporal based queries should be used to evaluate the temporal based datasets as:

1. Retrieve all locations that changed to a given value at a given time.
2. Retrieve all locations that changed between t_i and t_j with their new values.
3. Calculate the total area change between t_i and t_j

APPROACHES TO TEMPORAL REPRESENTATION

Snap shot approach

Over the past decades, relational Database Management Systems (R-DBMS) have resulted in modern GIS's based on multi-representational database designs i.e. the same data recorded in a database can be represented in many ways according the intended applications. As with the spatial aspects in a GIS, also the temporal aspect in such databases should be well suited for the envisaged applications, to enable GIS's to be used for sophisticated temporal analyses effectively in future.

The only spatio-temporal model that is available in existing GIS's is a series of 'snapshots' [Peuquet 1995]. The representation then becomes the sequentially display of the different images after each other (see fig. 2).

Amendment vector approach

Another approach uses the representation changes in geographic objects overlaying them (see fig. 3). These models rely on the concept of amendments, where changes are incrementally recorded to an initial situation and the representation is in a single temporal composite map; the time that any change occurs is given as an attribute of each amendment. This organisation allows the integrity of individual features and their spatial interrelationships maintain over time [Langran 1992].

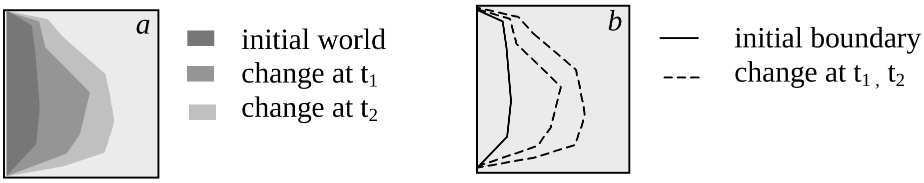


Figure 3: The amendment vector approach;
a) areal changes, b) amendment vectors.

4-dimensional approach

Hazelton [Hazelton 1990] proposed a 4 dimensional space-time Cartesian model with an extended hierarchy of nodes, lines, polygons, polyhedra and polytopes as the conceptual organisational basis. Each of these adds a single dimension to each other, making polytypes 4-Ddimensional enclosed features including 3-dimensional polyhedra with temporal information.

Following the general trend in IT object modelling, all of these models are feature based with thematic, spatial, topological and temporal attributes. An exception to this rule is the grid model, where temporal information is relative to the spatial position of the grid. In this sense tessellated spatio-temporal data is location based, since all information is stored relative to the specific grid location.

Time vector approach

In a time based approach location in time becomes the prime basis for registering changes. The sequences of events over time, representing the spatio-temporal change, is noted on a 'time-line' or 'temporal vector' and also indicated by the 'event list'.

Each location in time along the time line is linked to a particular set of features with their locations (space-time attributes) that changed at that particular moment and the feature's new value.

APPLICATIONS OF TIME REGISTRATION

ESTDM

The Event-based SpatioTemporal Data Model is proposed in [Peuquet 1995] and represents a specific example of the time-based vector approach that orders *temporally* changes to *locations* within a pre-specified geographical area.

ESTDM stores event lists at a specific temporal location t_i (t_i is called a time-stamp) where the associated changes in the event list relates to a single domain (e.g.: landuse or population). The set of changes in the event list shows a geographical location with the new value.

An disadvantages of this method is that the number of x,y,v triplets (where v represents the change at t_i) for raster type of geometry will be indicated for each pixel or other pixel representations as quads or runlenghts; for vector models this should be object based.

The structure for the data model uses pointers to connect adjacent and previous entries allowing easy addition of new change events.

The method specifically facilitates search and retrieval of specific temporal intervals both in forward and backward temporal direction and associated specific changes within these temporal intervals. This enables easy comparison of different temporal sequences in different domains within a GIS, or the same thematic domain in different geographical areas.

The hierarchical organisation of data in ESTDM allows looking at only the times at which events occur to establish an overall temporal pattern. Also the frequency and variability of change values, regardless their geographical location can be studied as well as location change.

Dutch Cadastre

Only about 10% of the objects change (of the about 6 000 000 parcels and 3 500 000 owners) in the Dutch situation of registering land property [Oosterom 1997], while change usually only occurs to full objects (either changes of parcels or changes of owners to parcels). Therefore, changes are recoded as attribute to an object (in the Dutch context parcels are called objects and owners are called subjects; they both form database objects) by adding two time-related attributes: `tmin` and `tmax`. The actual object description has a specific value `MAX-TIME` that is larger as any other time value.

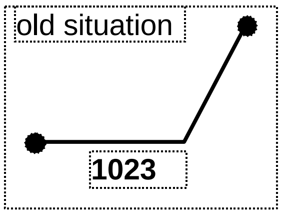
Distinction is made between system time (= transaction time, the time stamp of the actual change of the attribute value(s) in the database) and the ‘formal’ user time (= the time of the actual change of the object in reality), both in local time. `tmin` and `tmax` are both system times. The system also includes the formal time attribute `object-dt` or `valid-tmin`, the time stamp that the object was visited for a change. The attributes last `verification–dt` and `valid-tmax` may be added to the system in future in order to enable a bi-directional model, i.e. both the formal (local) time as the system time will supported.

Adding a new object to the dataset (see fig. 4), the value `tmin` is used for update value and `tmax` will get a special value `MAX-TIME`. Change of only an attribute value requires the creation of the whole object inclusive the `oid` (the unique object identification), which is a copy of the old object with the change of `oid`, the changed attribute value. The actual (system-) time for `tmax` in the old record is changed and also used for `tmin` in the new objects record. This allows a unique identification in space and time for every object through the combination of the attributes `oid` and `tmax`.

As the Dutch cadastral database includes topology for the parcel data, special attention is given to the topological linkage. This topological link refers to the `oid` and not to any time attribute. So, if any the topological change uses only the `oid` as a reference while the object itself is not changed the time attribute will also not be changed. However, when during an update the `oid` of an object changes because a new object is created a new version of the object is created and so all time attributes will be changed as indicated before.

A query on historic versions of a specific object (a standard query in a deed registered property registration) requires only the object's `oid` (and so avoiding to mention any time attribute). This does well for queries on simple object changes but not for changes as splitting parcel in several sub-parcels, combining several parcel into one parcel and other more complicated queries. However these queries are calculated form spatial overlap query.

Polyline 1023 (three point definition), created on Jan 12, 1998 is split on Feb 20, 2003 into tow objects: 1023 and 1268; on April 8, 2003 the quality of object 1268 is changed into level 2 quality.

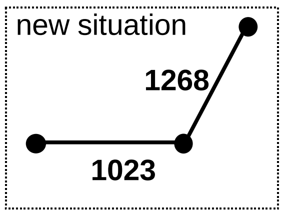


Situation per January 12, 1998

Oid	Geo	Q	t_{min}	t_{max}
1023	(0,0),(4,0),(6,2)	1	19980112	MAX-TIME

Situation per February 20, 2002

1023	(0,0),(4,0),(6,2)	1	19980112	20020220
1023	(0,0),(4,0)	1	20020220	MAX-TIME
1268	(4,0),(6,2)	1	20020220	MAX-TIME



Situation per April 8, 2003

1268	(4,0),(6,2)	2	20030408	MAX-TIME

Figure 4: Example of an update in the Netherlands Cadastral system.

During an update, a database should always change from a consistent situation into an new consistent database. Therefore, updates are performed in temporarily sub-datasets, extracted from the main database and marked '`for update`'. The database is transformed form the 'old' consistent situation into the 'new' consistent situation during the 'check-in' (update) and all changes of this single check-in will be assigned with the same `tmin` /`tmax`. This allows also for rollback operations on a very high level of operation.

Errors and mistakes, due to incorrect data mensuration or data input can be done by the creation of an `error-date`. So not the history is changed due to errors but additional error information is added as separate objects. Also changes due to geometric quality improvements of the data are not dealt with in the time change(s) because they are usually – many – very small changes and only provide users with the higher geometric quality.

Data dissemination only refers to the indication of and `t-end`. Update sub-datasets always contain two parts: the old situation (i.e. objects to be deleted) and the new situation (i.e. the new objects), assuming the a begin dataset is available at the customer's side.
Two different updates may be created:

1. the old situation at `t-begin` and the new situation at `t-end`, while the changed in between may be disregarded;
2. the old and new situation at `t-begin` and `t-end`, including all changes in between for historic information queries.

Other applications
In the previous time based applications were based on the actual change in an object registration. However, some applications like municipalities and utility companies deal with changes in objects that do not yet exist in reality. E.g. a house can be built and delivered at a certain time but before any construction in reality is done, the architect has designed the house and the constructor has asked for the permission to build the house. So in the municipal database objects have to be created about future objects in the reality. Since all of these refer to discrete changes, the above-indicated *time-based approach* for the design of a temporal database can also be applied for objects in 'pre existing' stage in reality.

CONCLUSIONS

Time registration in GIS enables very many potentials and many spatial applications require temporal information. However, the manners temporal information can be stored in spatial database are numerous: the ways, mentioned in this presentation are only a few. The way time is stored in the database and represented for different applications are closely related: the query type decides on the manner of storing temporal information. So, no temporal model is the best for every situation (same is true for spatial models: sometimes a raster model is the best, sometimes a vector model is the best). Also, the difference between handling natural versus man-made objects can determine the way time is recorded in a database.
However, not many applications are in use at present; the largest spatio-temporal database is the application of the Netherlands Cadastre, where the changes are discrete (at a certain time-stamp) and new records are replace the old ones, when a change occurs. To query the actual database is a simple operation based on time attribute query; forward change is taken care of, backward change requires extra links.

Temporal aspects in spatial data environment are becoming more accepted and standardised and become important to the use of databases for analysis, updating, quality control and visualisation. ISO TC211 and OpenGIS have recognised this and include temporal aspects in their standardisation process by developing ISO19108 [ISO 2002] and GML 3.0 [OGC 2003].

LITERATURE

Hazelton, N.W.J. , F.J. Laehy, I.P.Williamson (1990), On the design of temporally-referenced, 3-D geographical information systems: development of four-dimensional GIS, *Proceedings of GIS/LIS 1990*, Vol. 1, p.357-372, (ASPRS/ ACSM/URISA/AAG: Bethesda, MD, U.S.A.).
ISO (1992), *ISO 31*, Quantities and unit, Part 1: Space and time, Part 2: Periodic and

related phenomena, ISO 31: Geneva (CH) 1992.
ISO (2000), *ISO 8601*, Data elements and interchange formats – Information interchange, Representation of dates and times, ISO 8601: Geneva (CH) 2000.
ISO (2002), *ISO 19108*, Geographic Information/Geomatics - Temporal schema http://www.isotc211.org (accessed March 28, 2003).
Langran, G. (1990), Temporal GIS design tradeoffs, *Journal of URISA*, 1990. Vol 2, p. 16-25
OGC (2003), *OpenGIS*®, Geography Markup Language (GML), Implementation Specification, 3.0, http://www.opengis.org/techno/implementation.htm, (accessed April 2, 2003).
Oosterom, P.J.M. van (1997), Consistent bijhouden van topologie in een omvangrijke ruimtelijk-temporale database, *NGT Geodesia* 1997-6, p. 263-270 (in Dutch).
Peuquet, D.J. (1995), An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographic data, *International Journal Geographic Information Systems* 1995, Vol. 9, no. 1, p.7-24.