Topological and temporal modelling in GML

Wilko Quak, Marian de Vries
Delft University of Technology
Jaffalaan 9
2628 BX Delft
The Netherlands
w.quak@otb.tudelft.nl, m.d.vries@otb.tudelft.nl

Abstract

GML is becoming the world standard for the exchange of spatial data. Since the latest release of the GML standard (GML3) there is support for the modelling and exchange of topological data in GML. In this paper we take a closer look at the topological model in GML3 and discuss two issues that arose when we tried to encode the Dutch Cadastral model in GML.

The first issue that arose was that the topological model of the Cadastre is based on the winged edge data structure (which is used in many other systems, like for example Oracle Topology). The topological model of GML however is based on the boundary and co-boundary of topological primitives. Although it is very well possible to convert a winged-edge model to a co-boundary topological model this operation is not very efficient. In this paper we propose some minor changes to the GML topology model that makes efficient encoding of winged-edge data possible.

The second issue concerns the combination of topological and temporal modelling. In a temporal topology, a valid topology is maintained for any moment in time. This can be done by extending every node edge and face in the model with a time interval that indicates when that primitive is valid. In order to encode this model in GML, the temporal model of GML (Based on ISO19108) and the topological model of GML (based on ISO19107) need to be integrated in one model. However, this results in a few conflicts:

In the ISO topological model, topological primitives are grouped into topological complexes. One of the constraints on such a complex is that the geometries of the edges are not allowed to overlap. Because edges in the topological temporal model can overlap (as long as they are not valid at the same time) a topological temporal model stored in a GML topological complex will not result in a valid complex. In our paper we will propose to extend the ISO/GML topology model with a temporal topological complex.

Another problem that arises when storing a temporal topology in GML are the XLink references that are used to link between nodes, edges and faces: In a temporal topology different instances (at different times) of the primitives (nodes, edges, faces) can exist. As a result, the ID of the topological primitives is not unique anymore, which results in invalid GML documents. A solution where every new version of a primitive is given a new ID will not work either, because then it is not possible to give a reference to a primitive regardless of time. In our paper we will discuss on how to use XLink properly to find a solution to the problem.

1. Overview

The structure of this paper is as follows. Paragraph **Error! Reference source not found.** described the Dutch Cadastral temporal topological mode. Paragraph 0 describes the topological model of ISO/TC 211 as described in ISO19107 that is the basis for GML. Paragraph 0 handles one by one the conflicts that we found between these models when trying to express the cadastral model in GML. The paper ends with conclusions and suggestions in paragraph 0.

2. Introduction to Dutch Cadastral Model

The data model of Dutch Cadastre is a topological model, i.e. the geometry of the parcels is not stored, but can be inferred from the cadastral boundaries that are stored geometrically. The two most important tables in the cadastral database are the 'boundary' table with cadastral boundaries and the 'parcel' table with cadastral parcels. The concept behind this cadastral model is the winged edge topology, where parcel boundaries are edges in the topology and the parcels themselves are faces in the topology. The edges in the boundary table contain references to other edges according to the winged edge data structure (Baumgart, 1975), which are used to form the complete boundary chains (parcels). The edges also contain a reference to the left and right parcel (Oosterom and Lemmen, 2001).

Every parcel has exactly one reference to one of the surrounding boundaries and one reference to a boundary of each enclave. The structure of the topological references and the relationship between parcels and boundaries is visualised in **Figure 1**.

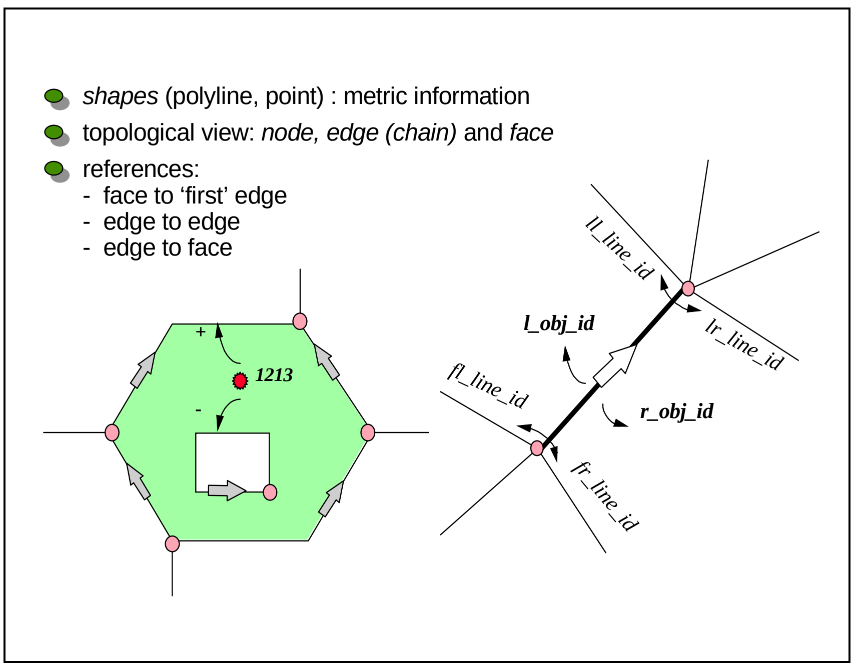


Figure 1 Topological model in the spatial DBMS of the Dutch Kadaster

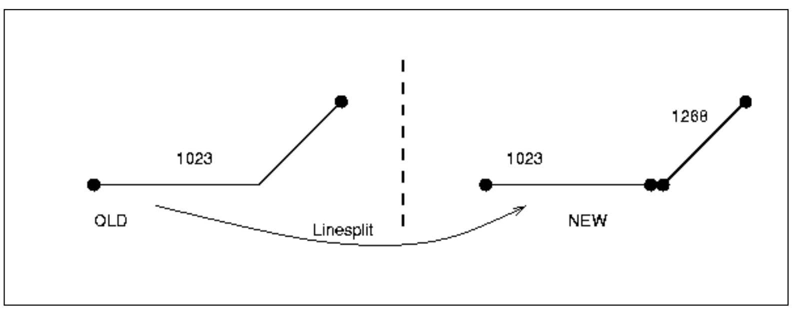
2.1 Time in the Cadastral model

In the cadastral model a complete history of all the parcels is kept in the database (oosterom, 1997). For this, all edges and faces are extended with two additional temporal attributes: tmin and tmax. These attribute indicate the time interval when the object is valid. Current object descriptions get a special value MAX_TIME, indicating that they are valid now. MAX_TIME is larger than any other time value.

When a new object is inserted, the current time is set as value for tmin, and tmax gets a special value: MAX_TIME. When an attribute of an existing object changes, this attribute is not updated, but the complete record, including the oid, is copied with the new attribute value. Current time is set as tmax in the old record and as tmin in the new record. This is necessary to be able to reconstruct the correct situation at any given point in history. As a result, there can be multiple instances of one object with the same object identifier. This means that the oid is not unique anymore. A unique identifier (key) is the pair (oid, tmin) for every object version in space and time.

For the topological references, only the oid is used to refer to another object and not tmin. In the situation that a referred object is updated and keeps its oid, then the reference (and therefore the current object) does not change. This avoids, in a topologically structured data set, the propagation of one changed object to all other objects as all objects are somehow connected to each other. In case the oid of a referred object has changed (becomes a different object), the referring object is also updated and a new version of the referring object is created.

The following example shows the contents of a database, which contained on 12 jan one line with oid 1023. On 20 feb this line was split into two parts: 1023 and 1268; see **Figure 2**.



oid	shape	tmin	tmax
1023	(0,0),(4,0),(6,2)	12 jan	20 feb
1023	(0,0),(4,0)	20 feb	MAX_Time
1268	(4,0),(6,2)	20 feb	MAX_Time

Figure 2 An edge is split in two parts

.

3. The ISO/TC 211 topological model

GML3.1 (ISO19136, 2004) is meant as an exchange format for the conceptual modelling framework described in the ISO19100 series. The topological model is described in ISO19107 Spatial Schema (2004). The model describes the correspondence of topological and geometric relationships up to 3 dimensions

In the ISO19107 model, there are four instantiable classes of primitive topology objects, one for each dimension up to 3D (TP_Node, TP_Edge, TP_Face and TP_Solid). In addition a topological complex (TP_Complex) is a set of topological primitives whose interiors are disjoint. Figure 3, Figure 4 and **Figure 5** show the UML models of the topological primitives. Because the Dutch Cadastre currently only models 2D geometries, TP_Solid is not used.

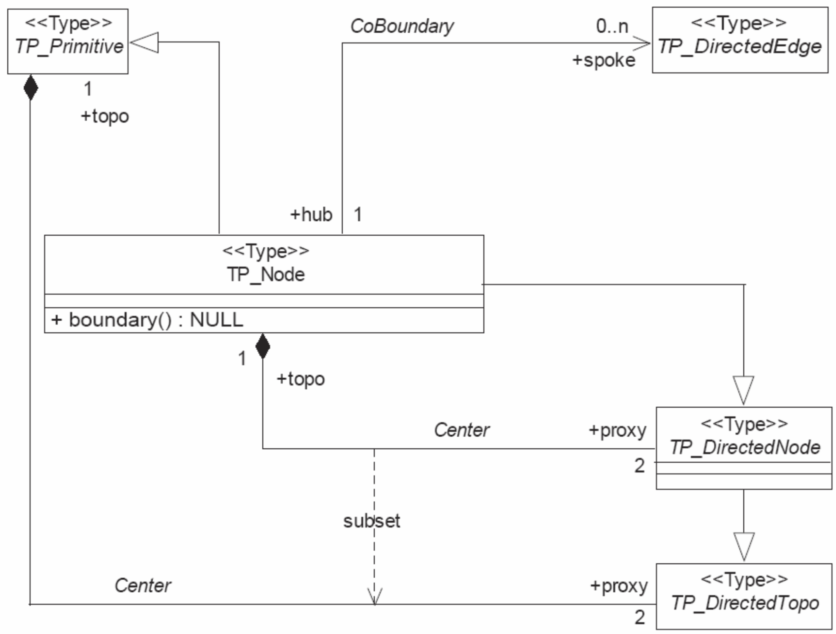


Figure 3 TP_Node (ISO 19107)

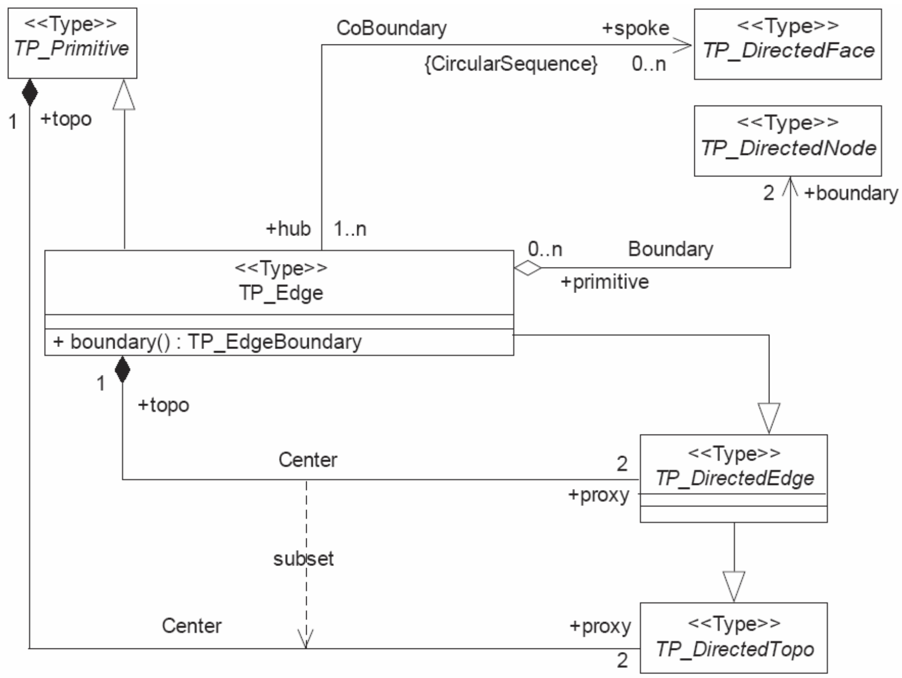


Figure 4 TP_Edge (ISO 19107)

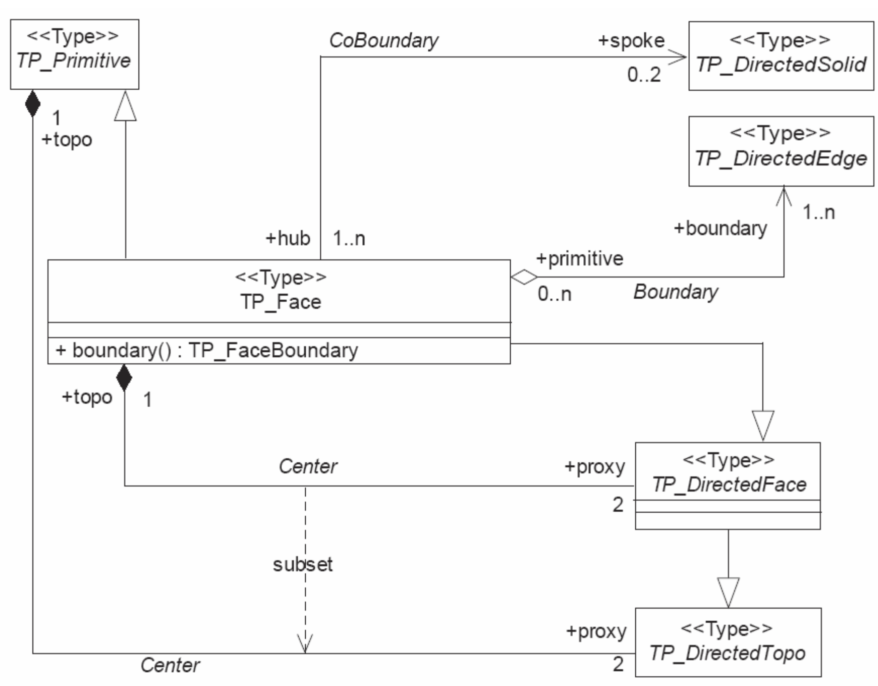


Figure 5 TP_Face (ISO 19107)

4. Mapping the cadastral model to ISO/TC 211

When creating a GML model for the cadastral data, a mapping from the cadastral model to the ISO19107 has to be created. Because there are differences in the models, this mapping is not always straightforward. The main conflichts between the models are that the Dutch Cadastre uses a winged edge structure thas does not completely fit in the ISO19107 topology model. The second problem is that the Cadastere stores a complete history of its parcels, this makes it hard to use the ISO topology model. In the next two sections, these problems and their possible solutions are handled in more detail.

4.1 Mapping the winged-edge model to ISO19107

Although it is possible to losslessly map to the ISO19107 model and back, this operation is quite expensive and in case GML is used to transfer topological data from one winged-edge system to another there will be considerable overhead. Below we propose a way to extend the ISO model in such a way that an easier mapping from winged-edge to ISO19107 is possible.

- The l_obj_id and r_obj_id for edges in the winged-edge model can be directly mapped to the CoBoundary relation in the ISO model. The co-boundaries are directed faces, one with positive orientation and one with negative orientation. These two orientation can be used to distinguish between ‘r_obj_id’ an ‘l_obj_id’.
- The face to first edge reference in the winged-edge model can be mapped to boundary relation in the ISO model
- Only the relationships between the edges (ll_obj_id, lr_obj_id,fl_obj_id,fr_obj_id) cannot be mapped to any relationship in the ISO model. In order to also preserve this information in the ISO Model we need to extend it. With the minimal extension of adding the EdgeNext relationship to a DirectedEdge all four elements can be uniquely mapped. (See Figure 6).

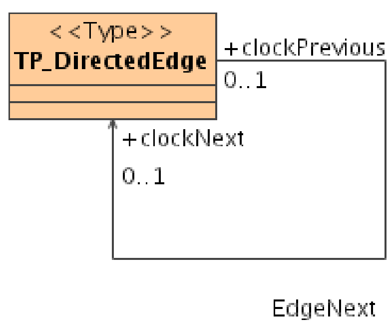


Figure 6: Extra relation to integrate winged with with the ISO model

4.2 Mapping the temporal topology to GML

The cadastral topological model is an integrated topological/temporal model that has a valid topology for any moment in time. This is implemented by adding a time interval to all topological objects (edges and faces) in which the object is valid. At any instance in time the collection of edges and faces that are valid at a given time form a complete instance of a non-temporal topology.

When mapping this model to the non-temporal model of GML, the complete cadastral map will be modelled as one ISO TM_Complex. In such a complex, the validity attribute will be just another temporal attribute and all the edges (regardless of time) will be valid at the same time. Parcel boundaries that cross, but that have disjoint valid time will result in intersecting edges in the Compex. According to ISO19107 only geometric complexes that consist of mutually disjoint geometric primitives will generate a topological complex without error. This means that modelling the cadastral model as a complex is not allowed. However, it is unlikely that any GML validating software will reject a GML file with overlapping edges. An alternative is to model the cadastral set just as a collection of vertices, edges and faces without grouping them together in a complex. Although ISO19107 does not allow this, GML does not require every topological primitive to be in a complex.

Unique object identifiers

Because in the cadastral temporal model there can be more than one instance of an object, the object idenfier (oid) of an object is not unique. In a GML document the id of an object has to be globally unique. The cadastral oids are only unique in combination with the tmin attribute. So to create unique identifiers in GML one can use combination of oid and tmax. However, topological references in the cadastral database, only use to the oid and not to tmin. In (Oosterom, 1997) it is explained why. As a result the GML xlink mechanism cannot be used to model the topological references. This makes it impossible to model the cadastral topological model. Another solution might be to exchange the cadastral data in multiple GML documents, in such a way that each document contains only one instance of each object.

5. Conclusions and Suggestions

With the current topological model of GML there is no 1:1 mapping possible between the cadastral topological/temporal model and the GML topology. This is mainly due to the fact that the ISO schema is a spatial schema, and not a spatio-temporal schema. In the current situation the best solution for the cadastre will be to use a model that is as close as possible to the GML topology but that differs in a few points. In the long term, an integrated spatio-temporal model in ISO19107 is the preferred solution. But before that time the following suggestion (that does not break current schema's in the ISO model) might make it more easy to map the cadastral model to GML:

- Add an optional EdgeNext relationship to the DirectedEdge of ISO.
- Add optional temporal attributes to the topology model and relax the constraint that the edges in a TopoComplex should be non-overlapping. An alternative would be to create a TempoTopoComlex.

References

Baumgart, B.G. (1975): A polyhedron representation for computer vision. *National Computer*

Conference, pages 589-596, 1975.

ISO/TC 211/WG 2 (2003) ISO/CD 19107, Geographic information — Spatial schema

ISO/TC 211/WG 4 (2004) ISO/CD 19136 Geographic information — Geography Markup Language, 2004

Oosterom, P.J.M. van and C.H.J. Lemmen (2001): Spatial data management on a very large cadastral database, *Computers, Environments and Urban Systems (CEUS). Volume 25, number 4-5, 2001.*

Oosterom, Peter van (1997). Maintaining Consistent Topology including Historical Data in a Large Spatial Database. *Proceedings Auto-Carto 13, Seattle WA, 8-10 April 1997, pages 327-336.*

Biography

Wilko Quak got his M. Sc. degree in computer science from Utrecht University in 1992. After that he worked as a PhD student at the University of Amsterdam in performance of spatial databases. Currently he is employed as an assistant researcher at the GIS Technology section of the OTB Research Institute for Housing, Urban and Mobility Studies, Delft University.