Some investigations on encoding of complex geographic information

Dr. Gerhard Joos
Prof. Dr. Wolfgang Reinhardt
M. Sc. Shi Wei

AGIS GIS-lab
University of the Federal Armed Forces Munich
D-85577 Neubiberg
Germany
Fax: +49 89 6004 3906
Wolfgang.Reinhardt@unibw-muenchen.de
Gerhard.Joos@unibw-muenchen.de
wie.shi@unibw-muenchen.de

key words: GIS, data model, data transfer, encoding, XML, GML

Abstract. As XML is becoming next generation data format for exchange and web publication, GML (Geography Markup Language [3]) is expected to be the next spatial data interchange format. The paper depicts the way of mapping from the hierarchically structured spatial data (like in MilGeo-data) onto the GML, publishing the GML documents from XML database on the Web.

1. Introduction

Besides online access via distributed computing platforms encoding is still a common way to enable the transport of geographic information (hopefully) without any loss. By browsing encoded data streams on the fly transfer mechanisms can also be used to serve GIS data on demand. Since in most transfer procedures the encoded data only contains the features themselves, this requires that the receiving GIS does already know the underlying model. Therefore it is desirable to encode the data schema as well. There are national, international, and vendor activities to utilize the encoding of geographic information. Many of them are using the Extensible Markup Language (XML) and its technologies. XML is the universal format for structured documents and data on the Web. To give two example of XML-based GIS transfer standards: ISO/TC 211 is going to release ISO 19118 “Geographic information - Encoding” as international standard and the Open GIS Consortium (OGC) has recently published the XML-based Geographic Markup Language (GML) as a OpenGIS® Recommendation. GML describes how to encode geographic information with simple geometry. To satisfy different needs of users GML 1.0 provides three profiles employing different XML technologies and GML 2.0 provides two abstract schema: feature schema and geometry schema.

The advantages of XML-based exchange of data are, that geospatial data can be described completely within the exchange mechanism; No a priori knowledge of the referring application schema is required; The data can also easily be published on the internet; XML separates content information from presentation; It's easy to link to non-spatial data; users needs no specific and expensive software; furthermore, interchange using GML can guarantee that application schemas conform to the OGC Feature Model.

In this paper we consider the whole process of data transfer including model description, schema mapping (like feature classes, properties, and relations), data transfer, and publication on the web. Different approaches are evaluated by a practical example of MilGeo-data.

2. Describing MilGeo data Model in hierarchy structure

Complexity of geo-spatial data can be expressed with several aspects. The geometry of features can be complex, which means, that there is a non-linear interpolation between vertices. Also the structure of modelling geo-spatial data can be complex. This is considered in our context where features are modelled in a hierarchical structure.

German Bundeswehr Geographic Office (MilGeo) is building and maintaining a spatial database called MilGeo-data which are required:

- as hierarchically structured data models,
- containing relationships between features,
- the topology is not considered.

The application schema for MilGeo-data uses a hierarchical object-based structure. Features belong to one of three types of hierarchical modeling, namely base feature, several levels of composite feature or theme.

A feature of type theme is at the top (root) of the feature hierarchy while base features are features at the bottom (leaf) of the hierarchy that describe concrete geographic objects. All features belong to a certain feature class. Base features have different properties: the geometric property is modeled with the geometric primitives point, curve, or surface. The characteristic attributes of the features are described by basic data types and coded lists.

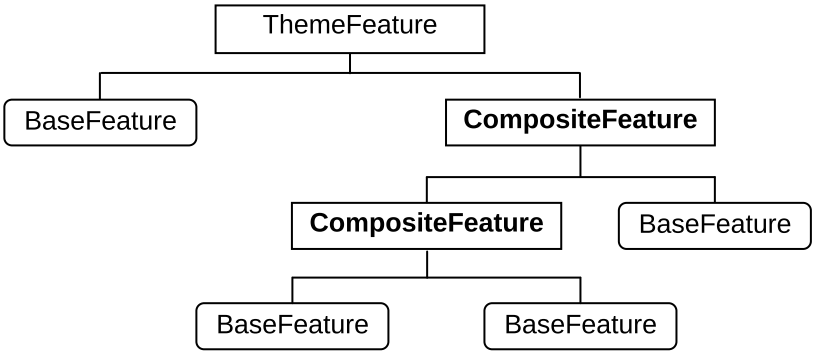

Figure1: The hierarchy of composite feature

Composite features are intermediate features (trunk) that are defined to clarify the conceptual relationship between theme and base features or to provide additional conceptual groupings of features. The hierarchy of composite feature is shown by figure 1.

3. Mapping from MilGeo-data to GML

The mapping contains three aspects: attributes mapping, geometry mapping, and relationship mapping. The attribute mapping corresponds to a mapping of standard data types and it can therefore easily be implemented. Additionally a mapping of coded lists for attribute values to enumerations in the attribute value domain has to be introduced. Since MilGeo-data is based on a geometry model which can be called simple, i.e. there is only linear interpolation applied and the features are not allowed to intersect themselves, the geometry mapping can be performed straight forward. The mapping for relationships between features is restricted to relationships between the hierarchical levels.

Though XML naturally supports an intrinsic hierarchical structure, the destination GML should satisfy two preconditions in order to realize the mapping from composite features to GML:

- support for collection type: GML 2.0 based on so-called simple features whose geometry model allows that geometry is a collection of other geometries of homogeneous or heterogeneous geometry type;
- support for the nesting mechanism: composite feature can be mapped to GML FeatureCollection type. The hierarchy depth, however, cannot be forecasted. Hence, GML should support nesting mechanism which can be implemented as following.

```
<element name="_FeatureCollection" abstract="true"
```

```
           substitutionGroup="gml:_Feature">
    <complexType abstract="true">
      <complexContent>
        <sequence>
          <element ref="gml:featureMember"
                   minOccurs="0" maxOccurs="unbounded"/>
        </sequence>
      </complexContent>
    </complexType>
  </element>
  <complexType name="FeatureMemberType">
     <complexContent>
        <sequence>
           <element ref="gml:_Feature" minOccurs="0"/>
        </sequence>
     </complexContent>
  </complexType>
```

In this fragment, element FeatureCollection contains element Feature; the element Feature can be replaced everywhere with the element FeatureCollection using substitution mechanism of XML. So nested hierarchy is realized with GML.

The two base GML Schemas provide meta-schemas, from which an application schema can be constructed in standard ways. Using GML collection type and nesting mechanism, theme feature and composite feature can be described as a collection of GML features, base feature replaces simply feature element of GML. The application schema is mapped to GML.

4. Publishing GML Document on the Web

GML has been designed with the concept of the separation of presentation and content. It leads to a fully self-describing stream of data, that can be stored or shipped through a distributed environment. For the online processing, GML can be interpreted on the fly, e.g. by a simple conversion within a standard web browser in a browser readable format.

By this processing, user submits requests to XML Database using XQuery through HTTP. As a result, XDB returns a GML object. This GML object is transformed to a simple graphic and a property representation by splitting its content into Scalable Vector Graphic (SVG) which describes geometry of features and Hyper Text Markup Language (HTML) which describes attributes of features. The geometry and attributes are connected through XLink/XPointer.

6. Conclusion and Prospective

Our experiments indicate that it is feasible and suitable to map hierarchically structured data to GML. MilGeo requires data exchange between multiple users on multiple levels:within organisations, between different users and among nations. This can be realized through the characteristic of XML that separates content from presentation.

Because of a large amount of geospatial data, the efficiency problem using XML spatial data needs to be solved. Since XML can be next generation format for databases (XDB), GML documents are to be administered by database technology. Our experimental (tested with Tamino) results indicate that the efficiency of managing XML and GML documents with XDB are improved notably.

Due to the combination of GML and G-XML, next versions of GML will express explicit topology and describe spatial object more completely. Persistent Identifiers for features are important for XLink/XPointer, and already in GML 2.0 as optional. In addition next generation web browser will completely support XML Schema, XLink/XPointer and XSLT/CSS so that GML can be displayed directly in the near future.

References

[1] ISO TC211, 1999. Geographic information http://www.statkart.no/isotc211/wg4/welcome.html

[2] OGC, 1999. OpenGIS Simple Features Specification, http://www.opengis.org/techno/specs.htm

[3] OGC, 2001. Geography Markup Language 2.0, OGC Recommendation Paper, http://www.opengis.org/

[4] OGC, 2000, Web Feature Server / Web Map Server, http://www.intl-interfaces.com/wmt2/

[5] W3C, 2000. XML Schema Part1: Structures. W3C Candidate Recommendation, http://www.w3.org/TR/xmlschema-1/

[6] W3C, 2000. XML Schema Part2: Datatypes. W3C Candidate Recommendation, http://www.w3.org/TR/xmlschema-2/

[7] W3C, 1999. XSL Transformations (XSLT), Version 1.0, W3C Recommendation, http://www.w3.org/TR/xslt/

[8] W3C, 2000. XML Linking Language (XLink), Version 1.0, W3C Proposed Recommendation, http://www.w3.org/TR/xlink/

[9] W3C, 1999. XML Pointer Language (XPointer), Version 1.0, W3C Candidate Recommendation, http://www.w3.org/TR/WD-xptr

[10] W3C, 1999. XML Path Language (XPath), Version 1.0, W3C Recommendation, http://www.w3.org/TR/xpath/

[11] W3C, 2000. XQuery, W3C Working Draft, http://www.w3.org/TR/xquery/

[12] W3C, 1999. Namespaces in XML, W3C Recommendation, http://www.w3.org/TR/1999/REC-xml-names-19990114/

[13] W3C, 2000. Scalable Vector Graphics (SVG) 1.0 Specification. W3C Candidate Recommendation, http://www.w3.org/TR/2000/CR-SVG-20001102/index.html/

[14] W3C, 2001. W3C, 2001. XML Fragment Interchange. W3C Candidate Recommendation, http://www.w3.org/TR/xml-fragment.html

[15] Software AG, 2000. Schema Definition, Version 2.2.1

[16] W3C, 1997, XML representation of a relational database, http://www.w3.org/XML/RDB.html

[17] XML:DB, 2000, the Standards Initiative for XML Databases, http://www.xmldb.org/

[18] Masatoshi Arikawa, Koichi Kubota, 2000, A Standard XML Based Protocol for Spatial Data Exchange, http://gisclh.dpc.or.jp/gxml/contents-e/

[19] Kelly Chan, 1999, DIGEST-A Primer for the International GIS Standard.