Requirements and Research Issues in Geographic Data Modeling

Anders Friis-Christensen
Comp. Sci. Dept., Aalborg Univ.
Fredrik Bajers Vej 7E
Aalborg, Denmark
afc@cs.auc.dk

Nectaria Tryfona Christian S. Jensen
National Survey and Cadastre
Rentemestervej 8
Copenhagen, Denmark

Comp. Sci. Dept., Aalborg Univ.
Fredrik Bajers Vej 7E
Aalborg, Denmark
{tryfona,csj}@cs.auc.dk

ABSTRACT

It is well-documented in the literature that geographic data have special characteristics that make the use of extensions to standard modeling languages and techniques, such as the Unified Modeling Language, attractive. Based on a real-world application from the Danish National Survey and Cadastre, this paper presents requirements to geographic data modeling notations. Existing notations are then evaluated against the requirements, and a case study is carried out. The result is an identification of pertinent aspects of geographic data modeling-including roles of geographic objects, constraints on objects, and quality of data-that are not handled satisfactorily by existing proposals.

Keywords

Requirements analysis, GIS, conceptual data modeling, geographic data.

1. INTRODUCTION

Geographic data and Geographic Information Systems (GIS) are being used increasingly. Along with the growing use of geographic data, it also becomes necessary to exchange and combine heterogeneous data, which refers to the same phenomena, between different authorities and systems. One example is the combination of data about buildings from the local authority with data from the National Register for buildings. Typically, such data are collected and organized differently based on non-standard data models, which are application-dependent. This makes the exchange of data a difficult or plain impossible task. Developing comprehensive, *standard* data models will create a basis for making the exchange of data an easier task. This calls for research in geographic data modeling.

The different kinds of geographic data have many common properties, such as location and extent, which are not supported in the standard modeling languages, e.g., the Unified Modeling Language (UML) or the Entity Relationship (ER) model.

This paper aims to identify and examine the common properties of geographic data, and to evaluate pertinent existing models with respect to their ability to accommodate these properties. First, a requirements analysis, based on interviews with users at the Danish National Survey and Cadastre (KMS), is described. A case study then explores how the requirements are supported by the modeling notations considered. This study reveals that the notations fall short in meeting all the requirements. To the authors' best knowledge, this is the first effort that evaluates existing work with respect to real users' requirements.

The paper is organized as follows. Section 2 presents requirements that aim to explicate the properties of geographic data that should be considered when developing geographic data modeling notations. Section 3 presents key existing techniques and notations for the modeling of geographic data. Section 4 describes the case study and the results derived from this. Section 5 summarizes and points to research directions.

GIS'01, November 9-10, 2001, Atlanta, Georgia, USA..

2. REQUIREMENTS ANALYSIS

The following requirements to modeling notations for geographic data derive from interviews with employees at KMS. The requirements are divided into three types: aspects related to objects, aspects concerning the relation between objects, and aspects related to the events of an object. The events of an object are initiated by outer circumstances that affect the object.

2.1 Geographic Objects

The aspects that need to be considered regarding objects in geographic data models are:

Location and extent. The spatial location and extent are defined by x, y, and z coordinates in a specific reference system and represented by either points, lines or polygons.

Temporal extent. It is necessary to keep record of the existence and change in time of an object. We need to represent valid, existence, and transaction time.

Complex spatial extents. It should be possible to associate complex spatial extents with objects. Thus, zero-, one-, two-, and three-dimensional extents consisting of, e.g., points (and multi-points), lines (and multi-lines), polygons (and multi-polygons), and rasters, should be possible. This enables objects to consist of more elements (e.g., a building may consist of two polygons).

Thematic values. An object has several attributes that define its properties, e.g., name, or ownership.

Fuzzy objects. The representation of a fuzzy object is related to its location and thematic values. When supporting fuzzy objects, it should be possible to model objects that have a location and an extent that to a certain degree may be referred to by the coordinates (x, y, z) and have thematic values that to a certain degree can be associated with a given class.

Entity vs. field based data. In the entity-based approach, the real world consists of fully definable disjunctive entities, such as roads and buildings. In the field-based approach, the real world consist of attributes assumed to vary over space as a continuous function [5]. This would be the preferred approach when representing phenomena such as heights and depths. It is important to be able to distinguish between entity and field based data because they have fundamentally different characteristics. Several of the preceding descriptions of properties only refer to entity based data.

Generalization. The generalization of an object relates to aspects such as scale and purpose. A single object can be derived from several objects at a larger scale, and information about derived objects or objects to be aggregated must be accessible. Furthermore, information about how the spatial extent depends on the scale is needed because it is used when visualizing an object.

Constraints. It should be possible to attach constraints to an object. An example can be that values of a certain attribute have to be limited to an interval, e.g., the z coordinate of object class low *buildings* is limited to l-5 meters, or the geometry of object class *windmill* should be points. A categorization of different types of constraints can be made, which would facilitate modeling process. The constraints are related to data quality: if they are not satisfied, this will negatively affect data quality.

Roles. An aspect, which is closely related to the representation, is the role of an object. A geographic object can be defined differently depending on the universe of discourse; in other words, the role or an object is application dependent. An example can be a coastline, which is defined by its representation on an aerial photo in a topographic map, while its definition in a nautical chart is based on the water level. So, different conceptual models describe different roles of an object.

Object id. A unique object id is necessary for data exchange purposes. When different institutions and authorities need to exchange data, unique object id's are a great help. A possible solution to this is to establish an authoritative object id supplier.

Data quality. The quality describes how "good' the data are. Such information is important when assessing the credibility of data and should be mandatory. The notion of data quality has some fundamental components that can be used in defining specific, application-dependent meanings of data quality, which is also called the "fitness for use." The fundamental components have to be measured against the application and the specifications, and the result is the "fitness for use."

The quality can be divided into quantitative and qualitative quality, where *quantitative quality* consists of components that are directly measurable. An example is the spatial accuracy which describes how accurate the position of an object is in comparison to a given reference object.

In contrast, *qualitative quality* consists of information that cannot be measured. Typically, this information relates to the data set as such and not to specific objects. An example, which also relates to the objects in a data set, is the origin, which describes how objects are created. Whether this information should be attached to an object itself, each of the geometric elements of an object, or both depends on several aspects. If a part of an object can be updated without updating the whole object, origin information should at least be attached to the geometry elements.

2.2 Relations Among Geographic Objects

The aspects to be considered about relations among objects in geographic models are:

Topological relationships. Topological relationships refer to the relations and connections among objects. Examples of binary topological relationships include contains, overlaps, passes trough, and touches [22]. In three-dimensional space, additional relations are possible, e.g., lies under, is on top of, and passes under.

Metric relationships. Metric relationships involve distance and depend on the absolute positions of objects relative to a given reference system [22].

Semantic relationships. These are relationships among objects, relevant at the conceptual level, that are neither topological nor metric. It could for example be that all land parcels have road-access, either directly or through an allowed access via a neighbor parcel. When representing the objects in less expressive models, some semantic relationship may be represented as a topological relationship.

Part-of relationships. An object can consist of other objects. An example is a county consisting of municipalities.

Relationship constraints. The relationship constraints are very important and are highly dependent on the type of relation between objects. A constraint can be topological: for example, it is not allowed to place a building in a lake. It is also important to be able to state exceptions. In the above example, it is may be allowed to build a house on pillars in a lake.

2.3 Events of Geographic Objects

Two aspects of the events that affect objects are important.

Visualization. The visualization of objects is essential. Objects can be visualized in different scales, e.g., a building can be visualized as a polygon in one scale and as a point in another (generalization). The placement of an object on a map generally relies on the neighbor objects. For example, a hedge can be "misplaced" if it is located next to a road that is purposefully displaced. This means that the visualization of an object is based on the properties and relations to other objects.

Change. As objects can change, information stating the "outer" circumstances that may cause changes is important, as is information about who is responsible for a change.

2.4 Interactions Among Aspects

Several of the aspects described in this section interact. For example, relationships between objects provide constraints. Also the notion of quality is related to constraints. When a constraint is not satisfied by an object, this affects the object's quality. In addition, object roles are related to their visualization. Typically, two different roles of an object have different visualizations because the purpose of a role affects the visualization. Because of these interactions, when we consider one aspect, we must be prepared to also take into account related aspects.

3. EXISTING MODELING TECHNIQUES

We proceed to survey existing proposals for the modeling of geographic data. Some of these proposals will be used in the case study in Section 4.

Two general modeling approaches to geographic data modeling have been adopted: the Entity Relationship approach and the object-oriented approach [22]. So far, the Entity Relationship approach has been the predominant; but as more application systems are becoming object-oriented (e.g., Arclnfo 8 and Oracle 8 support some notions of objects), this may change.

We more specifically proceed to review six of the most recent proposals for notations related to geographic and spatiotemporal data modeling:

STER The development of a spatiotemporal extension to the ER model, and the use of abstractions and modeling units in spatiotemporal modeling [20, 21].

Perceptory The development of a spatiotemporal plugin (within a Case Tool) for UML [1].

Ext. UML The development of a spatiotemporal extension to UML [17].

Patterns The use of design patterns in geographic data modeling using Object Modeling Technique (OMT) [10].

GeoFrame The development of analysis patterns for geographic databases, based on UML [7].

OMT-G Spatial data integrity constraints using an extension to OMT [3].

The different notations can be characterized by their approach. Table 1 shows if a notation uses the object-oriented approach or the Entity Relationship approach and whether it is an extension to existing languages or models.

Table 1: Characteristics of Notations

Model/Technique	ER	OO	Extension
STER	X		X
Perceptory		X	X
Ext. UML		X	X
Patterns		X	
GeoFrame		X	X
OMT-G		X	X

Tryfona and Jensen [20] develop an icon-based extension to the ER model, STER, that includes spatiotemporal entities, attributes, and relationships. STER offers support for the spatial elements, point, line, and regions, and for the temporal elements, existence time (for objects), valid time (for attributes and relationships) and transaction time (for all constructs). This notation was subsequently [21] used in the development of modeling units consisting of modeling parts that describe common spatiotemporal characteristics. Modeling units are similar to design patterns in terms of giving solutions to common design problems; however, modeling units concern data modeling whereas design patterns are used in software development in a broader sense [21]. No formal notation for constraints is considered other than those constraints provided by the spatial and temporal relationships, e.g., topological and metric relationships.

Bédard [1] develops a spatial and temporal plug-in for UML, which is implemented in a CASE tool called Perceptory [18]. Perceptory is an extension to UML based on stereotypes (icons) and adds support for spatiotemporal properties that are aligned with the ISO-standards for geographic information [4]: zero-, one- and two-dimensional spatial properties and zero- and one-dimensional temporal properties (instant, existence time (object), and valid time (attribute)). The spatiotemporal properties can be defined at class and attribute level. UML makes it possible to define constraints on objects and associations in textual form, e.g., using Object Constraint Language (OCL) or just natural language. But this is not formalized, and there are no predefined spatiotemporal operators, e.g., topological or metric, which makes the definition of constraints limited. Furthermore, Perceptory has no support for field-based data or for three-dimensional objects, and there is no support for transaction time.

Price et al. [17] develop a spatiotemporal extension to UML (Ext. UML) by adding a set of constructs that can be applied to classes, attributes, and associations. The extension is based on spatiotemporal symbols, and a specification compartment is used to describe the semantics of spatiotemporal properties.

The spatial properties consist of spatial extents, and the specification compartment supports the specification of dimension, interpolation type, and the model used (i.e., entity or field). The temporal properties are specified in more detail. For these, the specification compartment makes it possible to specify the time dimension (existence, valid, and transaction time), the interpolation method, the model (regular or irregular), and the unit of time used (e.g., instant or interval). Thematic properties represent thematic data. All types of properties can be used as nested properties, e.g., spatially dependent thematic property. Furthermore, it is possible to group attributes with the same spatiotemporal properties and to define, e.g., attributes that depend on the existence of an object. These properties are not considered in the other works described above.

Ext. UML provides a rich syntax covering the semantics of spatiotemporal data. It does not investigate constraints on associations or classes.

Gordillo et al. [10, 11] introduce design patterns as a conceptual tool for designing GIS applications using OMT. To address the problems of static models, a design pattern, termed ***decorator*** [9], is used to attach additional responsibility (e.g., location) to an object dynamically. Furthermore, new design patterns are developed to handle GIS domain problems, including roles of objects, reference systems, and the appearance of objects. There are several advantages in the use of design patterns for geographic modeling due to the dynamical change of properties of objects, which are essential to a GIS. A potential exists for further developing design patterns that handles the requirements of geographic data modeling.

Filho and Iochpe [7] develop analysis patterns for geographic databases, called GeoFrame, that rely on an extension to UML. The difference between design patterns and analysis patterns is that design patterns focus on design solutions and software implementations, whereas analysis patterns focus on the reuse of modeling solutions and use terms from the application domain [7]. Analysis patterns can also be characterized as modeling units, which have been defined earlier. GeoFrame adds support for a comprehensive list of spatial constructs for both entity-based data and field-based data. It adds no support for temporal properties or spatiotemporal associations. An interesting aspect of this work is the use of analysis patterns, which provide a useful solution in modeling situations where reuse is possible.

Borges et al. [3] propose OMT-G for defining spatial data integrity constraints. OMT-G is a spatial extension to OMT supporting both entity-based and field-based data. To ensure integrity in spatial databases, a number of rules (e.g., aggregation and connectivity) are defined. However, only spatial (as opposed to temporal) properties and constraints can be modeled using OMT-G.

In summary, it is noticed that all, except Borges et al. [3], offer notations for modeling spatiotemporal properties. Only Tryfona and Jensen [20, 21] and Price et al. [17] consider transaction time. Several papers consider constraints on classes and associations with spatiotemporal properties, but only Borges et al. [3] cover the issue thoroughly, albeit only for spatial constraints. Bédard [1] presents briefly a prototype for specifying constraints between spatial objects, but does not provide details. Finally, only Gordillo et al. [10, 11] consider other properties of geographic data than the spatiotemporal. Roles and appearance/visualization of objects are properties that are relevant to consider when using GIS, and patterns can be used to handle these issues [10, 11].

In the case study in Section 4, an object-oriented approach using UML is used. UML is chosen as the modeling language because it has become widely used in both academia and industry. It is now a standard defined by the Open Management Group (OMG) and is also proposed as an ISO standard (DIS19501) [13]. UML is furthermore used in the standardization work within Geographic Information done by ISO (ISO TC211, http://www.statkart.no/isotc211).

The notation for UML class diagrams used in this paper is depicted in Figure 1. For further readings about UML see, e.g., Booch et al. [2] or Fowler [8].

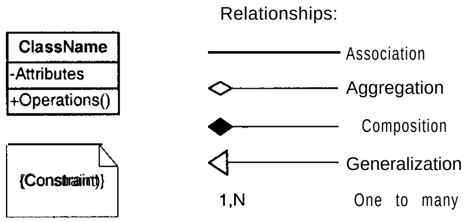

Figure 1: The UML Notation Used in This Paper

4. CASE STUDY

The following section explores the requirements described. We first classify the requirements with respect to their overall properties and characteristics.

- *Spatiotemporal* properties include the spatial and temporal extents, thematic values, object id, and entity- versus field-based data.
- *Roles* are as described in the requirements.
- *Associations* include topological, metric, semantic, and part-of relationships.
- *Constraints* include constraints on objects and relationships.
- *Data quality* includes both qualitative and quantitative quality.

The case study focuses on the above-presented requirements, but considers also interoperability, as this is related to requirement such as roles of objects. Issues from the requirements not considered here include fuzzy objects, generalization, and the object events. Consideration of these requirements is deferred because it seems appropriate to consider the remaining requirements initially.

The case study is based mainly on extensions to UML [1, 17] and design patterns [10, 11], but ideas from other works are also considered [3, 7, 20]. The purpose of the case study is to identify aspects of geographic data modeling not supported by existing work.

4.1 The Example

The examples are based on objects maintained by the Danish National Survey and Cadastre and consist of the following: *road, lot,* and *building.* These objects are chosen to illustrate some of the issues when modeling the requirements from above. Interoperability is an issue because the objects appear in two of the main databases at KMS (the topographic and cadastral databases), and *road* even exists in both as two different object classes and with no association other than that their references to the same real-world objects. To ease update and maintenance, it should be possible to handle, *e.g., road* as one object with different roles and visualizations.

4.2 Spatiotemporal Properties

As mentioned earlier, existing notations for modeling geographic data include several means of modeling spatiotemporal properties. An example of modeling the spatiotemporal properties of a *road* using Perceptory [1, 18] is seen in Figure 2, where also the basic icons are explained.

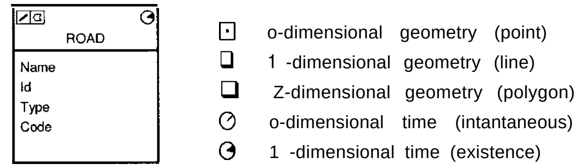

Figure 2: A Spatiotemporal *Road* and the Icons in Perceptory

These basic icons can be combined to obtain more advanced geometries. The combination of two (or more) icons separated by a line, as in the road object in Figure 2, means that the object (road) is represented of one or the other geometry (here, as either line segments or polygons). Another example is that if two (or more) icons are combined without a separating line, this indicates a complex geometry, i.e., multiple elements constitutes the geometry of an object.

As indicated in the figure, Perceptory supports zero-, one- and two-dimensional spatial extents, as well as only instantaneous and durable temporal elements. There is no immediate support for transaction time, which then must be added if needed; and there is no support for field-based data. The spatial data types can be specified in Perceptory, e.g., as Oracle Spatial geometry types 1 (point), 2 (line segment), or 3 (polygon), or as ISO geometries such as point and multi-point. Likewise, temporal data type can be specified, such as date or the ISO specified temporal data types, instant and period.

Support for transaction time is found in Ext. UML [17], which is more flexible than Perceptory. Ext. UML also it gives the possibility to specify whether data are entity-based or field-based. An example is seen in Figure 3.

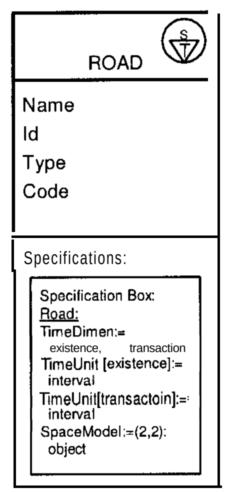

Figure 3: Ext. UML Using Road as an Example

Here, *road* is a temporally dependent spatial object-the spatial extent of road across time is recorded. Further, *road* is specified as a two-dimensional spatial object. In Ext. UML, it is not possible to define that an object is represented as, e.g., multi-points. This is possible in Perceptory, and in order to meet the requirements of the ISO-standards, this is necessary. Furthermore, Ext. UML seems less intuitive than Perceptory.

4.3 Object Roles

The object ***road*** has two different roles, and these are stored in two databases at KMS, the topographic database and the cadastral database. It is a well-known problem that two or more different objects in the database actually refer to the same real-world object. One means of modeling the object ***road*** is to use subclasses, as indicated in Figure 4.

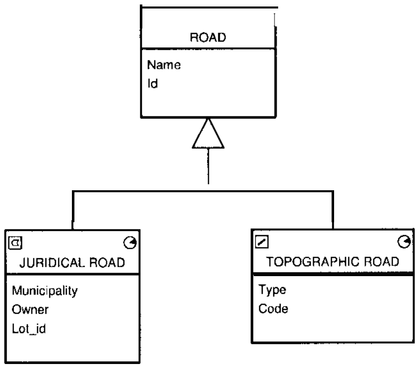

Figure 4: *Road* and Its Two Subclasses

However, there is a need to be able to include several types of the same object instance, which this solution does not allow. Furthermore it is important, to enable interoperability, that each type of ***road*** instance has the same global unique identifier-this makes it possible to relate information from both types of the same object.

One solution is to use the concept of roles, which is well-known in object orientation [16, 19]. Specifically, Gordillo et al. [10, 11] uses OMT and the role concept to model multiple roles of geographic objects.

In our example, a role design pattern can be added to the road object, which then can have a juridical and a topographic role. The role design pattern can be modeled as shown in Figure 5.

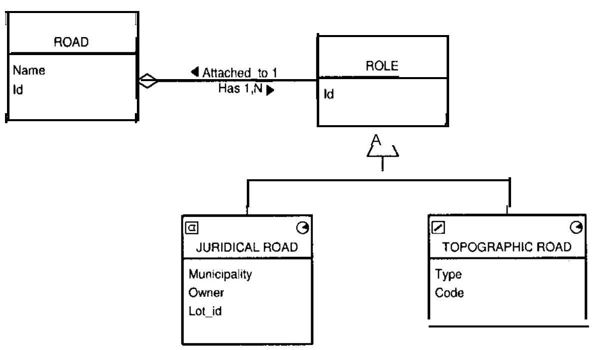

Figure 5: *Road* and Its Two Roles

The advantage of this solution is that an object can have associated more than one role, thereby enabling different information to be attached to the object. Like design patterns, modeling units [21] or analysis patterns [7] can also be used.

4.4 Associations

Associations represent relationships between objects, and different types exist. This section only considers simple associations, not aggregations and compositions.

To illustrate associations, the object owner is included, and Ext. UML is used in the example because Perceptory has no extension for spatiotemporal associations. A lot is owned by an owner as seen in Figure 6, where *lot* is temporally dependent on owner; an

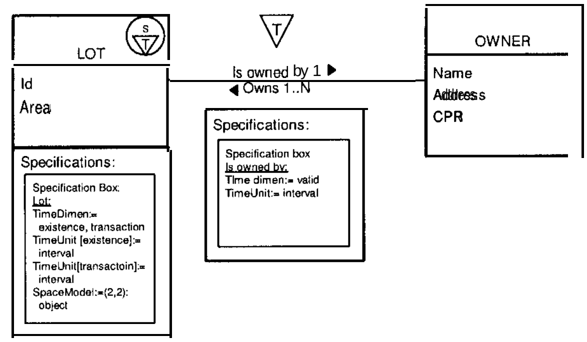

Figure 6: The Association Between Lot and Owner

owner can sell a lot, but it should be possible to go back in time to see information about previous owners. In Ext. UML, it is possible to define spatiotemporal dependencies and other types of relationships, e.g., the part-of relationship, between objects. However, there is no support for specifying the constraints considered in the next section.

4.5 Constraints

Constraints are important when modeling the world to be described. Most model elements inherently imply constraints, especially associations (e.g., the cardinality) and attributes (e.g., the type), however they cannot express all constraints of interest. Consider the example of associations between geographic objects. These associated objects are spatially related, and topological constraints are needed to describe the associations thoroughly. As stated earlier, there is the possibility of using OCL in UML when modeling constraints, and these can be attached to classes, attributes, operations, and associations between classes.

Ditt et al. [6] classify constraints according to the data involved in the constraints:

1. Constraints referring to an attribute of a single object.
2. Constraints referring to at least two attributes of a single object (a single tuple in a relational database).
3. Constraints referring to all objects of a single class.
4. Constraints referring to an object and its associated objects of various classes.

Constraints referring to operations of objects are not mentioned in this classification, but also need to be considered. Only Borges et al. [3] investigate constraints that are not inherent in the model elements. They focus on items 3 and 4 above and define constraints for certain spatial properties and associations. An example is the binary association ***containment*** which is defined as: (i) the geometry of the containing object must encompass the geometry of the contained objects, and (ii) the boundaries of no contained object can exceed the boundaries of the containing object [3]. Borges et al. do not consider temporal types such as synchronization. Furthermore, the solution is based on predefined textual rules and there is no formal language in which new constraints can be specified.

When modeling geographic data, all types of constraints have to be considered for the geographic properties. This is not considered in any of the related work. We consider next a brief example of modeling data including constraints using a formal language, like OCL. In this example, two classes, *building* and ***lot,*** their association, and their constraints are described.

However the problem is that there are no spatiotemporal operators defined within OCL, e.g., is-inside. In Figure 7, the association between building and lot is illustrated using OCL notation for constraints and Perceptory notation for the spatiotemporal properties.

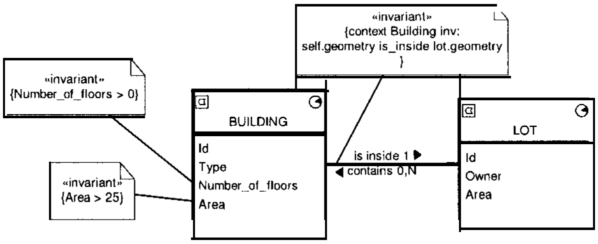

Figure 7: Association and Constraints Between Building and Lot

The constraint `is-inside` attached to the association is not valid within standard OCL and has to be defined as a user-defined extension to OCL, as do other relationships such as touches, overlaps, intersect, and distance relationships.

The constraint on the attribute `Number-of-floors` demands that a building has at least one floor. `Area` requires that the area of a building should be at least $25m^2$. These constraints can be termed thematic constraints, and they use a standard operator (>) supported by OCL. However, it is not logical that area is a thematic property of the object; instead, it should be considered as a property of the object's geometry and thus use a constraint like: `context Building inv: self.geometry.area > 25.`

A research direction is to define these geographic types and operations in a user-defined extension within OCL, that then can be used when modeling geographic data.

4.6 Quality

Quality aspects are important for all types of data, including geographic, and it is important to have access to this information, as data always approximate the real world. The aspects of quality in geographic data modeling are not investigated in previous work but, e.g., a systematic approach of using patterns or modeling units [7, 11,211 can be used.

An important aspect of quality is that it cannot be assessed without taking into account the application of the data. It implies two parts of quality assessment: the identification of fundamental components in quality and the assessment of quality taking the application into account. The inclusion of quality in geographic data modeling only concerns the first part. This is a difficult endeavor, due to the large number of involved aspects. In this paper we choose to define quality as consisting of the following fundamental components:

- Completeness (omission and commission). Are some objects missing from the data set or are there objects in the data set that are not found in "reality"?
- Thematic accuracy (quantitative and qualitative). Is the classification of an object correct (qualitative) and are the measured values correct (quantitative)?
- Temporal accuracy (measurement, consistency, and validity). Are the data valid, and are the time measurements consistent?
- Spatial accuracy (absolute and relative). Are some objects misplaced?
- Logical consistency (domain, topological, and format). Is the topology consistent (e.g., lines shall intersect in a node) and do the data obey the domains specified for them?
- Qualitative quality (lineage, origin, and usage). How are data created and what is the intended use?

Here, we consider those that are most relevant to the modeling process. For a thorough study see, e.g., Guptill and Morrison [12] or the standardization work done by IS0 within data quality in geographic information [14, 15].

The internal or the external approach can be used when assessing quality. The external approach checks the data against some kind reference, e.g., to see if an object in a given data set also exists in the reality. The internal approach is to check data for consistency using an internal test, e.g., to check if a topological rule is observed; for example, two lines have to intersect in a node. The quality components can be classified according to the assessment approach. For example, the external approach has to be used when assessing completeness and spatial accuracy, whereas logical consistency can be assessed using the internal approach.

The different quality components are related to different geographic model elements. For example:

- Completeness is related to object instances.
- Thematic accuracy is related to attributes.
- Spatial accuracy is related to the geometry properties of an object (likewise temporal accuracy is related to temporal properties).
- Logical consistency is related to attributes and associations.
- Qualitative quality is additional metadata.

So the quality components can be seen as additional properties to objects, attributes, and associations, but they are also related to constraints, e.g., consistency is related to a domain constraint on an attribute.

5. CONCLUSIONS AND RESEARCH DIRECTIONS

This section summarizes and briefly discusses the findings from the case study. Specifically, the strengths and weaknesses of the models incorporating patterns/modeling units or standard object-oriented concepts are considered, as they offer the most promising outsets for future work.

Table 2 characterizes STER [20], Extended spatiotemporal UML (Ext. UML) [17], Design patterns for geographic applications (Patterns) [10], GeoFrame [7], Perceptory [18] and OMT-G [3] according to their support for spatiotemporal properties, associations, roles, constraints, and quality. The category "constraints" indicates those constraints that are specified explicitly, not those that are implicit, e.g., in associations.

In the table, Yes indicates either complete fulfillment or that only small modifications are needed to fulfill a requirement. Next, Partial denotes partial fulfillment, meaning that certain elements are not possible to encompass in the model. Finally, No indicates that a requirement is not fulfilled.

None of the models satisfy all of the requirements considered, and additional research is needed to investigate aspects such as quality, roles, and constraints. All the object-oriented approaches can be used when modeling roles of geographic objects, but only Gordillo et al. [10, 11] consider the topic thoroughly. However several aspects about roles need to be investigated further. Such

Table 2: Fulfillment of Requirements

Model	Properties	Assoc.	Roles	Constraints	Quality
STER	Yes	Yes	No	No	No
Ext. UML	Yes	Yes	Partial	No	No
Patterns	Partial	No	Yes/Partial	No	No
GeoFrame	Partial	Partial	Partial	No	No
Perceptory	Yes/Partial	Partial	Partial	Partial	No
OMT-G	Partial	Partial	Partial	Partial	No

aspects include the questions of what defines a role (e.g., different semantics or purpose), how roles can be used in a distributed environment, and a specific notation for roles.

Most of the work involving constraints, except that by Borges et al. [3], is limited to the constraints implicit in associations and attributes. A development of an extension to UML that handles geographic data types and operators is needed because it is important to thoroughly specify objects to avoid inconsistency and ambiguity.

Most of the models address spatiotemporal properties and associations. Very different approaches have been taken, and not all models satisfy all requirements. For example, Perceptory has no formal specifications of spatiotemporal associations, as has Ext. UML.

It is important that there is a balance between the comprehensiveness and ease of use of a notation, Perceptory is very intuitive and easy to use, but, e.g., does not support some aspects of the temporal dimension. Ext. UML covers most aspects, but is less intuitive and more difficult to use. The challenge is to balance the two or to improve both simultaneously.

The study reported in this paper indicates that future research should be directed towards aspects such as object roles, constraints, and quality. Also areas not considered in the case study, including generalization and fuzzy objects, are important research directions. When considering interoperability of geographic data, it is an advantage to have standard modeling notations and implementations for all these aspects.

Acknowledgments

The authors wish to thank the employees at the National Survey and Cadastre for providing the information for the requirements analysis phase. This research was supported in part by the Wireless Information Management network, funded by the Nordic Academy for Advanced Study through grant 000389 and by a grant from the Nykredit corporation.

6. REFERENCES

[1] Y. Bédard. Visual Modelling of Spatial Databases: Towards Spatial PVL and UML. *Geomatica*, 53(2):169–186, 1999.

[2] G. Booch, J. Rumbaugh, and I. Jacobson. *The Unified Modeling Language User Guide*. Object Technology Series. Addison-Wesley, 1999.

[3] K. Borges, A. Laender, and C. Davis. Spatial Data Integrity Constraints in Object Oriented Geographic Data Modeling. In *Proceedings of ACM GIS*, Kansas City, USA, 1999.

[4] J. Brodeur, Y. Bédard, and M.-J. Proulx. Modelling Geospatial Application Databases using UML-based Repositories Aligned with International Standards in Geomatics. In *Proceedings of ACM GIS*, Washington DC, USA, 2000.

[5] P. Burrough and R. McDonnell. *Principles of Geographical Information Systems*. Oxford University Press, 1998.

[6] H. Ditt, L. Becker, A. Voigtmann, and K. Hinrichs. Constraints and Triggers in an Object-Oriented Geo Database Kernel. Bericht 3/97-I, Institut für Informatik, Westf. Wilhelms-Universität, 1997.

[7] J. Filho and C. Iochpe. Specifying Analysis Patterns for Geographic Databases on the Basis of a Conceptual Framework. In *Proceedings of ACM GIS*, Kansas City, USA, 1999.

[8] M. Fowler. *UML Distilled: Applying the Standard Object Modeling Language*. Object Technology Series. Addison-Wesley, USA, 1997.

[9] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns. Elements of reusable Object-Oriented Software*. Addison Wesley, USA, 1995.

[10] S. Gordillo, F. Balaguer, C. Mostaccio, and F. D. Neves. Developing GIS Applications with Objects. A design Patterns Approach. *GeoInformtica*, 3(1):7–32, 1999.

[11] S. Gordillo, F. Balaguer, and F. D. Neves. Generating the Architecture of GIS Applications with Design Patterns. In *Proceedings of ACM GIS*, Las Vegas, USA, 1997.

[12] S. Guptill and J. Morrison, editors. *Elements of Spatial Data Quality*. Elsevier, 1995. The International Cartographic Association.

[13] ISO. Information Technology - Unified Modeling Language (UML). Draft International Standard (DIS) 19501-1, ISO, 2000.

[14] ISO. Geographic information - Quality evaluation procedures. ISO/TC 211 19114, ISO, 2001.

[15] ISO. Geographic information - Quality principles. ISO/TC 211 19113, ISO, 2001.

[16] B. Pemici. Objects with Roles. In *Proceedings of the Conference on Office Information Systems*, pages 205-215, Cambridge, MA USA, 1990.

[17] R. Price, N. Tryfona, and C. Jensen. Extended Spatiotemporal UML: Motivations, Requirements, and Constructs. *Journal of Database Management. Special Issue on UML*, 1 1(4), 2000.

[18] M.-J. Proulx and Y. Bédard. Perceptory, Centre de Recherche en Geomatique, Universite Laval, Quebec, Canada., 2000. http://sirs.scg.ulaval.ca/perceptory.

[19] J. Richardson and P. Schwarz. Aspects: Extending Objects to Support Multiple, Independent Roles. In *Proceedings of ACM SIGMOD*, pages 298-307, Denver, USA, 1991.

[20] N. Tryfona and C. Jensen. Conceptual Data Modeling for Spatiotemporal Applications. *GeoInformatica*, 3(3):245–168, **1999.**

[21] N. Tryfona and C. Jensen. Using Abstractions for Spatio-Temporal Conceptual Modeling. In *Proceedings of ACM SAC*, Como, Italy, 2000.

[22] M. Worboys. *GIS: A Computing Perspective*. Taylor and Francis, 1995.