An Object-Oriented Method for Designing Geographical Information Systems*

Jonás Arturo Montilva C.
Departamento de Computación , Universidad de Los Andes
Facultad de Ingeniería, Grupo de Ingeniería de Datos y Conocimiento
Mérida, Mérida, 5101, Venezuela

ABSTRACT

The need for using an object oriented approach to develop geographical information systems (GIS) has been widely recognized by the GIS community. However, very few object oriented methods for developing GIS have been published. Most of them focus only on the process of designing GIS databases. The analysis of the GIS context and the need for dealing with multiple views of the database are ignored in these methods. This paper introduces an object-oriented method for dealing with the process of analyzing and designing GIS. The notion of *organizational system* is introduced here to deal with the GIS context. An organizational system is a human activity system whose processes require information about the spatial entities of a geographical region. A GIS provides the spatial and thematic information needed by users in the organizational system to perform its processes. Analyzing and modeling the organizational system is, therefore, crucial for identifying the requirements, entities and users of the GI S.

The method is based on an object-oriented model that represents the entities of an organizational system, their spatial and thematic properties and relationships and their behavior. The main output of the method is the design of a spatial database and the application programs needed to satisfy the requirements of a GIS and its organizational system. Steps are included in the method for mapping an object-oriented database schema to relational and vector database schemas.

Keywords:
Geographical Information Systems, Spatial Databases, Object-Oriented Methods.

1. INTRODUCTION

During this decade, there has been a growing interest in the IS community for enhancing the semantics capabilities of traditional information systems with new and multiple types of data, including spatial, temporal, audio, still images, and video data. Geographical information systems (GIS), in particular, are becoming increasingly important because of the emergence of business information requirements concerned with the manipulation of image and spatial data. The need for dealing with geo-referenced data is common to many different contexts. Besides the current applications in Geography and Earth Sciences, new application domains are emerging in many different areas, including electrical, civil, transportation and chemical engineering, marketing and tele-communications.

GIS are concerned with the collection, organization, storage, analysis and visualization of geo-referenced data about spatial entities of the earth. The central component of a GIS is its database. A GIS database is composed by spatial and non-spatial data, also called thematic data. Objects in a GIS database represent a collection of real-world entities associated with a given geographical space. Spatial data capture the geometrical and topological properties of the entities; whereas thematic data are related to non-spatial properties of those entities. The main difference between a GIS and other types of information systems rests, precisely, on the ability of a GIS to integrate spatial and thematic data and the functionality provided to manipulate and visualize the spatial data.

Although, there exists an extensive collection of methods for developing information systems, most of them are very inadequate or limited in their ability to deal with GIS. This is because a GIS imposes special requirements upon the process of designing a system. Additionally to the manipulation of spatial data, a GIS demands a high degree of human-computer interaction and the integration of alphanumeric, graphical and imaging data types. These requirements cannot be easily satisfied with traditional methods for

* Published at the Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics - SCI´97. N. Callaos, S. Y. Lee, and C. Temponi (Eds.). Caracas, Venezuela. Julio, 1997.Vol. 1, p.145-152.

developing information systems. On the other hand, methods oriented exclusively to GIS application development are scarce [6, 9, 12]. There is, therefore, a growing need in the GIS community for new methods and techniques that take into account the specific requirements and characteristics of the GIS technology.

We propose, in this paper, a new method for dealing with the process of analyzing and designing GIS. The method is built upon proven principles, methods and techniques from Object-Oriented Software Engineering, Object-Oriented Databases and Systems Engineering. The object modeling notation of OMT [15] is used for representing the GIS application domain and designing the spatial database. Database integration is used for merging different views of the application domain.

The paper is organized as follows. Section 2 presents the fundamentals concepts of the GIS technology. It introduces the notion of organizational system as the context of a GIS and an object oriented spatial data model, which is used by the method for designing the spatial database. The general overview and features of the method are given in Section 3. The phases of the method are described, in detail, in Section 4. Finally, the concluding remarks are discussed in section 5.

2. GIS CONCEPTS

GIS are a special kind of information system that have the ability to create, access, analyze, and manipulate spatial data, about a given region of the earth, using digital maps as the unit of manipulation. A map is a model of the spatial distribution of significant entities of a portion of the earth' surface [5].

As any other kind of information system, a GIS is part of a human activity system which uses the information produced by the GIS to perform many different processes concerned with spatial entities. In this section, we discuss the nature and characteristics of the context of a GIS, as well as the object-oriented data model used as the conceptual basis for designing the spatial database of a GIS.

2.1. THE CONTEXT OF A GIS

A GIS provides information to a wider system called here the *organizational system.* An organizational system is a human activity system whose processes require different resources, such as human effort, money, raw materials, technology and information. The objective of a GIS is to produce the spatial and thematic information needed by users in the organizational system to perform its processes. Information provided by a GIS is concerned with the entities of a given geographical application domain, i.e., that portion of the earth' surface that is of interest to the organizational system.

An organizational system may be seen as an organized set of activities or processes that are performed by its actors with the purpose of achieving a set of pre-defined objectives. Processes are activated by the occurrence of events. An event is an action of very short duration that takes place inside or outside the organizational system. In the execution of processes are involved a set of spatial and non-spatial entities. Each entity has a set of properties whose values define the state of the entity. Both processes and events changes the state of the entities. The main components of an organizational system and their relationships are shown in Figure 1.

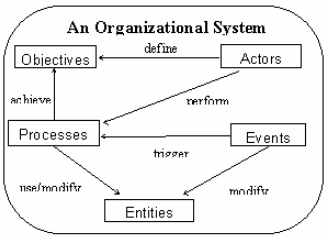


Figure 1. Components of an Organizational System

Properties of the spatial entities in an organizational system are classified into spatial and non-spatial properties. Spatial properties are, in turn, classified into geometrical properties (e.g., location, form and direction) and topological relationships (e.g. adjacency, connectivity and inclusion between two or more entities). Non-spatial properties, also called thematic properties, are those properties of an entity that are not directly related to space (e.g., chemical, physical, psychological, social, temporal, economical, and naming properties).

A GIS database is a model of the geographical application domain of an organizational system. Each relevant entity of the application domain is represented in the database by a software object. Information is produced by querying and manipulating the database.
We believe that analyzing the organizational system of a GIS is a crucial factor for the process of identifying the requirements and designing the GIS database. Our method is based on this premise.

2.2. A N OBJECT-ORIENTED GIS DATA MODEL

In a very simplified manner, a GIS is composed by a spatial database, a graphical user interface and a set of processes and programs for data input, database management and information production and visualization. A spatial database management system (spatial DBMS) is commonly used for updating and querying the database, as well as ma intaining and visualizing spatial data.
Central to the process of designing a GIS is the design of its spatial database. Since our method uses an object-oriented approach for designing the database, an object-oriented data model that describes the constructs for database modeling is required. We propose a very simple data model which provides a set of system defined classes, as shown in Figure 2.
According to this model, a spatial database is a collection of spatial objects. Each spatial object represents an entity of the GIS application domain. A spatial object captures the spatial and thematic properties of an entity.
The structure (attributes) and behavior (operations) of a spatial object are defined by the *Spatial_Object* class. Each spatial object is an instance of *Spatial_Object*. This system-defined class provides a rich set of spatial analysis operations, such as translation, rotation, distances, intersection, area, perimeter or length calculation, as well as Boolean topological operations for determining adjacency, contiguity and inclusion relationships between two spatial objects. Additionally, the *Spatial_Object* class provides a set of operations required to manipulate the object in the database and display its geometrical representation (e.g., save, get, and display).
Based on their geometric properties, spatial objects are classified into three types: point spatial object, linear spatial object and polygonal spatial object. A point spatial object has a shape defined by a point graphical object. A linear spatial object has a shape made of a set of contiguous segments of line. A polygonal spatial object has a shape made of one or more polygons. The structure and behavior of these types of spatial objects are defined by the subclasses *PointSpObject*, *LinearSpObject* and *PolygonalSpObject*, respectively.
Our model provides independence from the vector and raster spatial data models traditionally used in implementing GIS. The classes *PointRepresentation*, *Line-Representation* and *PolygonRepresentation* hide the details of implementing the geometrical and topological properties of spatial objects. This feature is important because during the conceptual design of a spatial database the designer should not be concerned with the implementation details imposed by low level spatial data models, such as the vector and raster data models.
Maps are the unit of visualization of spatial data in a GIS. Each map has associated a set of spatial objects of the same or different class. A map is visualized by displaying the geometrical properties of its spatial objects in a window. We assume here the use of a unique scale for all spatial objects in the database, but it is also possible to represent and manipulate spatial objects at different scale, as shown by Roberts, et al [14] and Montilva [10, 11].
Town, *River* and *Road* are user-defined classes that specialize the system-defined classes of the model. They inherit the spatial structure and behavior defined by *Spatial_Object*. It should be noted, in the given example, that all non-spatial or thematic attributes are user-defined. They are proper of each class defined by the user.
A designer creates the conceptual schema of a database by specializing the system-defined classes provided by the model, as exemplified by Figure 3.

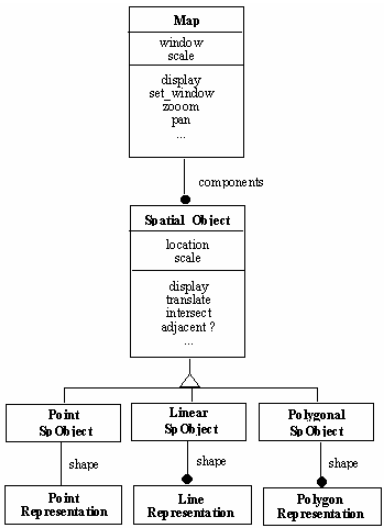


Figure 2. The schema of the OO-GIS model core

3. AN OVERVIEW OF THE METHOD

We assume that two previous activities have been performed before the method can be applied:

1) Management Commitment.- The need for developing a GIS must be recognized by a group of managers and users directly involved in its future use. Financial and human resources must be assigned to the project based on a project planning process.
2) Project Planning.- A feasibility study and a project plan must have been defined before initiating the process of developing the system. A development team must be organized. It is comprised by software engineers, users and specialists in Earth Sciences.

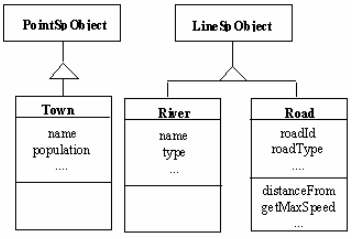


Figure 3 Using the OO data model

The proposed method is divided into four different phases, as shown in Figure 4. Each phase is divided into steps and each step into tasks, as explained in Section 4.

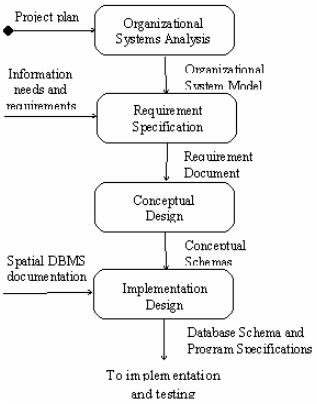


Figure 4 The phases of the method

The first phase, Organizational System Analysis, is devoted to understand the context in which the GIS will operate. A good understanding of this context is achieved by modeling the structure of the organizational system.
The second phase, Requirements Specification, determines what the GIS is expected to do. The details of the information to be produced by the GIS, as well as the input data processes, the spatial analysis operations and the cartographic modeling requirements are some of the requirements to be specified in this phase.
The Conceptual Design phase produces a set of object-oriented external schemas, one schema for each main business process identified in Phase 1. An integrated conceptual schema is then produced through a database schema integration process. Procedures and programs required for spatial analysis and cartographic modeling and map production are outlined here.
Finally, the Implementation Phase maps the integrated conceptual schema to an implementation schema. The latter is based on the data model supported by the spatial DBMS used for implementing the GIS.
The main results of the method is a set of design specifications which includes the design of the spatial database and the detailed design of procedures and programs required for spatial analysis, cartographic modeling and map production.

4. THE PHASES OF THE METHOD

Phase 1: Organizational System Analysis
The basic premise of this phase is that a good understanding of the organizational system of a GIS is essential for a comprehensive definition of its requirements. Knowledge acquisition about the organizational system is a success factor for the development of a GIS. An organizational system can be known through a system modeling process that captures and represents the system components: its objectives, its actors, its processes, its entities and its events. The geographical application domain is, also, defined in this phase. The steps of this phase are defined based on these components.
The main inputs to the phase are the existing documentation and the knowledge that users have about the organizational system. The output of the phase is a system model that describes the components of the organizational system.
Step 1.1: Defining the organizational system
This step involves the identification of the organizational system, the definition of its objectives and the delimitation of its geographical application domain.
Let us take, for example, the design of a GIS for the prevention and control of accidents in an industrial plant. The objectives of the organizational system are: (1) to reduce the probability of accident occurrences and (2) to minimize the impact caused by accidents on people, environment and plant installations. The geographical application domain is the geographical region surrounding the plant installations that can be affected by an accident.
Step 1.2. Modeling the organizational system processes
To achieve its objectives, an organizational system structures and performs a set of processes or activities. This organized set of processes has usually a very complex structure that can be modeled using a hierarchical functional decomposition process. The result is an acyclic directed graph whose root is the main function of the organizational system. Its internal nodes are processes and its leaves are tasks. The higher level nodes represent processes at a high level of abstraction, meanwhile the low level nodes stand for detailed activities and tasks. Figure 5 illustrates the process model of the organizational system for prevention and control of accidents (PCAS).

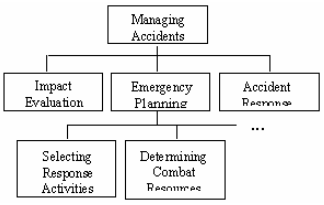


Figure 5. A model of processes of an organizational system

Step 1.3. Determining the actors

Actors are the people who are involved in the execution of processes of the organizational system. Actors could be external or internal to the system. Most of the actors will be users of the GIS. Internal actors are members of the units (e.g. departments, divisions or groups) that comprise the organizational structure of the system.

Step 1.4. Identifying the main types of entities

The execution of processes in the organizational system uses, consumes, produces, requires or involves a set of entities. An entity is a concrete or abstract object that is of interest to the organizational system. Entities have properties, some of them could be related to space (e.g. the location and the shape of the entity). Based on their common properties, entities are grouped or classified into entity types or classes. The purpose of this step is to identify the classes or types of entities that are relevant to the organizational system. In the PCAS example, the main type of entities are: installations, roads, rivers, lakes, oil pipelines, vegetation, towns and emergency plans.

Step 1.5. Establishing the events

An event is any action or incident that occurs into or outside the organizational system. The execution of processes into the organizational system is caused or triggered by events. An event may change the state of an entity. The state of an entity is the set of values associated with its properties at a given time. Examples of events in the PCAS system are the occurrence of an accident, the failure of an equipment and the occurrence of a natural disaster.

Phase 2. Requirements Specification

A GIS is a system that provides a set of information services to an organizational system. This set of services, as well as the operating restrictions and characteristics that the GIS must comply, are some of the requirements that must be established and documented in this phase. A requirement is a detailed description of a service, need, feature, property, restriction or standard to be satisfied by a GIS. Requirements are the main input to the design phases and are used for verifying the design and implementation of the GIS.

We do not enforce any particular technique for specifying requirements. The teamwork should decide which technique is the most appropriate for each case. There exists a wide set of techniques for requirement specification, including user interviews, questionnaires, data flow diagrams and prototypes.

The output of this phase is a document that describes and organizes the set of requirements. Requirements could be classified as follows:

1) Information requirements.- They are those information services to be provided by the GIS, including base and thematic maps, diagrams, statistics and tabular reports. Queries to the database are also a kind of information requirements.
2) Spatial data analysis requirements.- They are concerned with the manipulation of spatial data. Spatial data operations could be classified into: thematic variable analysis, spatial analysis and cartographic modeling [1].
3) Database requirements.- This type of requirement is concerned with the structure and management processes involved in creating, using and maintaining the spatial database. They are closely related to the software tool used for implementing the system.
4) Interaction requirements.- They define the characteristics of the user-GIS interaction, as well as the input of spatial and thematic data and the production of spatial information. The spatial data sources and the processes of capturing and digitizing the spatial data should be specified as interaction requirements.
5) Quality attributes.- They specify a set of factors o conditions that determine the quality of the system. Some of these attributes are the following: database interoperability, portability, efficiency and usability.
6) Restrictions.- The use and development of a GIS could be restricted by a set of factors such as costs, computing resources available, management policies, etc. For instance, the use of a particular spatial DBMS could be imposed by a corporate management policy.

Requirements of types 1 and 2 should be defined for each process included into the hierarchical structure that model the processes of the organizational system. This structure is described in the step 1.2.

Phase 3. Conceptual Design

The objective of this phase is to design the conceptual schema of the GIS database. This schema represents the structure and behavior of the spatial and non-spatial entities that are associated with the organizational system. A schema determines the conceptual content or structure of a GIS database.
Each process of the organizational system involves a set of entities. This set may be different from one process to another. The same entity may be understood differently by users at two or more processes. It is, therefore, convenient to build a schema for each process and then integrate the resulting schemas to produce a global or integrated schema. The schema associated with a process is called a partial conceptual schema and the result of integrating all partial schemas is called a global conceptual schema.
For modeling the schemas we use the object modeling notation employed in OMT [15]. Any other object-oriented notation (e.g., Booch [4], Fusion [7] or UML [13]) could also be applied for this purpose.
Schemas integration is a well known process that is widely described in the literature [2,3]. This phase is adapted from the method proposed by Batini, Ceri and Navathe [3].

Step 3.1: Designing the partial conceptual schemas

For each major process of the organizational system, as identified in the hierarchical structure of the step 1.2, a conceptual schema is built using the OMT notation. Figures 6 and 7 show two partial conceptual schema corresponding to two different processes of the PACS example.

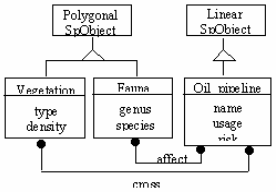


Figure 6. A partial conceptual schema

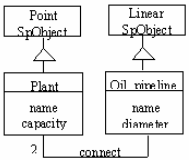


Figure 7. Another partial conceptual schema

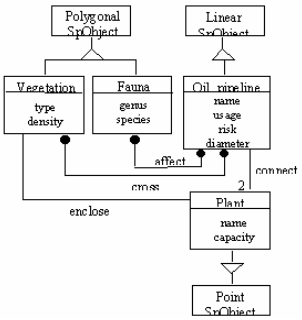
Step 3.2: Preintegration

The number of partial conceptual schemas to be integrated depends on the complexity of the hierarchical structure of processes. If this number is greater than three, defining an order for integration becomes then necessary. Schemas are integrated by pairs. The result of integrating the first pair of schemas is integrated with a third schema and so on.
Structural and semantic differences between schemas should be identified before initiating the integration. Some of these differences are due to homonyms and synonyms used when naming classes, attributes and associations in two schemas.

Step 3.3: Integrating process conceptual schemas

The partial conceptual schemas are integrated in this step using the order defined above. The structural and semantic conflicts between the components (i.e., classes, attributes, domains, relationships between classes) of each pair of schemas should be solved before merging them.
Figure 8 illustrates the result of integrating the partial schemas shown in Figures 6 and 7.

Figure 8. An integrated conceptual schema



Step 3.4: Verifying the integrated conceptual schema
The global conceptual schema produced in the step 3.3 is now verified using the information requirements specified in the Phase 2. Verifying the schema consists of determining if the schema contains all the data needed to produce each information requirement. The global schema should be improved, if it fails to satisfy any information requirement.

Phase 4. Implementation Design
The implementation design is the process of converting a global conceptual schema to one that can be directly implemented in the spatial DBMS being used for developing the GIS.
Each spatial DBMS uses its own spatial data model. For instance, ARC/INFO® uses two different models for storing and manipulating GIS databases. The Arc/Node model is used for spatial data; whereas the Relational Model is used for thematic data.
Step 4.1: Converting the global conceptual schema to an implementation schema
Before converting the global schema to an implementation schema, it is necessary to define a mapping between the constructs used in the source object model and those used in the implementation data model. This process is explained here using the OMT object model as a source, and the ARC/INFO model as a target.

1) Each spatial class of the object model is converted into one or more coverages in the Arc/Node model and one table in the relational model. For example, the Vegetation class in Figure 8 is converted into one coverage and one table. The Vegetation coverage is a collection of polygons. One or more polygons for each type of vegetation existing on the geographical application domain. The Vegetation table contains two columns: type and density. An identifier attribute should be defined and added to both the coverage and the table. It is used to link the coverage with its corresponding table.
2) Each non-spatial class of the object model is converted into a table in the relational model. The process of converting an OO schema into a relational schema is described in detail by Rumbaugh, et al [15].

Step 4.2: Designing the procedures and programs
Some spatial data analysis requirements, specified in Phase 2, are too complex to be satisfied by a single operation into the language provided by the spatial DBMS. A procedure or program should be designed in order to satisfy these types of requirement.
Similarly, the processes of digitizing spatial data and producing cartographic information could depend upon the characteristics of the spatial DBMS. These processes should be described here in detail based on the guidelines provided by the DBMS supplier.
Step 4.3: Verifying the implementation design
Finally, the implementation design must be verified using the requirements. As mentioned before, the process of verifying a design consists of determining the ability of the implementation schema to satisfy the requirements specified in Phase 2. Iteration of phases 2 - 4 could be necessary to improve the schema and accommodate new users requirements before proceeding to the implementation and testing phases.

5. CONCLUSIONS

An object-oriented method for designing geographical information systems has been introduced in this paper. The distinctive features of the method are: (1) its ability to represent the context of a GIS in terms of

objects, processes and events; (2) its use of the notion of organizational systems to gain a better comprehension of the requirements; and (3) its applicability to existing spatial data models and tools.
To the best of our knowledge, very few methods specifically devoted to the development of GIS exist in the literature. General ideas about how to design or model an object-oriented GIS database are described by Davis and Albuquerque [8], Nativi [12], and Sacchi y Sbatella [16]. A more complete method, proposed by Clarke [6], does not take into account neither the analysis of the organizational system nor the use of object orientation.
The method has been applied to the design and development of a GIS for studying and estimating the impact of mass movements and earthquakes on the population and service infrastructure of the city of Mérida in Venezuela. It has also been extensively used as a teaching instrument for developing small projects in several courses on GIS conducted at the University of Los Andes in Mérida.

6. REFERENCES

[1] Bosque Sendra, J. *Sistemas de Información Geográfica* (in Spanish). Ediciones RIALP, Madrid. 1992.

[2] Batini, C. and Lenzerini, M. A Comparative Analysis of Methodologies for Dtabase Schema Integration. *ACM Computing Surveys*, Vol. 18, No. 4, 1986.

[3] Batini, C. Ceri, S. and Navathe, S. *Conceptual Database Design: An Entity-Relationship Approach.* Addison-Wesley, 1992.

[4] Booch, G. *Object-Oriented Analysis and Design with Applications*. Second edition, Addison-Wesley, 1994.

[5] Burrough, P.A. *Principles of Geographical Information Systems for Land Resources Assessment.* Oxford University Press, 1986.

[6] Clarke, A.L. GIS Specification, Evaluation, and Implementation. In Maguirre, D.J., Goodchild, M.F. and Rhind, D.W. (Eds.) *Geographical Information Systems.* Vol. 1., Longman Scientific Technical. 1991

[7] Coleman, D. et al. *Object-Oriented Development: The Fusion Method*. Prentice-Hall. 1994.

[8] Davis, C.A. and Albuquerque, K. Object-Oriented GIS in Practice. *Journal of the Urban and Regional Information Association (URISA)*. 1994. pp. 786-795.

[9] Laurini, R. and Thompson, D. *Fundamentals of Spatial Information Systems*. Academic Press, London, 1992.

[10] Montilva, J.A. *An integration methodology applied to the design of a data/knowledge model for multimedia and spatial applications*. Ph.D.Thesis. University of Leeds, School of Computer Studies, Leeds, UK, January, 1993.

[11] Montilva, J. A. A methodology for the design of specific-purpose data models applied to multimedia and geographical information systems. *Proc. of the 5th Int. Symp. on Systems Research, Informatics and Cybernetics - ISAS´95*, Baden-Baden, Alemania, Agosto, 1995.

[12] Nativi, S. A Conceptual Modeling for the GIS Developing. *Proc. of the Five European Conference on Geographical Information Systems* (EGIS`94), 1994.

[13] Rational Software Corporation. *The Unified Modeling Language: Notation Guide* (http://www.rational.com), 1997.

[14] Roberts, S.A., Gahegan, M.N., Hogg, J., and Hoyle, B. Application of object-oriented databases to geographic information systems. *Information and Software Technology*, Vol.33, No.1, 1991, pp.38-46.

[15] Rumbaugh, J. et al. Object-Oriented Modeling and Design. Prentice-Hall, NY. 1991.

[16] Sacchi, L. and Sbatella, L. An Object-Oriented Approach to Spatial Databases. *Proc. of the Five European Conference on Geographical Information Systems* (EGIS`94), 1994.