APPROACHES TO DATA MODELLING IN GIS

Irena Mitášová, Jana Višòovcová, Milan Hájek
Laboratory of Geoinformatics
Department of Theoretical Geodesy of Slovak Technical University, ArtInAppleS, Ltd.
Radlinského 11
813 68 Bratislava, Slovakia
e-mail: mitasova@svf.stuba.sk, jvi@artinapples.sk

Spatial data, regardless incoming sources and methods, define and describe particular data model of environment. As an objective of this paper, process of data modelling with several approaches is briefly discussed. Emphasis is given to object-oriented data model.

Data model of environment is a starting point in process of digital maps, databases and GIS creation. In presence, creation of new digital data models is needed, as well as the conversion and actualisation of existing analogue models.

Present spatial databases are of different purpose and structure. They also use various sources of spatial data. The most often it is the case of multiscale data from maps, photogrammetric and satellite images or geodetic data of various structure and accuracy. For management and implementation digital spatial data various technologies are used. They can be aimed at automated mapping (AM) or geoinformation systems (GIS).

In the following discussion of digital spatial data we will follow two basic assumptions. First one is that data capture, maintenance and usage is supported by computer technologies. Secondly, acceptance of specific differences of digital maps from spatial databases of GIS is needed. We will come from an idea that presentation of data in the form of maps is one of GIS outcomes, however, information functions of GIS are much more wider and flexible.

1. Data Modelling of Environment

Object of data modelling of environment are both its basic components: physical-geographical sphere and social-economical one as well. Process of data modelling of environment can be simply imagined on the basis of Fig 1. Model of subjected environment created by geodetic, cartographic and photogrammetric methods should have structure, contents and accuracy enabling to re-create it functional data model. This data model should not only be able to be processed by computer technologies, but also to be simply used for many purposes and users.

General tool for description of environmental elements is **entity-attributes-relational concept**. It is based on classification of environmental entities and their classes, as well as on definition of attributes and relations among them. There can be found all types of relations in environmental data models - from 1:1 and n:1 to n:n.

Layering-relational concept of modelling is based on mapping and relational database technologies. Environmental entities are situated in the layers of digital map, their attributes are processed by relational database management system (RDBMS). RDBMS is based on organised tables. RDBMS has "borders" mostly in actualisation, redundancy and data exchange between users and systems.
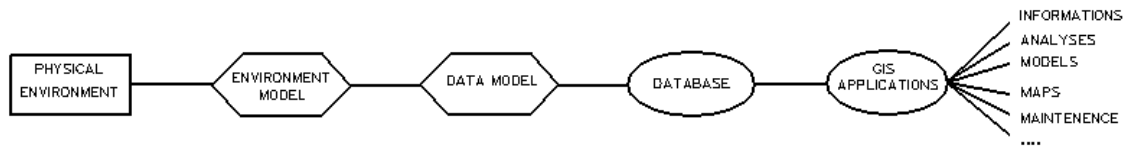
Fig 1 Procedure of environmental data modelling

In an **object-oriented concept** of data modelling are real world entities represented by objects. Object is characterised by properties as well as methods for their changing. Objects are grouped to classes. Among classes is functional mechanism of hierarchy and inheritance. Object oriented database management systems (OODBMS) are more progressive from the point of view of complicate structures of data actualisation and usage of spatial data for multimedia and computers networks.

Data modelling is a process, in which particular types of models with various information ability enter into relations. Successive increase of information functions in the data modelling process is described on Fig 2. The most accurate model of subjected area is picture model, e.g. photogrammetric or satellite image. Images are processed using transformations and digitalisation in order to get model of area in defined coordination system suitable for technical purposes. This way is obtained second environmental model (respectively its approach), in the form of original digital map. Digital map contents is usually divided into levels, it has standard cell library and geometrically-graphical elements (points, lines, areas, text). Such a digital map can be created as a geometric model or object model. In the first case are points connected by lines, signed by cells and description and terminology is given. This form of digital map model can be used for graphical presentation and documentation. Object model of digital map contains complex graphical elements representing real world entities. Lines and areas are closed and exactly defined. Object model of area description is more effective and able to be later improved. Attaching user attributes data to geometric description of graphical objects, e. g. in the form of relational database tables is obtained third model of subjected area. In such a model, manipulation and queries can be made and supported by graphical functions (e.g. filling of selected elements). Following model is get by defining of mutual relations among objects - topology. Topologic model of environment is the base of spatial analysis so important for GIS. It can be realised on various level using various software tools.

2. Topological Data Model

Topology, without using it as a mathematical discipline, can be applied on modelling of basic relations among real world entities. By topology, they can be exactly defined by planar graphs, their nodes and polygons, borders of shapes. Using this tools, we can express:
- continuity - vertexes are connected by edges,
- polygonization, i.e. definition of areas - by connecting edges is created polygon abounding an area,
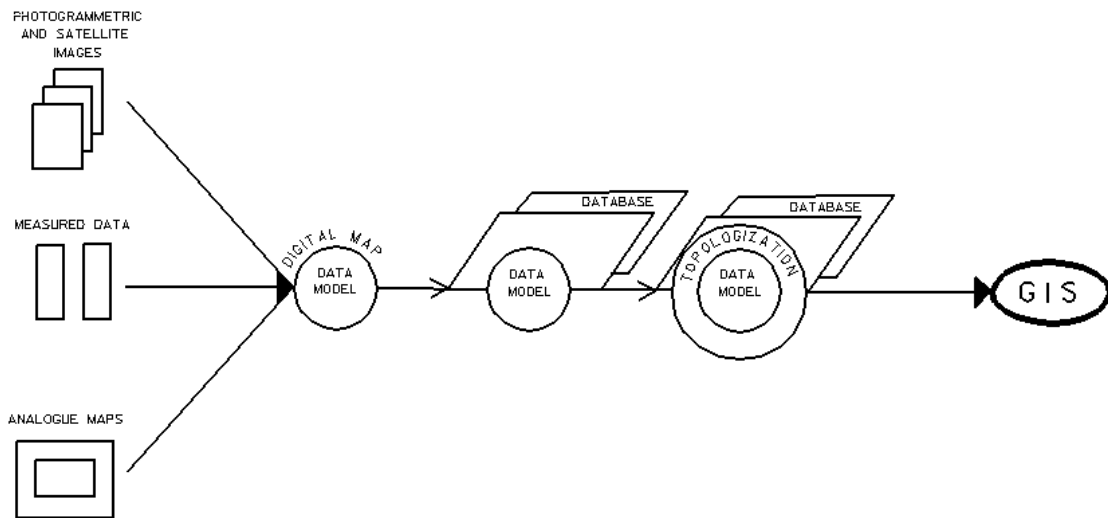- adjacency - oriented edges have right and left polygon (area).

Fig 2 Increase of information functions in the process of environmental data modelling

Sample of topology of area part with nodes and vertexes as elements, and their geodetic coordinates, borders and cartographic symbols is depicted on Fig 3 and in Tables 1-4. As can be derived also from the figure, topology is relevant for vector or vectored data.

In spite of that topology is for surveyors and cartographers term easy to understand, it is far from being a trivial matter. Many CAD products are not provided by topology functions. Some products, declared as ones on GIS level do not have tools for topologization and its usage. Or, this functionality is covered by separate modules. However, without topology is not consequently possible to solve spatial analyses. Example of spatial analyses types can be: which lands are affected by projected infrastructure; which parcels are touched by water, air or soil contamination from particular source and so on. Already in the process of digital spatial data creation we make topologization (and that is why also their usage not only for GIS) simply or complicate.
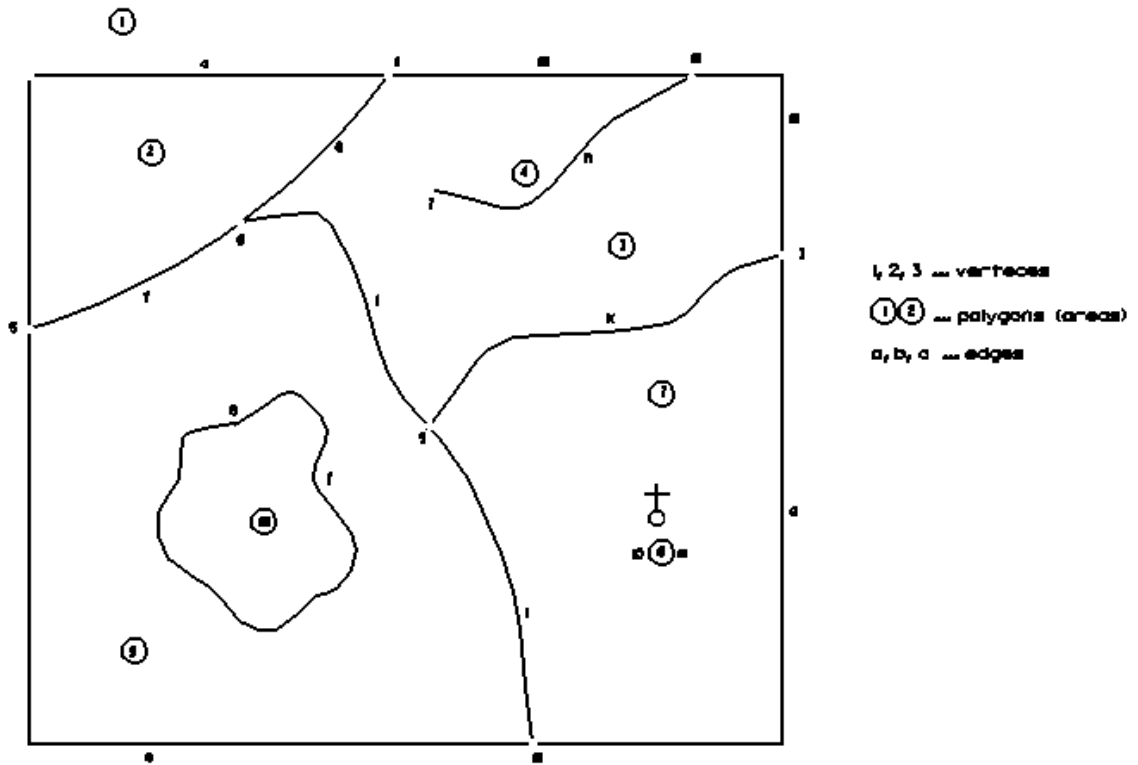
Fig 3 Topological description of area

Table 1 **Topology of polygons (areas)**

Polygon	Edges
1	surrounding
2	a,f,g
3	b,c,k,i,g,1,h
4	h
5	e,f,i,l,0,j
6	j
7	d,l,k
8	m

0 - following edge is an island (hole) in the polygon
1 - following edge is free from one side

Table 2 **Topology of edges**

Edge	from vertex	to vertex	right polygon	left polygon
a	5	1	2	1
b	1	2	4	1
c	2	3	3	1
d	3	4	7	1
e	4	5	5	1
f	5	6	2	5
g	6	1	3	2
h	7	2	3	3
i	6	9	5	3
j	8	8	6	5
k	9	3	7	3
l	9	4	5	7
m	10	-	8	-

Table 3 **Incidence of edges and vertexes**

Vertex	Edges
1	a,b,g
2	b,c,h
3	c,d,k
4	d,e,l
5	a,e,f
6	f,g,i
7	h
8	j
9	i,k,l
10	m

Table 4 **List of coordinates of vertexes and edges' points**

Edge	List of coordinates x,y
a	100,300 100,500 400,500
b	400,500 650,500
c	...
...	...

3. Data Models of Environment

For an effective creation of graphical and non-graphical digital data, various data models are created. They came from basic approaches to environmental data modelling, from methods and technologies for further processing and usage. Three data models can be mentioned in connection with the general approaches to data modelling: relational, object-based and object-oriented model. Their basic principles, characteristics, possibilities and drawbacks were worked out in ArtGIS Educational Training Program [3]. This document and software was developed by co-operation of people of Laboratory of Geoinformatics, common office of Faculty of Civil Engineering of STU Bratislava, Slovakia and ArtInAppleS, Ltd.

3.1 Relational Data Model

Relational data model is characterised by clear distinction between the graphical and descriptive data of territorial elements. Graphical entities are stored in layers of digital computer map, attribute data in form of relational database tables. After mutual interconnection between the graphical and descriptive data, analyses can be made in a model created this way. A land in the digital map is linked through the identifier with the record of the relational database table, containing further descriptive data of the land (Fig 4). Land border and land identifier can be created by any constructive drawing function with arbitrary graphical attributes. Topologization of this way created data requires special software.

Conception of relational data modelling is at present widespread (MGE, ArcInfo, GeoGraphics...) and for many, especially smaller GIS applications, fully satisfactory. Graphical systems and relational databases as well as complex GIS environments have reached throughout their development high efficiency. Nevertheless, existing duality in representing territorial elements can be a potential source of spatial data inconsistency. Limitations of relational data modelling are from practical point of view shown out especially with need of actualisation of bigger and more compound systems.
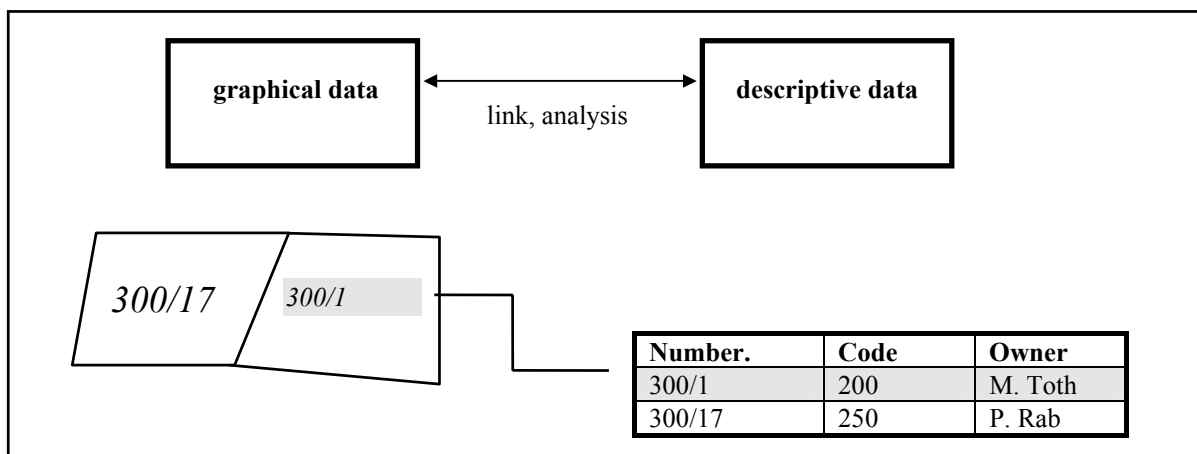

Number.	Code	Owner
300/1	200	M. Toth
300/17	250	P. Rab

Fig 4 Relational data model

3.2 Object-based Data Model

Object-based data model of graphical objects in CAD systems is a partial improvement of relational technologies in the field of geodetic and graphical data. It compounds complex graphical elements representing components of territorial reality (e.g. forest, water area etc.). Single graphical elements have distinctively specified graphical attributes and are connected with their identifiers. Using CAD system MicroStation, forest can be created by function Shape in level 20, violet colour, ..., number of forest land can be attached as a tag with defined graphical attributes to the forest border (see Fig 5). Such an object-based approach is much more exact for expressing of territorial elements. It enables to make contents and technology of data model uniform and standard.
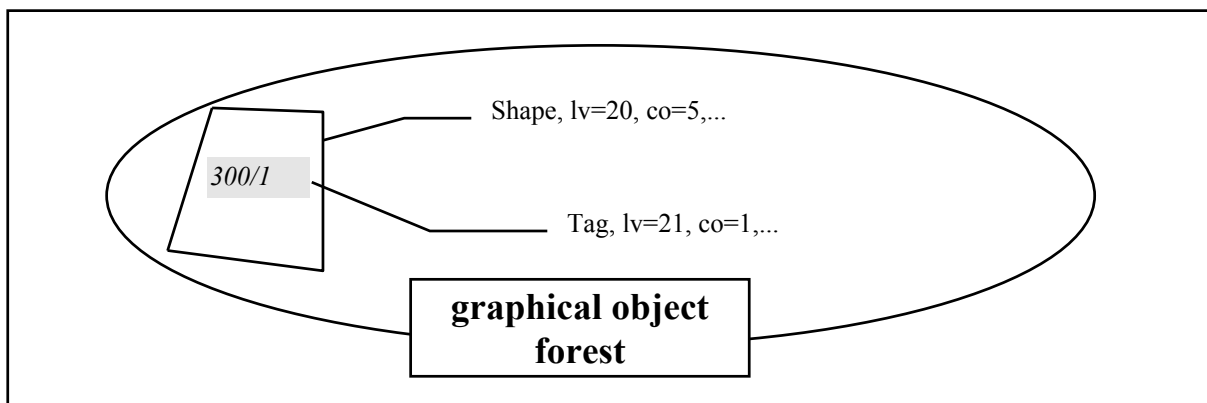

Fig 5 Object-based data model

Object-based model of spatial data is thus a transition between classical relational and "pure" object-oriented conception of territory data modelling. Its repletion gives to processed digital map greater semantic value, larger information contents. We can see the real objects behind graphic image of map's elements and the map is no more "just" a means of visual representing and output creation.

3.3 Object-Oriented Data Model

Object-oriented data model compounds a set of objects interacting one another. Object is characterised by its attributes (static properties) and behaviour (methods, i.e. functions which can object perform). E.g. object forest has its definite attributes as number, code, owner, ..., and it is able to change these attributes. Object can be represented in a different form - graphic image in map, descriptive information, picture etc. It is always just one object, there is no separation of graphics from descriptive data, there are only different forms of expressing the same object (Fig 6). The result is that if some attribute is changed it changes automatically in the "map" and "database" i.e. on all places of its appearance.
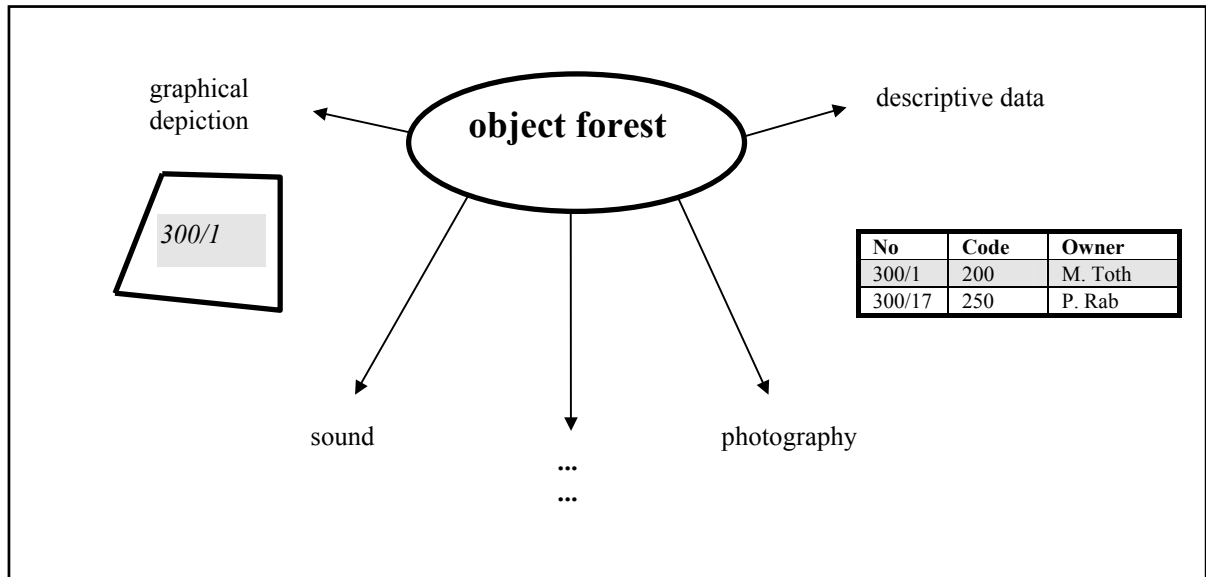
Fig 6 Object-oriented data model

In the object-oriented approach to data modelling the difference between graphical and non-graphical data disappears which ensures their consistency in an application. Object-oriented data model supports theoretically unlimitedly difficult and structured data types and applications. Object-oriented model of spatial objects is a representation of a real world in its full intricacy. It corresponds with the physical application model and topology better than spatial models used up till now.

4. Digital Spatial Data Creation

When considering GIS spatial data quality and usability, it is supposed that quality of spatial data is a compromise between their georeference accuracy and demands for their capturing , processing and usage. These aspects are considered during the project phase of GIS. Then, creation of digital spatial data is influenced by several factors:
- data sources,
- methods and technologies of data capture,
- methods and technologies of further processing and usage.

Sources and methodology of data capture influence mainly their accuracy and actuality. Methodology and technology of data processing determine data structure and form.

Spatial data, regardless their incoming sources and methodologies, describe and define some territorial data model. Computer model of environment is expressed by digital spatial data. In these data is encoded spatial relationship to the places of creation and usage, topology of area, geodetic and geometric definition of area elements. Classification and methods for changing the elements as well. Geodetic and geometric definition of area elements and thematic classification was also used in mapping and creation of analogue maps up till now. Creation of territorial data models, topologization and consequent spatial analysis are inherent in entrance to computer technologies for data modelling and to databases and map creation as well.

5. Conclusion

Maps being used up till now are more or less analogue territorial model. Conversion of this model to digital spatial models, needed for GIS, is connected with necessary data actualisation on one side and with parallel transformation of analogue maps to digital ones without information and accuracy lost on the other side. It is usually complicated process, with georeference, object and topological definition of area elements as its part. To finish this process successfully, it is necessary to use know-how of fast evolution of geoinformation methods and technologies, standardisation, usage and unifying of spatial databases. In an opposite case, huge amount of one-purpose spatial data or isolated "information islands" may be created. And this would definitely be not efficient nor economic.

References

[1] Meruòka, V. - Polák, J.: Objektovì orientovaná analýza a design. Softwarové noviny è. 1, 1995, page 46-57.

[2] Mitášová, I: - Geoinformatika a priestorovo orientované informaèné systémy. Geodetický a kartografický obzor, 41, 1995, No 9, page 183-186.

[3] ArtInAppleS, Ltd.: ArtGIS Training Program. ArtInAppleS, Ltd. Bratislava, www.artinapples.sk.