Design and Implementation of a Spatial Database for Traded Real Estate

Ewa DĘBIŃSKA, Poland

Key words: spatial database, geographic information system, Unified Modelling Language

SUMMARY

A real estate surveyor needs fast access to relevant information to do his job effectively. Every surveyor collects the data on real estates that were subject to market transactions. Depending on the type of surveyed real estate, the data can contain transactions regarding real estate that is developed, undeveloped or residential. Additionally, the database is collected for different areas, districts or cities. While producing the value estimation, the surveyor has to analyze the entire database and select transactions relevant to the type of the real estate under consideration, its localization and update the database if necessary. The above activities make the spatial databases, which are basic for every spatial information system, an ideal solution for real estate surveyor. In addition, the basic tasks allocated to geographic information systems, like entering, storing, analyzing and annotating spatial data are also basic tasks necessary in real estate survey.

At the moment, the most popular database design process is Computer Aided Software Engineering - CASE. Using relevant CASE tools allows increasing effectiveness of database development. The logical model of the database which can be recorded using object modelling language – Unified Modelling Language can be used for automatic generation of the database schema that complies to the specification.

The paper proposes a model of a database on real estate that were subject to market transactions. The ready-made data model contains object classes allowing locating undeveloped, developed and residential estate, as well as object classes relevant for the local spatial development plan, for districts and areas. On top of object classes containing the geometry data, we have also designed the tables for the information of notarial acts and real-estate registers. For the defined classes and tables we have also defined the attributes that are the most relevant ones in the surveying process. We also have defined the database creation and data entry methods, as well as presented using GIS in the surveying process at the stage of data entry and later, database management.)

Design and Implementation of a Spatial Database for Traded Real Estate

Ewa DĘBIŃSKA, Poland

1. INTRODUCTION

A real estate surveyor needs fast access to relevant information to do his job effectively. Every surveyor collects data on real estates that were subject to market transactions. Depending on the type of surveyed real estate, the data can contain transactions regarding real estate that is developed, undeveloped or residential. Additionally, the database is collected for different areas, districts or cities. While producing the value estimation, the surveyor has to analyze the entire database and select transactions relevant to the type of the real estate under consideration, its localization and update the database if necessary. The above activities make spatial databases, which are basic for every spatial information system, an ideal solution for real estate surveyor.

2. THE DEFINITION OF A SPATIAL DATABASE AND GEOGRAPHICAL INFORMATION SYSTEMS

A database is a set of integrated data from a certain domain. A spatial database is a kind of database broadened by a spatial factor, which means that the database includes a record of the geometry of objects, except for the descriptive attributes of objects. Spatial data belongs to understand as data with reference to surface to the Earth (Bielecka 2006).

Spatial databases are a fundamental element of geographical information systems (GIS). The GIS is a system for creating, storing, analyzing, modernizing and managing spatial data (Gaździcki 2001). It needs hardware, specialized software data as well as an educated group of people, which served mentioned components. Good organization and cooperation of mentioned elements can cause thrift of time and costs of work. In order for the system to function correctly, the presence of all elements is essential; however, data seems to be the most important element. Without the correct data, on proper level of minuteness and accuracy, even the best educated group of people and specialized programming cannot give answers to the question assigned by the user, such as: Where is ul. Szeroka located in Krakow? Or which lots are in the neighborhood of lot number 256?

The database decides on the range of capability of the application of the GIS. Therefore, the formative stage of the database is very important. Proper selection of a feature class for the description of a selected piece of the real world during the phase of project design is not only important for essential maintenance of the information range of the system, but it also has an influence on the efficiency of later management data (Makowski 2005).

3. THE PROJECT OF DATABASE

The *Microsoft Office Visio 2003* was used for record a project of database of real estate. Visio is software belonging to CASE tools. CASE means Computer-Aided Software Engineering and is a kind of software which improves the process of the project and is characterized by sophisticated tools for the editing of graphics. CASE tools can use different techniques for the visualization of projects. One of the most often used is the Unified Modeling Language (UML). UML is now the most common standard used for object-orinted modeling. The main idea of UML is to make possible the use of simple means to create a graphic model presenting any fragment of the reality. It is a very good tool for accurate expression of thoughts in the form of a graphic model. Notation UML is a set of legible symbols and marks which can be understood by anyone with a minimum level of knowledge of computer science.

In the suggested model of a spatial database (Fig. 1), there are planned feature classes for residential estates such as developed, undeveloped and residential areas. The aforementioned types of estates are most often encountered in the real estate market. Taking into consideration the fact that suggested database is in development phase, other types of property have been omitted.

The source of information at the spatial nature according to valued estates for surveyors is land and building cadastre (LBC). The basic and the most important surface object in LBC is a parcel. A parcel is possible to define as a continuous area of ground located within the boundary of one unit, homogeneous with respect to law, delimited from surrounded by means borderlines (Hycner 2007). By reason of the above-mentioned, in the suggested model of a database, the parcel is also the main feature class, and based upon this, it is possible to define the location of developed and undeveloped estates. There is a clear similarity between developed and undeveloped estate law classification, because in both cases, there is freehold or perpetual usufruct the main law. Whereas dwelling areas characterize clear dissimilarity with relation to the above-mentioned types of real estates. The main law for dwelling areas is not the freehold but ownership to residential; therefore, interrelation residential location with parcel would be not enough precise information.

In a situation of many buildings on one parcel, one building can be situated more advantageous and another less, and this could exert an influence on the final value of the estate. Therefore, a dwelling area in the model of the database was connected with the feature class *Building,* due to that the concrete dwelling area is recorded to proper location in space - building.

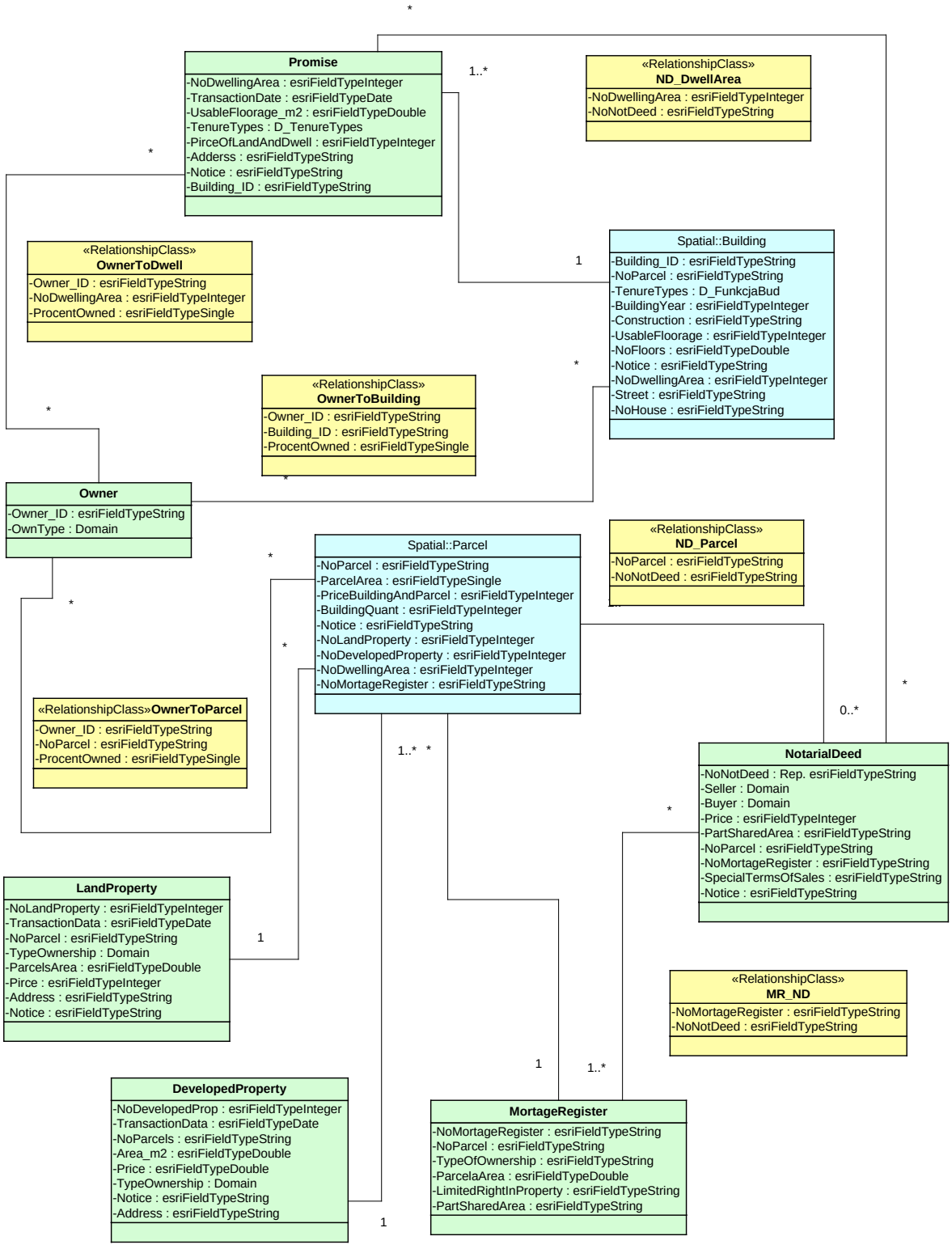

Fig. 1. The project of a spatial database for traded real estate; written in UML.
Source: own recourses

Premise is defined as a table, not as feature class, in which objects have spatial attributes. Obtainment information of site of room estate will be possible through a defined relationship within the table named *Premise* and the feature class *Building.* There are two types of relationships possible: simple and composite. Simple relationships are relationships between two or more objects in the database that exist independently of each other. In the case of a composite relationship, the lifetime of one object controls the lifetime of its related objects. Practical application of a composite relationship could be, for example, in buildings demolition, in which are dwellings. When we delete a demolished building from the database record, all dwellings related to the building will automatically be deleted as well. In schema of database for traded real estate where the more transactions the better that relationship would not be good result. Therefore, the relationship between *Building* and *Promise* is defined as a simple relationship. However, if this situation takes place, a user will have access to all transactions regarding non-existing dwellings, but he does not have the possibility of locating on map a building where there were once dwellings.

In order for the designed database to be fully functional, there are defined tables named *AD* and *RD* for information from authenticated deeds and registers of deeds. In the model of the database, there are feature classes named *Roads* and *LSDP*, which allows for the storing of information about roads and Land Spatial Development Plans in one database. However, it could be difficult to enter vector data for the *LSDP* class. Later in the article, a proposal is provided for a slution to this problem.

At the end of description of designing of database, it is worth to say that this project was forwent consultations with real etates surveyors. As a result of this, every feature class has defined the most essential attributes for producing a value estimation. However, the proposed range of the described information is not final, because the schema of the model UML makes possible his mulitple use and free modification objects and attributes. As a result, a spatial database for trated real etates can fit to the needs and dispositions of the surveyor.

4. IMPLEMENTATION OF A SPATAIL DATABASE

After the design phase, the model was checked by *Semantics Checker* macro. The next step was to automatically generate an empty geodatabase in *ArcGIS*. This process was documented in a report.

Preparation of the geodatabase for use by surveyors includes import spatial data regarding districts, road network, land spatial development plans, parcels and buildings. The *ArcGIS* system allows for the import of data from many formats such as *.shp, *.dgn, *.dwg and *.dxf. Such a prepared database must be complete with information about transactions. From delved result that most surveyors collect data on real estates in Microsoft Excel. Many of them also use the software WALOR. WALOR is a professional tool for appraisers. It records data in Microsoft Access, but a user can revise data only through the interface of the software or through export to the *.xls format. Conversion from the XLS format to the geodatabase is possible by middle format dBASE. Defined in database schema relationships between feature classes and tables make possible final integration of data.

5. APPLICATION OF GIS FOR VALUATION

A database for real estates in GIS is an ideal source of spatial information for surveyors who needs a piece of information for appraisal. Every software belonging to the group described as desktop GIS has tools for finding, sorting and analyzing the described and geographic information in a geodatabase. How we can use the created database in an *ArcGIS* system is described below.

ArcGIS enables the searching of data and later the selection of records in a database by building SQL expressions (Fig. 2). We can use comparison operators like: equal (=), not equal (<>), greater than (>), less than (<), greater than or equal (>=), less than or equal (<=), logical operators like AND, OR, NOT and arithmetic operators, e.g. SELECT *from *Promise* WHERE: [PULokalu_m2] > 40 AND [PULokalu_m2] < 60 AND [DataTransakcji] < 31-01-2008 AND [DataTransakcji] > 31-01-2007

A selection's results can be sorted in one or more columns recorded in the reports as well as in new feature classes or tables. Furthermore, each record in a non-spatial table, e.g. *Promise*, *DevelopedEstate*, can be located on the map thanks to special tools and a previously defined relationships. The localization of an object is carried out singularly, i.e. for each record individually. This is a big limitation, because it is impossible to run spatial analyses for those records. In this case, the best solution is geocoding the tables: *Promise*, *UndevelopedEstate* and *DevelopedEstate.* "Geocoding is the process of assigning a location, usually in the form of coordinate values, to an address by comparing the descriptive location elements in the address to those present in the reference material" (ArcGIS Desktop Help). The result of the geocoding process is a new point feature class. Each point has the same attributes as the related record in the table. The representation of every transaction as a point, which has an XY coordinate, makes possible selection by location during the collection of data by a surveyor. Selection by location allows for the selection of objects based on their location relative to other features, e.g. 'are within a distance of', 'are completely within', 'contain'. Though sometimes just displaying transactions on a map could show their mutual relationships and their dependence on other objects in a neighborhood (Fig. 3.). Most important are the spatial relationships, which are invisible during revision of the data recorded in tables.

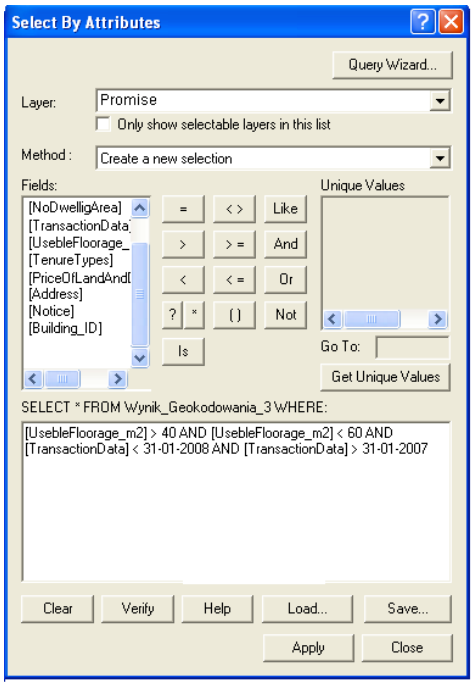

Fig. 2. Select By Attributes dialog box.
Source: own recourses

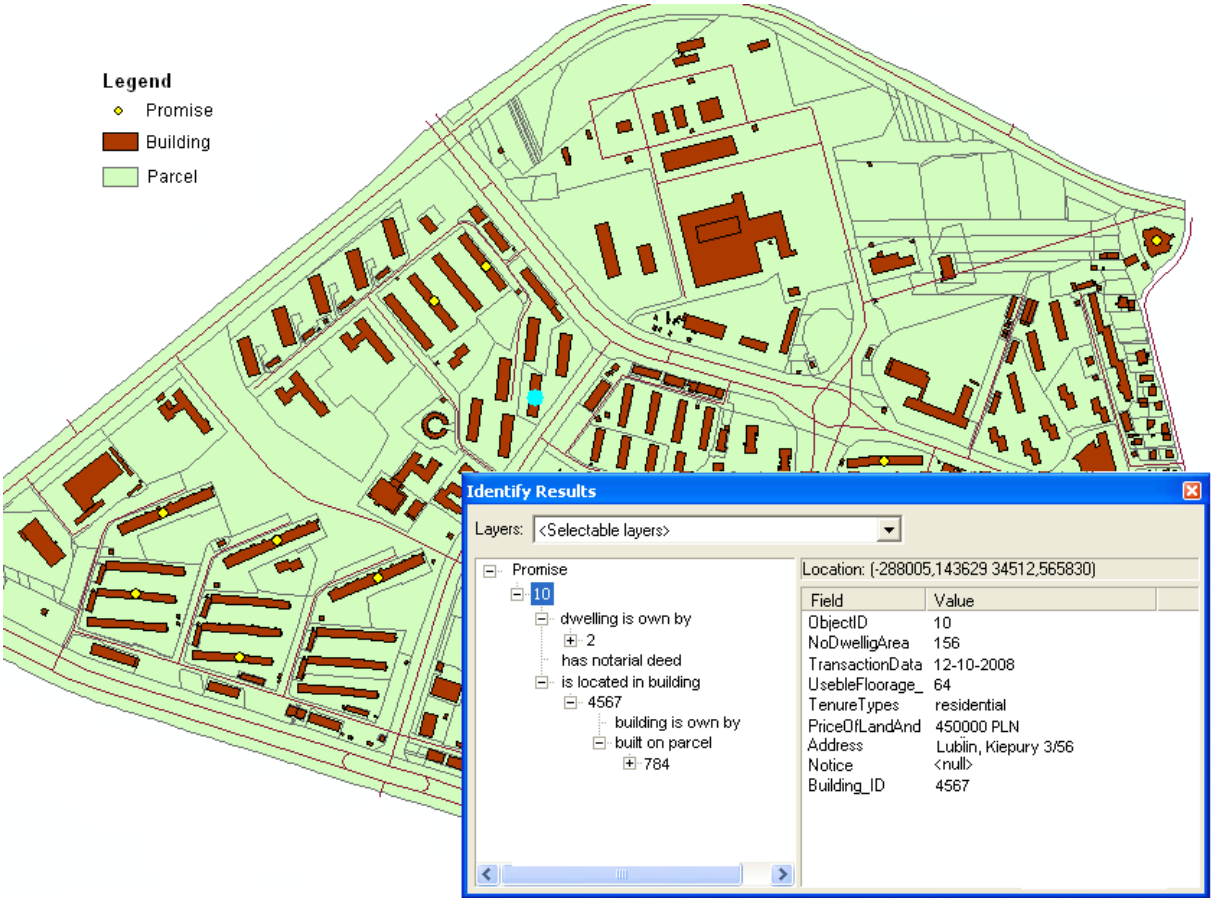

Fig. 3. Location and identification of the selected property's attributes.
Source: own recourses

The GIS organizes geographic data into a series of thematic layers and tables. It makes possible the overlaying of layers in any order. It would be perfect if all the layers displayed on a map could be vectors, but the reality is different, and in Poland surveyors must work with rasters, which often happen according to LSDP. In the case of this data, we can use geo-referencing, which means to assign the coordinates from a known reference system to the page coordinates of a raster (image) or a planar map. It is enough to raster data can be displayed, queried and analysed with other geographic data. Finally, the user decides which layers will be displayed and how the data will be visualized on the map. After all it only need to just one look of surveyor on the map to answer the questions: in which unit is located the valuated property? What are the numbers next units? Is there the LSDP and what is its establishing?

6. SUMMARY

Presented in the article database model is initial proposal. The proposed feature class as well as its attributes could be modified. The author realizes that the proposed database in *ArcGIS* is not handy for people who have never had contact with desktop GIS. Therefore, the author is working on the creation of an application which could make usage of the database easier for traded real estates.

REFERENCE

Bielecka E. 2006. *Systemy informacji geograficznej. Teoria i zastosowania*. Wyd. PJWSTK, Warszawa
Gaździcki J. 2001. *Leksykon geomatyczny*. Wyd. PTIP. Warszawa
Hycner R. 2007 Notes from lecture: Fundamentals of cadastre and real estates management, academic year 2007/08
Kucharska-Stasiak E. 2006. *Nieruchomość w gospodarce rynkowej*. Wyd. PWN. Warszawa
Makowski A. 2005. *System informacji topograficznej kraju*. Wyd. PW. Warszawa.
http://webhelp.esri.com - ArcGIS Desktop Help

CONTACTS

Ms. Ewa Dębińska
AGH University of Science and Technology
Department of Geomatics
Al. Mickiewicza 30
30-059 Kraków
POLAND
Tel. +48 12 617 23 00
Fax + 48 12 617 22 77
Email: debinska@agh.edu.pl