A Scaleless Data Structure for Geographic Information

Markus Mayr

Vienna University of Technology
Department of Geodesy and Geoinformation
Gusshausstrae 27-29, 1040 Vienna, Austria
mayr@geoinfo.tuwien.ac.at
http://www.geoinfo.tuwien.ac.at

Abstract. To handle the problems arising from representing and merging map information on multiple scales, so-called Multi-Representation-Databases (MRDBs) are used. Within these, information is stored in different layers with different levels of detail which in turn are connected by automated generalisation processes. This is still not solved satisfactorily and is, in the opinion of the author, essentially a problem of the data-model used to store geographic information. Changing the way of storing features away from any fixed constraints concerning positional information to a lattice-oriented, topographic, tagging based way to define geographic features, it should be possible to develop a flat, scaleless geographic storage which allows the use of data of any scale.

Keywords: geographical databases, topology, data storage, ontology, scaleless

1 Introduction

To handle multiple scales, so-called Multi Representation Databases (MRDBs) are used. These store multiple layers with different levels of generalisation in one database. The different representations of objects in different levels are linked with one another (see fig. 1). This is necessary to ensure consistency. When an object of any level gets changed, an update has to be triggered along its connections to other objects. This is true for attributive as well as geometric changes. With this setup it is easy to apply changes to the highest level of detail and push the changes to the lower levels (e.g. re-run certain generalisation operators) but it is hard to propagate any update the other way around.

An example of a MRDB is a project called "GiMoDig" [5] which uses such a database to deliver adequate cartographic results. While showing a general map, certain elements are displayed in more detail to emphasise their existence. Hampe [3] is using the advantages of an MRDB to construct a system which allows the collection of data of different detail with the help of a web-interface. The database ensures integrity and delivers data in various scales.

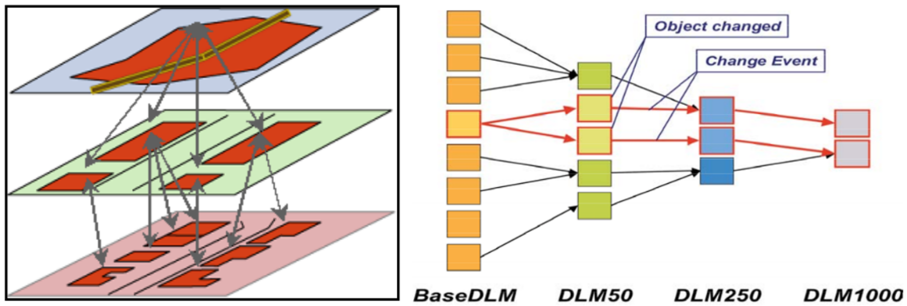

Fig. 1. Links between different levels of generalisation in a MRDB ([5],[4])

2 Scaleless, Scaleable and Scaleful

By *scaleless*, the author means any form of representation or storage which does not split data according to their scale but uses a flat way to handle this data. This is true for access as well as storage. In contrast to that, *scaleable* refers to any form of system that can handle data of multiple scales (e.g. a MRDB) or offers interaction on different scales. In this sense, a scaleless storage can also be a scaleable one, but not all scaleable databases are automatically scaleless. A system can also be *scaleful*. In this case a storage or access method requires the compulsory notion or assumption of a specific scale (even if it is just one fixed scale). Table 1 gives an overview of the authors understanding of the terms scaleless, scaleable and scaleful respectively to data access and storage.

Table 1. Scaleless, scaleable and scaleful access and storage

	data access	*data storage*
scaleless	no scale is given nor is it expected	no specific notion of scale
scaleable	offer interaction on multiple scales	operate on multiple scales
scaleful	specify or expect a scale	explicitly distinguish between scales

The term scaleless is used by different authors (e.g. [8] or [7]) to describe their finding which would be better described as scaleable or scaleful since the scalelessness is directed towards an access method but not the actual storage of data which in their case is scaleful. Vermeij et al [8] describe a data structure which makes it easy to access geographic information of any scale stored in one database by using a multi-dimensional access method. Despite being called this way, this data structure is not really scaleless since data inside the database has to be stored in manipulated form according to its scale to allow access to it. Van Oosterom [7] writes about a so-called Reactive-tree which is described as a data structure which allows fast access to and storage of geographic information at multiple scales. Nevertheless, data storage and access itself is not scaleless, but scaleful.

3 General Idea

In contrast to a classical hierarchical MRDB architecture (figure 2a), the database can be organized in a centralized way. The actual central data storage can be held in a scaleless state (figure 2b) . Other possible representations are generated from this single, scaleless database. While access to the complete database can indeed be scaleful, the primary data store is not. But how to keep a data storage scaleless?

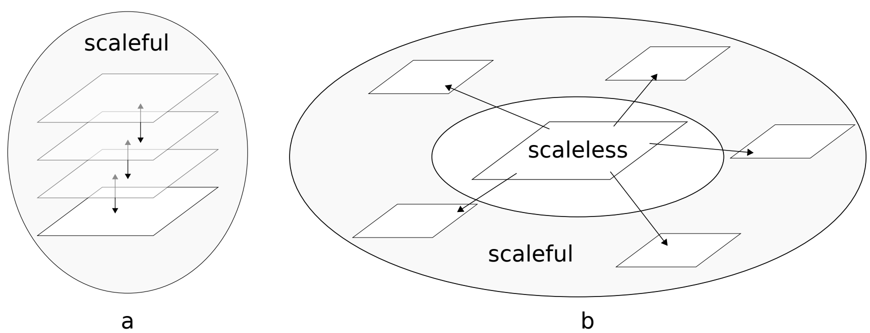

Fig. 2. Two types of scaleable database architecture: (a) Typical hierarchical MRDB structure (b) centralised structure with scaleless main database

Usually, when dealing with geographic databases, the complete database depends on exact measurements concerning positional values and attribution. No matter how any kind of topology is defined, it boils down to every object having some kind of coordinative reference - at least when using information about position or shape. Circumventing this necessity, it should be possible to use a data model which allows storing geographic information of different granularity in one single database. Topological constraints ensure the consistency of data. Together with stored data they form an ontology. So, the storage and use of topological information is of importance (see figure 3).

Interesting examples of topological data structures are given by e.g. the US Geological Survey. In their data standard for the Digital Line Graph (Enhanced) ([1],[2]) a topological data structure is described. While this is very simple, it could provide a starting point for a more advanced topological ontology.

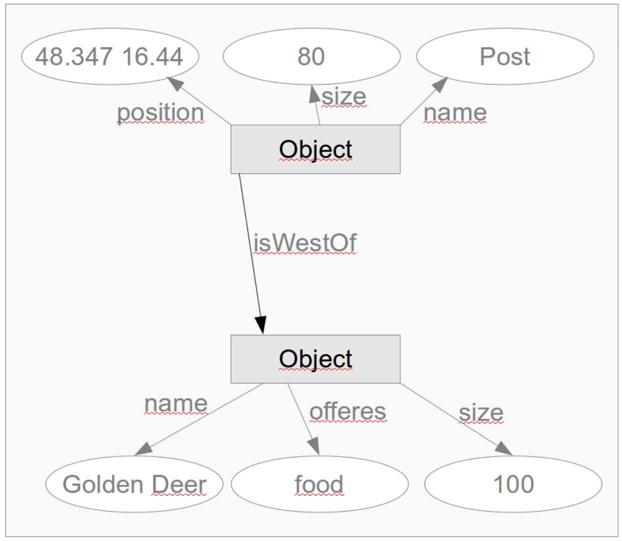

Fig. 3. Possible layout of a simple scaleless data structure

Another important condition for a scaleless database is to avoid typification as much as possible. This includes typification of geometry as well as of category of the entity. The reason for this is to offer an unreserved foundation for later evaluation. Furthermore, this allows mixing geometries: A certain point-of-interest may be represented by any kind of geometry. When for example an object is given by a polygon but the query performed on the database requests a scale which would represent this information as a point, a centerpoint can be calculated. Geometrical information may be abandoned completely when at least some geometric reference points are existing. Geographic information about the object may be calculated from its topological relations to other objects. The avoidance of categorising objects allows to dynamically perform operations on the data. This can be achieved by using a tagging-based attribution schema similar to the one used by OpenStreetMap [6]. By this, when e.g. generalising a line network, objects of similar properties can be treated similarly. Depending on the needs of the query, categories can be defined on-the-fly. The level of (attributive) detail of single database entries can be ignored since there is no fixed minimum information defined.

4 Arising Questions

From this vague concept of a scaleless data storage different questions arise:

– Is there a topology/ontology/reasoning mechanism which is able to cope with non or loosely standardised geometric and attributive information?
– Is the definition of an object-granularity (features, entities, objects, components, ...) necessary?
– How to integrate geometric information as an option?
– Can typification of the type of entities (river, street, ...) be avoided?
– Can typification of the type of geometry (point, line, ...) be avoided?
– Is any form of scaleless *access* possible/practicable?

A concrete question of research has yet to be found, but the most promising directions are the issues of typification and the definition of a scaleless data structure.

References

1. W.R. Allder, A.A. Elassal, and Geological Survey (U.S.). *Digital line graphs from 1:24,000-scale maps*. U.S. Geological Survey circular. U.S. Dept. of the Interior, Geological Survey, 1984.
2. Kenneth J. Boyko, Michael A. Domaratz, Fegeas Robin G., Hedy J. Rossmeissl, and E. Lynn Usery. An enhanced digital line graph design. *USGS Numbered Series*, 1048, 1990.
3. M. Hampe. *Integration einer multiskaligen Datenbank in eine Webservice-Architektur*. PhD thesis, 2007.
4. Mark Hampe, Karl H. Anders, and Monika Sester. Mrdb applications for data revision and realtime generalisation. In *Proceedings of the 21st International Cartographic Conference (ICC)*, pages 192–202, Durban, South Africa, 2003.
5. Mark Hampe, Monika Sester, and Lars Harrie. Multiple representation databases to support visualization on mobile devices. In *ISPRS Congrese*, volume XXXV of *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 135–140. Lund University, Dept of Physical Geography and Ecosystem Science, 2004.
6. F. Ramm, J. Topf, and S. Chilton. *OpenStreetMap: Using and Enhancing the Free Map of the World*. Uit Cambridge Limited, 2010.
7. Peter Van Oosterom. The reactive-tree: A storage structure for a seamless, scaleless geographic database. *Auto-Carto 10: Technical Papers of the 1991 ACSM-ASPRS Annual Convention*, (6):393–407, 1991.
8. Maarten Vermeij, Peter Van Oosterom, Wilko Quak, and Theo Tijssen. Storing and using scale-less topological data efficiently in a client-server dbms environment. *7th International conference on GeoComputation*, 11 2003.