CARTOGRAPHIC GENERALIZATION: INTERFACE ISSUES

Clodoveu A. Davis Jr.
Centro de Desenvolvimento e Estudos, Empresa de Informática e Informação do Município de Belo Horizonte – PRODABEL, Brazil

Keywords: Automated cartography, Cartographic generalization, Geographic data modeling, Geographic information systems, Multiple representations, Multiple presentations

Contents

Summary

The creation of series of maps using conventional cartographic techniques defines a hierarchy in which less detailed, smaller scale maps are produced from more detailed, large scale ones. This process is usually handled by experienced cartographers, who use both their technical skills and aesthetics sense to decide which features are kept in the transition, which are discarded, and which are to be modified in order to keep the map readable and visually pleasing. There are efforts towards replicating this process, called cartographic generalization, in a computer environment, using a Geographic Information System (GIS) and based on a detailed geographic database. This article shows how this can be done, and describes the operations that must be implemented in order to accomplish such a task, and illustrating the process with an example. There are issues regarding the implementation of operations that can replace the cartographer's

aesthetic sense, but it is clear that the geographic database creation process has to be aware of the need for such artistic improvements of the maps, while striving to generate data that are useful for the widest possible array of users and to avoid redundant efforts. The ideal is to try to create very generic spatial databases from which, using algorithmic transformations such as the ones presented in the article and possibly using aesthetic criteria from cartographic practice, maps and other kinds of visual products can be obtained. The incorporation of these aesthetic criteria can be performed by introducing cartographers as a part of the process, making their work more efficient through the use of adequate technology.

1. Introduction

In conventional cartography, the region of interest is frequently divided into coverage areas, or map sheets. These are again divided, so as to produce more detailed maps, in a larger scale, thus generating a series of articulate maps. Each of the levels of such hierarchy establishes a new appearance for the objects that compose the map, trying to filter out excessive detail, and to maintain a constant density of information. In any given level, each mapped phenomenon has a single graphic aspect, chosen as a function of the map’s scale and intended use. The symbols used for presenting the various objects or phenomena covered by the map are usually exemplified in a legend. Depending on the intended use, on the space available for presenting data, and on the density of printed information, the cartographer selects, among features of the same nature, which ones will appear on the map, and how the should be presented, leaving behind those considered less important.

This process is applied both to the creation of a map from the physical reality and to the transformation of a more detailed map into a less detailed one. A problem with this process is that this kind of work is usually performed manually, based on the cartographer’s common sense and empirical knowledge. The decisions for the creation of a map are taken based both on well-known and extensively documented cartographic techniques, and on the cartographer’s expertise, including his/her aesthetic sense. This process is known as *cartographic generalization*, and its incorporation to geographic information systems (GIS) and digital cartography is proving to be rather complex.

Observe that the use of the term *generalization*, in a cartographic context, has a meaning that is similar, albeit distinct, to the usual database technology concept. In both cases, the term is used in a context that intends to reduce the complexity of information. In databases, generalization means *abstraction* of information, the suppression of detail in order to give the data a broader meaning. In cartography, the interest lies in the suppression of unnecessary detail, to produce a new, less detailed, version of a map.

However, the construction of geographic databases cannot be restricted to the cartographic paradigms, since the demand for georeferenced information is getting broader and more complex (see *Advanced Geographic Information Systems*). Typical cartographic limits are no longer universally acceptable. The division of an interest area into map sheets, for instance, is not desirable in a geographic database, since it is necessary to retain the capacity for querying, viewing, or producing maps covering any region, regardless of the sheet boundaries. High costs usually associated with data conversion efforts also contribute for the development of initiatives that intend to broaden the use of geographic databases, by stimulating data sharing among applications – some of which may even be interested in producing cartographic documents. Therefore, it is important to be able to work on an interest area with varying levels of detail, and also to be able to count on tools that enable viewing data in the most appropriate way, considering the application's purposes. At the same time, it is necessary to manage the geographic database using a consistent, efficient, and safe database management system (see *Spatio-temporal Information Systems*).

These issues and others are at the foundation of a major paradigm shift, that is taking place nowadays: the evolution from paper cartography to geographic databases, from which cartographic documents and many other information products can be generated. The driving power behind this paradigm shift is the increased efficiency and economy that comes from collecting data with multiple applications in mind, and organizing geographic databases considering a wide range of possibilities in terms of analyzing, visualizing, comparing, updating, and distributing information, in an adequate representation for the purposes of each different application. New techniques that allow us to interact dynamically with the geographic database also pay a part in this evolution, along with techniques to encode and store temporal variables in a meaningful and useful way. For an overview on interface problems and cartography, see *Interacting with GIS - from Paper Cartography to Virtual Environments.*

This article will present the alternatives currently available for the representation of geographic information in a computer-based system such as a geographic information system (GIS), indicating the tools that are required to adequately achieve multiple uses of the same data, including cartographic product generation included. The approach adopted starts by formally defining representations and presentations (Section 2), and by presenting a comprehensive set of representation alternatives (Section 2.2). Then, the possibilities for transformation between representations and for generating presentations from representations are explored, by presenting a set of well-known operators from the fields of automated cartography, computational geometry, and spatial analysis (Section 3). These operators are analyzed and organized for the generation of new representations or presentations from primary representations (Section 4). Finally, the paper discusses the need for operators that can fulfill, at least partially, the role of the cartographer in some aspects of the generation of cartographic documents from a geographic database (Section 5).

2. Representation and Presentation

It has been said that geographic information, if appropriately managed and distributed, can be much richer than cartographic information, unless it is only used to imitate conventional maps. Nowadays, users want their GIS to be able to manage a multiple-usage database, from which it must be possible to produce at least the usual presentation alternatives employed by cartography, with varying symbology and adequate information density. The GIS is also required to have a continuously varying range of scales, and must not restrict information use to a predetermined set of standard scales, like in paper cartography. Furthermore, the visual aspect of the geographic objects must be adequate to the application's needs, on the screen or other media, regardless of whether or not it will be used to generate paper maps. In order to achieve that, GIS needs to (1) be able to maintain *multiple representations* of the same georeferenced entity or phenomenon, and (2) be able to produce several different *presentations* from a given representation.

This article uses the term *representation* in reference to the way the characteristics of spatial objects, including their geometric shape, are managed by a computer database, and *presentation* in reference to the visualization or graphical aspect of the spatial objects, on the screen or on paper. The concept of representation is therefore related to the notions of resolution, spatial dimension, precision/accuracy, detail level, and geometric or topological behavior. On the other hand, the concept of presentation is related to the display and/or output of the object as it must be visualized by its users, therefore involving parameters such as color, line type, and fill pattern.

The distinction between representation and presentation is crucial for the understanding that maps are, ultimately, becoming one of the many forms of output of the contents of a geographic database. In order to ensure that the same database serves many other purposes, besides mapping, its structure must be carefully modeled, using adequate techniques. It is also necessary to count on features of the geographic database management systems that allow the implementation of multiple representations, and on interactive features of GIS to allow the implementation of multiple presentations for each representation.

2.1 Multiple Representations and Multiple Presentations

The need for multiple representation features in GIS and geographic databases can be explained in terms of two main demands. The first is the need to deal with phenomena whose representation varies with scale, eliminating excessive detail and simplifying the visual appearance and controlling the object density, in order to benefit human interpretation and analysis skills. The second is the need to accommodate different perceptions of the same real world phenomenon, regardless of whether it is a physical or socioeconomic phenomenon.

The first demand is the general object of the field of cartographic generalization, traditionally solved through the intervention of a human cartographer who, using both technical and aesthetic criteria, decides which elements must be included on the map, which must be suppressed, and which should be simplified, considering a specific scale and a specific set of uses for the map. A complete automation of such a process has not been achieved yet, and there are many doubts on whether it is at all possible to do so, because of the difficulty involved in formulating clear rules that can be implemented in a computer for automated generalization, based on the cartographer's expertise.

The second demand correspond to the need to integrate the requirements of each geographic application, as expressed in their conceptual schemas. It is also a complex problem, considering the possibility of significant differences between the perception of a phenomenon by two or more distinct groups of users. In this context, the solution is carried out through a *geographic data modeling* standpoint, and it is necessary to count on a model that allows the specification of spatial aspects of the information –

according to multiple user views – and, at the same time, allows the specification of all required visualization alternatives, as demanded by the applications. (For a discussion on modeling, see *Conceptual Modeling of Geographic Applications).*

Both demands can be integrated when the problem is analyzed from the database creation point of view. Since in this approach the map is understood as an output obtained from a geographic database, many cartographic generalization operations can be implemented as special transformation procedures, for which the input is a given detailed representation of a real-world phenomenon, and the output is a simpler, less detailed representation. The actual visualization or plotting of a given representation is carried out using another kind of transformation, this time from a representation to a presentation.

Of course, automated procedures are harder to achieve in the case of cartographic generalization operations of a more artistic nature, which will probably require the manual intervention of an experienced cartographer. Nevertheless, there are attempts in that direction, requiring the compilation of a set of formal “knowledge rules” from the cartographer’s practice. But since such developments have not reached commercial GIS products, this paper will from here on concentrate on well-known procedures and algorithms that can be used for the implementation of transformations between representations or transformations from representation to presentation. These usually are studied in the fields of automated cartography, computational geometry, and spatial analysis. But first, a brief introduction of the usual geographic representation alternatives is necessary.

2.2 Representation Alternatives

Choosing a representation alternative for a real-world element involves, in general, the definition of the parameters for the *discretization* of its geometric shape, making it simple enough for a computerized system to handle. This discretization is carried out according to the nature of the observed phenomena. Individual elements, such as rivers, trees or property parcels are more easily represented using simple data structures, called *geo-objects*, which intend to reproduce the most important aspects of its geometric shape. Elements which correspond to continuously varying phenomena, or *geo-fields*, are understood from a set of samples, obtained at sites whose quantity and spatial distribution are chosen according to the nature of the data collection process.

Regardless of the choice between geo-objects or geo-fields, the complexity and the nature of the geometric representation are fundamentally dependent of the intended use, as defined by the application. For instance, an application that needs to record the location of gas stations for marketing purposes might only need to represent each station as a point, therefore using only one pair of geographic coordinates. On the other hand. An application that sees gas stations as an environmental threat for the underground water supply needs to have much more detail than that when representing them, including the location of underground reservoirs, the soil type and the relief of the region. The representation of the relief, for this application, needs to be very detailed, while a much coarser representation of the relief of the same region is sufficient for an erosion potential application. (For an additional discussion on objects and fields, see *Detail Filtering in Geographic Information Visualization.)*

There are six alternative representations for geo-objects (*point*, *line*, *polygon*, *network node*, *unidirectional arc*, and *bidirectional arc*), and five for geo-fields (*isolines*, *samples*, *tesselation*, *planar subdivision*, and *triangulated irregular network*). Table 1 presents the formal definitions for each alternative.

Representation	**Definition**
Point	A *point* is an ordered pair (x, y) of spatial coordinates.
Line	Let $v_0, v_1, \ldots, v_{n-1}$ be n points on the plane. Let $s_0 = \overline{v_0 v_1}, s_1 = \overline{v_1 v_2}, \ldots, s_{n-2} = \overline{v_{n-2} v_{n-1}}$ be a sequence of n - 1 segments, connecting those points. These segments form a polygonal *line* L if, and only if, (1) the intersection of consecutive segments is only the endpoint shared by them (i.e., $s_i \cap s_{i+1} = v_{i+1}$), (2) non-consecutive segments do not intercept (i.e., $s_i \cap s_j = \varnothing$ for all *i*, *j* such that $j \neq i+1$), and (3) $v_0 \neq v_{n-1}$, i.e., the polygonal line is not closed.
Polygon	A *polygon* is the region of the plane limited by a closed polygonal line.
Network node **Unidirectional arc** **Bidirectional arc**	A network is a structure formed by a set of *nodes* N = {n_1, n_2, ..., n_m} and a set of arcs A = {(*i*, *j*), (*k*, *l*), ..., (*s*, *t*)}, where each arc (*i*, *j*) connects pairs of nodes belonging to N (therefore $i, j \in$ N). Zero or more arcs are connected to a node, and each arc is connected to exactly two nodes. If the pair of nodes to which the arc is connected is an ordered pair, *i.e.*, the direction of the arc is relevant, it is called an *unidirectional arc*. If it is not, the arc is called a *bidirectional arc*.
Isolines	*Isolines* are polygonal lines that represent the intersection between a surface and a set of given planes, parallel to the XY plane, and conveniently spaced, usually at regular intervals.
Tesselation	A regular *tesselation* is a division of the space into cells of uniform size and shape, so that to every point of the modeled space there is only one corresponding cell.
Samples	*Samples* are sets of points to which the value of a field is associated. Supposedly the quantity and the distribution of the samples is good enough to consider them representative values of the field, that is, it is possible to determine the value of the field anywhere in the plane using an interpolation procedure between samples.
Planar subdivision	A graph *G* = {*N*, *A*}, composed of a set *N* of nodes and a set *A* of arcs, is said to be planar if it can be inserted on the plane without crossing any arcs. This insertion forms a partition of the plane called *planar subdivision*, in which any point of the plane either (1) coincides with a node, (2) belongs to an arc but does not coincide with any nodes, or (3) belongs to the interior of a single subdivision.
Triangulated	A *triangulated irregular network* (TIN) is a planar subdivision

irregular network (TIN)	structure in which the arcs are line segments and the polygons are always triangular.

Table 1. Representation alternatives

Several observations must be made concerning the set of representations presented in Table 1. First, there is a subtle, though important, distinction between point, network node and sample. Points are used to represent individual entities, for which there is no need for a more detailed geometric depiction. One example is the representation of cities in a state-wide perspective: each city is an individual entity, for which no additional geometric details beyond its geographic location are needed. If there are any values corresponding to a continuously varying phenomenon associated to all points in a set, then it must be considered a set of samples, and not simply a group of points. Likewise, if a set of points functions as nodes in a network, then it must be modeled as a set of network nodes.

A similar observation must be made concerning lines, isolines, and network arcs. Even though the geometric shape is similar, the role of the shape in the representation of a real-world phenomenon implies a different meaning for each of these alternatives. The same can be said about polygons, planar subdivisions, and TINs: even though all essentially represent polygons, the latter two cases carry the meaning associated with the notion of fields, mainly what is known as the *planar enforcement rule*: every point of the plane corresponds to a single specific value of the field. This can usually be estimated by interpolating among nearby points (samples, isoline vertices, triangle vertices, cells) for which the value of the field is known. Individual polygons, such as those that are used to represent blocks in a city, do not fulfill this rule, and are therefore used to represent geographic objects.

A separate representation for bidirectional arcs is not mandatory: it would suffice to admit that, whenever bidirectional arcs are needed, two unidirectional arcs in opposed directions can be used. However, the distinction between these two representation alternatives helps to bring the database’s conceptual schema closer to an important set of geographic applications. This option also helps to avoid redundancies and to reduce the number of arcs required to represent networks that are essentially bidirectional, such as the ones used in power and telecommunications infrastructure.

The most usual type of cell used in a tesselation is square, and in this case the tesselation corresponds to a *digital image* or a *grid*. As opposed to digital images, grids are usually employed for the representation of surfaces, while images are more common in the representation of a large volume of regularly spaced samples, such as the ones obtained by remote sensors or scanners. Notice that cells do not have an individual meaning when separated from the tesselation itself.

These eleven representation alternatives cover the majority of what’s available both in current GIS products and in geographic data models. The most important alternatives included in the literature, but not covered here, include three-dimensional objects other than surfaces, which are quite rare in current GIS, and complex objects, recursively composed of parts which, ultimately, belong to one of the classes presented here.

3. Transformations

This section presents a selection of transformation algorithms, generically called here *operators*, each of which has been presented in the literature aiming at specific applications in the fields of computational geometry, automated cartography, and spatial analysis. Each algorithm can be used to produce a transformation in the representation of a geographic object or to produce a presentation for a given representation. More complex transformations can be achieved by applying the algorithms sequentially.

-
-
-

Bibliography

Bailey, T. C., Gatrell, A. C. (1995) *Interactive Spatial Data Analysis*, 413pp. London: Longman. [A textbook on spatial analysis.]

Brassel, K. E., Weibel, R. (1988) A Review and Framework of Automated Map Generalization. *International Journal of Geographical Information Systems*, 2(3): 229-244. [An early paper on map generalization, at the edge of the development of computer-based generalization tools.]

Buttenfield, B. P. (1993) *Multiple Representations*. Closing Report for the National Center for Geographic Information and Analysis Research Initiative 3, State University of New York, Buffalo. [This text contains the results of a research initiative by the NCGIA on multiple representations, at a time when commercial GIS products were at their infancy.]

Edelsbrunner, H. (1987) *Algorithms in Combinatorial Geometry*, 423 pp. Berlin: Springer-Verlag. [One of the most important textbooks on computational geometry, including the analysis of several of the algorithms listed in the article.]

Lagrange, J. P., Ruas, A. (1994) Geographic Information Modelling: GIS and Generalization. In *Proceedings of the Sixth International Symposium on Spatial Data Handling*, 2: 1099-1117, Edinburgh. [A seminal text on generalization using GIS.]

McMaster, R. B., Shea, K. S. (1992) *Generalization in Digital Cartography*, 134 pp. Washington, DC: Association of American Geographers. [A comprehensive book on map generalization operations.]

Müller, J. C., Lagrange, J. P., Weibel, R. (editors) (1995) *GIS and Generalization: Methodology and Practice*, 257 pp. London: Taylor & Francis. [A collection of texts on map generalization considering the potential use of GIS or GIS-based tools.]

O'Rourke, J. (1994) *Computational Geometry in C*, 346 pp. Cambridge: Cambridge University Press. [A textbook on computational geometry, contains practical descriptions of several important algorithms that are used in cartographic generalization.]

Preparata, F. P., Shamos, M. I. (1988) *Computational Geometry: an Introduction*, 398 pp. New York: Springer-Verlag. [One of the most important textbooks on computational geometry, including the analysis of several of the algorithms listed in the article.]

Biographical Sketch

Clodoveu Augusto Davis Jr. graduated in Civil Engineering in 1985 from Universidade Federal de Minas Gerais (UFMG). He also has M.Sc. and Ph.D. degrees in Computer Science, also from UFMG, in 1992 and 2000, respectively. He led the team at Prodabel that conducted the implementation of GIS technology in the city of Belo Horizonte, Brazil, and coordinated several geographic application development efforts. Currently, he works at Belo Horizonte's Municipal Information Technology Company (Prodabel), coordinating research and development efforts at the company's Development and Studies Center. He is also the scientific editor of *Informatica Publica*, a Brazilian journal on information technology for the public sector (www.ip.pbh.gov.br). His current research projects range from spatial database modeling to GIS interoperability, with a focus on the practical applications of such technology for Brazilian municipal administrations. This includes his work as the coordinator of the geoprocessing group in Belo Horizonte's metropolitan high-speed networking consortium, a local Internet-2 research initiative funded by Brazilian federal government institutions. On the education side, he lectures as a visiting professor for UFMG's graduate course on geoprocessing, and at the graduate course on information technology for the public sector, organized and implemented by Prodabel with the support of Pontifícia Universidade Católica de Minas Gerais (PUC-MG). Much of his published work, including academic papers on geographic information technology, his bimonthly column on *InfoGeo* (a Brazilian geomatics magazine), and many articles on urban applications of GIS can be found at Prodabel's web site (www.pbh.gov.br/prodabel/cde) and at his personal web page (www.cdavis.hpg.com.br). His main interests are urban GIS, geographic databases, map generalization, and multiple representations in GIS.