TOWARD A PRACTICABLE MODEL OF CARTOGRAPHIC GENERALISATION

Mark Monmonier

Department of Geography
Syracuse University
Syracuse, New York 13244-1160
U.S.A.

ABSTRACT

A system for fully automatic cartographic generalisation must be able to deal with severe reductions in map scale by preserving the salient characteristics of geographic features while avoiding overlapping symbols. An "intelligent database" treating features as assemblages of landscape units can further reduce the complexity of the generalisation algorithm. A hybrid data structure is needed to allow the algorithm to track features and detect overlap. Skeletal representations of key features can guide the selection of features and displacement of symbols. A supervised approach, similar to supervised classifiers for remotely sensed imagery, affords a practicable means of choosing options and setting tolerances.

INTRODUCTION

Attempts to evaluate the cost effectiveness of programs for collecting digital cartographic data must recognize two levels of use. Primary applications are those uses the program was designed to serve. Because primary uses commonly relate to the mission of the agency supporting the project, their benefits generally are comparatively easy to assess. Secondary applications are less predictable, and often serendipitous. But although their benefits are difficult to define and measure, in the long run secondary uses might well yield a greater return on investment than primary applications. Included in the broad range of secondary uses are applications in which the data are related to other thematic coverages, are used at some future date as a base for generating change data, and are generalised to yield smaller-scale displays and other less detailed derivative products.

This paper addresses the use of computer-assisted cartographic generalisation to extend the utility of digital geographic data. It presents and explores several principles essential to the design of an operational system for cartographic line generalisation. And it recommends strategies in feature-coding to extend the applications of digital data to a wider range of uses at smaller scales.

HYBRID DATA STRUCTURE NEEDED

Perhaps the most obvious requirement for an operational generalisation model is a hybrid, vector-and-raster data structure. Although point-symbol data can be generalised in vector mode and areal data in raster mode (Monmonier 1983), linear features would require elements of a raster structure to represent their proximity to neighboring features. An operational model must recognize that generalisation involves three processes: selection, smoothing, and displacement.

Vector-only approaches, such as the widely used Douglas algorithm (Douglas and Peucker 1973), are successful only where scale reduction is comparatively minor and the density of features is sufficiently low so that overlap--or worse, criss-crossing--tends not to occur. These algorithms address only the need for feature smoothing, and then only by a process of point removal (Van Horn 1985). It is perhaps symptomatic of the operational limitations of vector-only algorithms for line generalisation that in demonstrations and evaluations, such as the exhaustive comparative study by McMaster (1983), these algorithms are applied only to isolated strings of points unencumbered by encroaching neighbors. Yet, solutions acceptable for well-spaced features may prove inappropriate for neighboring features, such as parallel railways, roads, streams, and contours sharing a steep-sided valley or water gap, as shown in Figure 1.

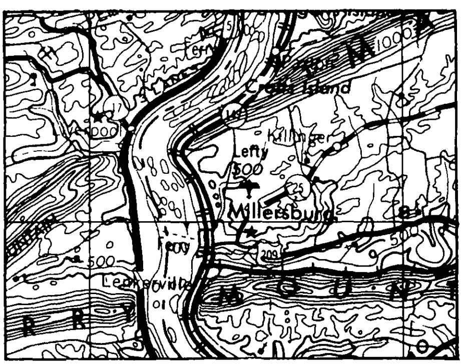

FIGURE 1--Crowded linear features, such as the river, roads, and railway sharing this watergap, require separation if overlap of symbols is to be avoided. [From Harrisburg, Pa., U.S. Geological Survey 1:250,000-scale topographic map, 1969.]

In contrast, efforts at generalising lines in raster mode can yield equally embarrassing results. In an investigation of the potential of raster data for shifting apart close features that might otherwise overlap, Wilson (1981) demonstrated that a high local feature density can produce unwarranted wrinkles in an otherwise smooth feature, as in the generalised map on the right side of Figure 2. This example includes only the boundaries of New York State (with Long Island excluded) and the Adirondack Forest Preserve, the state's principal railway lines, and the New York State Thruway from Buffalo to New York. Note the indentation in the east-central part of the region, where the mutual displacement of the road and railway produced an inward dent in the Forest Preserve boundary. Iterative raster-mode algorithms that examine only a small part of the map at once cannot easily preserve salient trends in linear features.

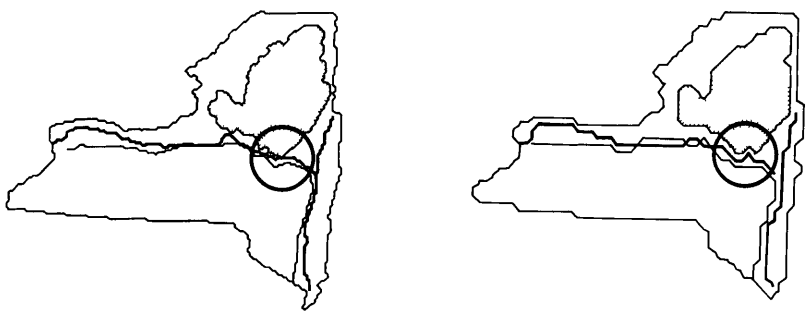

FIGURE 2--Original (left) and generalised (right) versions of raster-mode representations of selected linear features in New York State. Circle identifies dent in Forest Preserve boundary produced by mutual displacement of railway (heavy line) and road. [Redrafted from Figures 4.30 and 4.32 in Wilson (1981).]

Vector-mode displacement of cartographic features has been partly automated but seems doomed to the interactive guidance of a human cartographer. Most of this effort was carried out in the 1970s in West Germany, largely in the Department of Topography and Cartography at the University of Hanover (Lichtner 1979). Efforts have focussed on the generalisation of topographic maps, with scale reduced by factors of 1/2 to 1/4. Advances have included a series of algorithms for displacing a stream from a road, a road from a triangulation station, and a road from itself on a double-hairpin bend (Schittenhelm 1976). Where the scale reduction is moderate, say, from 1:50,000 to 1:200,000, a human cartographer enters displacement vectors to direct the shifting apart of parallel linear symbols (Christ 1976). Where the scale reduction is comparatively minor, as when the street grid on a 1:25,000-scale topographic map is generalised for portrayal at 1:50,000, apparently only interactive checking and occasional minor adjustments are needed (Leberl, Olson and Lichtner 1986). For more severe scale reductions, though, considerable human intervention would be needed to help these algorithms cope with densely packed features.

Generalization for severe scale reduction has been programmed in raster-mode for the less demanding case of land-cover polygons, reduced in detail from 1:250,000-scale digital data for presentation at 1:2,000,000 (Monmonier 1983). As shown in Figure 3, a raster-mode algorithm can cope with the need to remove clutter by eliminating some features and merging or thickening others. But land-cover polygons can be divided, whereas roads and railways require continuous symbols of uniform length, and must either be displaced to avoid overlap or eliminated altogether. Raster data can afford a convenient means of identifying contested parts of the graphic plane but cannot efficiently and independently deal with all pixels representing a long, thin, uniformly wide polygon.

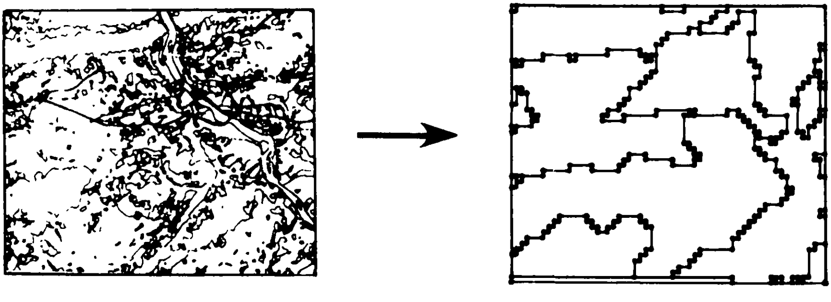

FIGURE 3--A raster-mode generalisation algorithm can remove the clutter inherent in the boundaries of land-cover polygons reduced from 1:250,000 to 1:2,000,000 (left). Elimination of thin, graphically unstable features and merger of nearby polygons with the same land cover yields a far less cluttered set of boundaries (right). [From Figures 13 and 18 in Monmonier (1983).]

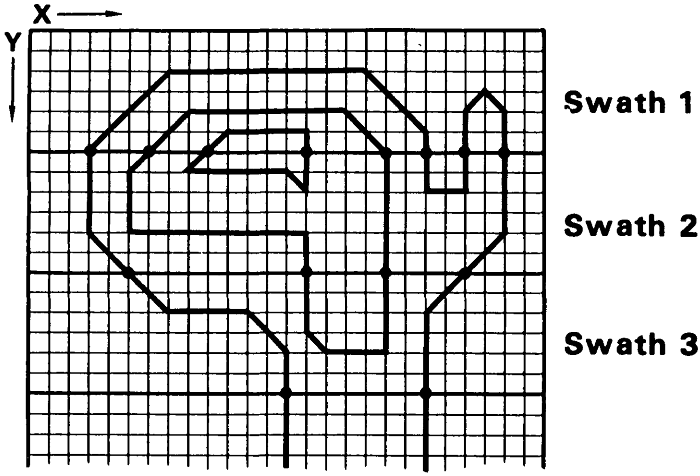

FIGURE 4--Elements of the vaster data structure, a hybrid combining a raster organisation of horizontal swaths aligned with the rows of a grid and a vector organisation of chains representing sections of each linear feature falling within the swath. [Adapted from Figures 2a and 2b in Peuquet (1983).]

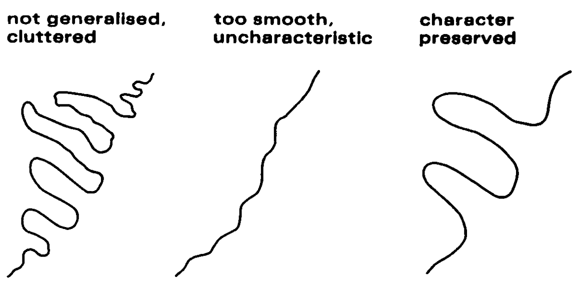

FIGURE 5--A severe change in scale might warrant the selective elimination of some features of a given type and the enlargement of others. For example, a reduction in the number of inlets not only prevents clutter but preserves the character of a fjorded coast. Merely smoothing the coastline yields a symbol with less information about the form of the coast and its geomorphic history.

A hybrid data structure, such as the "vaster" structure examined by Peuquet (1983), might promote the efficient identification and resolution of overlapping symbols. The vaster structure divides the raster grid into a number of swaths, several rows wide (Figure 4). Vector chains represent portions of linear features falling within the individual swaths, and the initial nodes of these chains are recorded as scan-line data. Chains can be followed easily between neighboring swaths, and adjoining chains can be detected easily within a swath. For feature displacement, though, elongation of the swaths in one direction (as in Figure 4) seems inefficient--displacing horizontal features upward or downward in horizontal swaths is inherently simpler than displacing vertical symbols in the same data structure. A more appropriate hybrid structure might subdivide the grid into square units, instead of long, thin swaths. Perhaps a vector-quadtree hybrid--call it a "vectree" structure--is warranted.

LANDSCAPE UNITS IN AN "INTELLIGENT DATABASE"

What might be called intelligent databases will greatly simplify cartographic generalisation. By "intelligent" I mean that the data would specify a feature's membership in a landscape category relevant to scale reduction. Perhaps the classic example of where this might be appropriate is a fjorded coastline. Points around each embayment should be similarly tagged, as should headland points between fjords. With these data a generalisation algorithm might efficiently preserve the geographic flavour of the coast at smaller scales by retaining and widening some fjords while eliminating others. Figure 5 illustrates how a cartographer trained in geomorphology would both promote graphic clarity and retain salient physiographic traits. We should expect no less from an operational model. Yet common sense suggests that much of the requisite intelligence might more efficiently be added to the data than to the algorithm.

An intelligent database might represent useful linkages between social and economic elements of the landscape. These linkages, which often are implicit in digitally encoded data on land use, land ownership, and administrative areas, are just as frequently missing from electronic representations of the street and road network--to mention but one common type of digital map. In the United States, for instance, route numbering of state highways is often the vestige of a haphazard, incremental strategy of roadbuilding, and individually numbered routes need not reflect the principal lines of travel between interacting cities along the same route. In contrast, the newer and more heavily used Interstate network would almost always warrant selection whenever roads are to be shown: its corridors are comparatively smooth and direct, and its links are usually the principal intercity routes. Moreover, its numbering scheme differentiates intercity routes from beltways and spurs, which might need to be suppressed for very-small-scale displays.

One general-purpose geographic database that would benefit from the identification of landscape units is the U.S. Geological Survey's digital line graph (DLG) data (Allder and others 1983). These quadrangle-format files of vector data represent hydrographic, boundary, transportation, and other "culture" features shown on large-scale topographic maps. Linear data for streams indicate flow direction and braided or artificially confined channels. But addition of simple stream-order numbers indicating branching structure and network geometry (Richards 1982) would be highly useful in identifying minor tributaries, which might readily be eliminated for clarity when scale is reduced.

Overlaying DLG data and the Geological Survey's land-use and land-cover data might permit a rule-based generalisation algorithm to discern the relative importance of some linear features. For example, short roads solely in residential areas are likely to be side-streets, whereas longer roads through commercial and industrial areas might well be thoroughfares. But such assumptions are problematic and far from foolproof. Specifying this kind of intelligence directly in the database would benefit cartographic generalisation as well as applications in transportation planning, computer-assisted highway navigation, and real-property assessment, among others.

"EXPERT GUIDANCE" FOR A RULE-BASED SYSTEM

Relative importance can be specified for individual points as well as for entire features. SAS/GRAPH, an American software package for data graphics and simple statistical maps, uses this approach to guide the generalisation of its county, state, and country boundaries--each point has a "density level" derived by applying the Douglas point-elimination algorithm five times, with the tolerance incremented in steps to yield five progressively more generalised sets of points (Carter 1984). A point-tagging strategy seems less suitable, though, for a large cartographic database with many features, some of which will have to be not only smoothed but displaced.

A plausible solution might be to provide a highly generalised version of the map at a greatly reduced scale. Designed by an experienced cartographer or geographer, this skeletal set of carefully selected, smoothed and displaced features could then guide the computer-assisted generalisation of the data for a range of intermediate scales. Deveau (1985) employed a similar philosophy in generalising coastlines. He controlled the degree of generalisation by finding first a "smallest sufficient subset" of points to express a feature's salient character and then eliminating details falling within the tolerance used to specify the degree of generalisation. A rule-based system might use a smallest sufficient subset of features and landscape units to guide the selection of features and to suggest directions for their displacement.

A "SUPERVISED-GENERALISER" APPROACH

An intelligent database and the guidance of a smallest sufficient subset of feature characteristics should simplify the design and implementation of an operational algorithm for fully-automatic generalisation. But experience with the comparatively simple raster-mode generalisation of land-cover data suggests that the user will need to select and set an unwieldy number of options and tolerances (Monmonier 1983). Particularly prominent among these choices are the relative priorities for selecting various types of geographic features, and for preserving their character and conserving positional accuracy. The map author must be cautious in setting specifications for the algorithm--at any given reduced scale markedly different generalisations might be produced from the same database. A "supervised" strategy, similar to those used to classify remotely sensed imagery, would be useful in allowing the cartographer to work out the choice of options and tolerances with "training data" for one or more known areas. Trial runs could be used to calibrate the algorithm to assure mapped patterns appropriate to project goals and presentation constraints. Once calibrated, the generalisation model could be applied to a much larger region.

An expert-systems methodology might also be useful. A computer could, no doubt, be programmed to examine a series of generalised maps prepared at various scales to meet specific presentation goals, and to establish a set of options, tolerances, and decision rules likely to yield suitable results for similar goals when applied to similar data. But this level of sophistication seems unwarranted because most, if not all, needs might be met with a far simpler "supervised-generaliser" approach based on good data. Intelligence in the data can obviate the need for highly elaborate algorithms.

REFERENCES

Allder, W. R. and others (1983) Digital line graph attribute coding standards. U.S. Geological Survey Circular, no. 895-G.

Carter, J. R. (1984) *Computer mapping: progress in the '80s*. Washington, D.C., Assn. of American Geographers, 56-57.

Christ, F. (1976) Fully automated and semi-automated interactive generalization, symbolization and light drawing of a small scale topographic map. *Nachrichten aus dem Karten- und Vermessungswesen*, Reihe 2, no. 33, 65-74.

Deveau, T. J. (1985) Reducing the number of points in a plane curve representation. *Proceedings of Auto-Carto 7: Digital Representations of Spatial Knowledge*, Washington, D.C., 152-160.

Douglas, D. H. and Peucker, T. K. (1973) Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *The Canadian Cartographer*, 10, no. 2, 112-122.

Leberl, F. W., Olson, D., and Lichtner, W. (1986) ASTRA--a system for automated scale transition. *Photogrammetric Engineering and Remote Sensing*, 52, 251-258.

Lichtner, W. (1979) Computer-assisted processes of cartographic generalization in topographic maps. *Geo-Processing*, 1, 183-199.

McMaster, R. B. (1983) A mathematical evaluation of simplification algorithms. *Proceedings of the Sixth International Symposium on Automated Cartography*, Ottawa, vol. 2, 267-276.

Monmonier, M. S. (1983) Raster-mode generalization for land use and land cover maps. *Cartographica*, 20, no. 4, 65-91.

Peuquet, D. J. (1983) A hybrid structure for the storage and manipulation of very large spatial data sets. *Computer Vision, Graphics, and Image Processing*, 24, 14-27.

Richards, K. (1982) Rivers: form and process in alluvial channels. London, Methuen, 32-39.

Schittenhelm, R. (1976) The problem of displacement in cartographic generalization attempting a computer-assisted solution. Nachrichten aus dem Karten- und Vermessungswesen, Reihe 2, no. 33, 65-74.

Van Horn, E. K. (1985) Generalizing cartographic data bases. Proceedings of Auto-Carto 7: Digital Representations of Spatial Knowledge, Washington, D.C., 532-540.

Wilson, S. (1981) Cartographic generalization of linear information in raster mode. Unpublished M.A. thesis, Syracuse University, U.S.A.