Appeared in **Cartouche, Number 39, Fall 2000.**

National Atlas of Canada Producing First Map Using Automated Generalisation of Framework Data.

Rupert Brooks
Brooks@nrcan.gc.ca
September 22, 2000

The National Atlas of Canada has a long tradition of making high quality reference maps of the country. Over the last two decades the organisation has switched from a traditional to a digital mapmaking environment where the majority of maps remain online and are never sent to press. It has been a challenge and source of pride to continue the paper reference map series using the most advanced technology, while maintaining or improving the aesthetic and technical standards of the past.

The entire GIS industry has gradually increased in sophistication since its inception. In the beginning stages, only the simplest spatial data structures, and corresponding similar products could be produced digitally. Improvements in technology have allowed for the replacement of manual processes by digital ones. This resulted in mapmaking being carried out digitally using a very similar production process to the traditional paper approach. Scribers were replaced with mice, screens replaced the drawing table. The sequence and structure remained similar. As a result, the data layers produced for mapping corresponded very closely to the scribed sheets that would have been produced in the past.

The next logical progression is the development of a geospatial data infrastructure. Cartographic visualisation is just one of many applications that the geospatial data infrastructure must support. This is leading to a complete change in mapmaking procedures. In the past, multiple cartographic base layers would be kept at a range of scales. Each such layer was a relatively simple group of features for a map, rendered in a style and a level of detail appropriate for the scale. Attributes, if present at all, were often solely related to the symbolisation of the feature. Name information was usually present as annotation near the feature on the map, but there was no explicit linking of the feature name to the feature record in the database. Data infrastructures, however, operate on the principle of collecting a minimal number of different datasets, as close as possible to their source. In addition to supporting mapping, the data is designed to answer analytical questions, and so contains a high degree of attribution and topological structure.

The National Atlas of Canada's contribution to the data infrastructure is a series of 1:1M framework data layers. These are being designed to contain considerable intelligence and topology. The hydrology layer has undergone a great amount of work in the past while. From the outset, the data was intended to support cartographic visualisation at a variety of scales. In fact, the need for automatic generalisation was one of the driving factors in the design of the dataset. Derived from the VMAP0, these layers are relatively detailed as compared to other National Atlas products[1].

Deriving cartographic products from a feature database is certainly not new. For example, for a long time now the Canadian topographic map series has been created from features in the National Topographic Data Base (NTDB) and Canadian aeronautical charts have been generated based on features extracted from the Canadian Aeronautical Charting database (CANAC). For

[1] Brooks, R. 1999. The New and Improved Base Framework for National Atlas Data. Proceedings of ICA '99. Also online at http://atlas.gc.ca/english/about_us/index_pres_e.html#framework

these products, the features in the database are a close approximation to the features that must appear on the printed map. The National Atlas of Canada required a significant scale transformation as part of the cartographic process, which added considerable challenge.

As the work on the hydrologic layer of the framework progressed, work had also begun on preparing a new reference map of Canada's Northern Territories. The existing map was drawn in 1974, with a partial names update in 1982 and some new information overprinted since then. While still popular, the map was now significantly out of date. The base material was never digital, and at a scale of 1:4M, it fell far enough between the usual Atlas scales of 1:2M and 1:7.5M that recompilation was essential. It was an ideal candidate to use for testing the theoretical automated generalisation. Instead of manual recompilation, the cartographic team would work with the «frameworks» team to create a hydrologic layer for this map. This layer would then be quality controlled and edited as part of the usual cartographic process.

As the process began in the fall of 1999, the team was faced with a number of challenges. The organisation ran on a base software of ESRI Arc/Info, but it had quickly become clear that no off the shelf tools within that software were adequate to handle the task. In addition, the base was not yet complete, and had to be fully structured and updated before this could be applied. While research examples of automated generalisation software existed, they had not proved robust enough in practice to be used for production. As the data was being structured and revised, generalisation software had to be rapidly prototyped, tested and brought into the production process.

Automated generalisation for map production can be divided into two generalisation processes – database generalisation (selection) and cartographic generalisation (rendering). The preferred application is to apply a selection process to the full database to yield the set of features to be present. These features are then rendered using cartographic generalisation operators to prepare the result for presentation on paper.

The selection process was based on algorithms developed by Richardson and Thomson[2], and extended in the course of this project by Thomson and Brooks[3]. Lake selection was guided by a measure of area and shape. River selection is based on the stream ordering of the river network. Previous efforts at using stream ordering for generalisation, when compared against manual efforts had achieved accuracies greater than 80%[4]. While impressive, even 20% remaining error

---

[2] Richardson, D. E. 1994. Generalization of spatial and thematic data using inheritance and classification and aggregation hierarchies. In: Waugh, T. C. and Healey, R. G. (Eds.), Advances in GIS Research 2, London: Taylor and Francis, 957-972.

Richardson, D. E., and Thomson, R. C. 1996. Integrating thematic, geometric and topological information in the generalization of road networks. Cartographica, 33(1), 75–83.

Thomson, R. C., and Richardson, D. E. 1999. The 'Good Continuation' Principle of Perceptual Organization applied to the Generalization of Road Networks, Proceedings of the ICA 19th International Cartographic Conference (Ottawa, 14-21 August), 1215–1223.

[3] Thomson, R. and Brooks, R. 2000. Efficient Generalisation and Abstration of Networks using Perceptual Grouping. Proceedings of GeoComputation 2000. University of Greenwich, UK.

[4] Rusak Mazur, E., and Castner, H. W. 1990. Horton's ordering scheme and the generalisation of river networks, The Cartographic Journal, 27, 104-112.

was too much for a production process. By developing techniques that could integrate supplemental information from the database the team was able to exploit the sophisticated attribution of the framework data to achieve a highly accurate result.

In addition to these primary rules, two rules were applied to guide the interaction of lakes and rivers during the generalisation. The outflow path from a lake to the ocean would be kept if the lake was kept. Conversely, small lakes at the very tip of river systems were preferentially kept if the river system was kept. Finally the process was tuned to provide slightly more features than were required, since it was far easier to manually remove a feature, than to add one.

There was very little experience in automatic cartographic generalisation within the organisation. It quickly became clear that off the shelf tools would not suffice to provide an aesthetically pleasing cartographic solution. Custom software was required to process the selected data for visual effect.

The expertise of the cartographic team was essential to this development. This author began with only a theoretical understanding of the problem and could not have come up with a successful solution without the patient assistance of the experienced cartographers. The software was rapidly prototyped by producing test generalisations which were then scrutinised and commented on by experienced cartographers. After approximately two months of development, and many such test cycles, the results were imperfect, but acceptable. The remaining problems could be fixed by an experienced cartographer in a reasonable amount of time.

The importance of the personal dynamic cannot be overemphasised. GIS advocates, computer scientists and engineers have claimed to be “on the verge” of producing automated generalisation techniques for nearly 40 years. However, the success stories have been few and far between. In practice, automated techniques have rarely demonstrated the capability that even a relatively inexperienced cartographer can muster. Conversely, cartography has gained a reputation among engineers for being subjective and very prone to the individual idiosyncrasies of the cartographer. We were fortunate that the project team was able to dispel both of these myths.

The generalisation procedures have been a success in practice. Nevertheless, there are many potential areas for improvement. Several systematic problems remain which must be corrected manually. Small spikes are frequently created by the line simplification techniques. Although not usually visible at the final scale, these are being corrected manually before printing. The system has also shown a strange preference for narrow lake features, which can sometimes create a “noodley” appearance. Again, these stylistic problems must still be corrected by hand.

Overall the experience was challenging but successful. It took longer to carry out the automated generalisation than it would have taken to manually generalise the data, but this time includes considerable development time. When the dead ends and learning experiences are factored out, the process is reasonably efficient. Ultimately, however, the real reason to seek procedures like this is to be able to invest effort in maintaining a single sophisticated database that supports many applications rather than multiple simplistic map layers. Indeed, over the course of the year, numerous other applications have surfaced.

The map *Canada: Yukon Territory, Northwest Territories and Nunavut* will be released by the end of this year. It will be the first map produced by the National Atlas of Canada that has used automated generalisation techniques for a significant portion of its production. Nevertheless, the entire team is pleased with the outcome and consider it to be on a par with any of the other maps produced here.

Appeared in **Cartouche, Number 39, Fall 2000.**

Figures 1-4 show the evolution of the hydrology and coastline on the south coast of Baffin Island through various stages of processing.

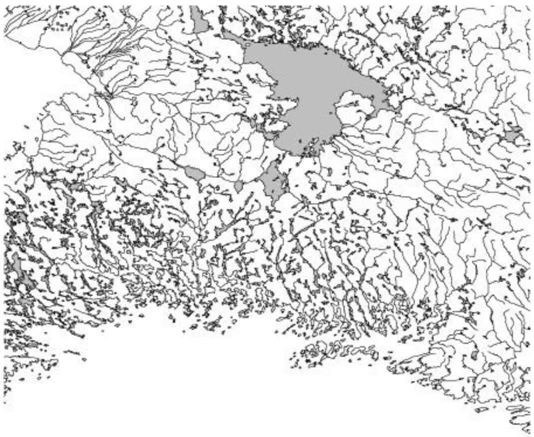

Figure 1: Raw 1:1M data

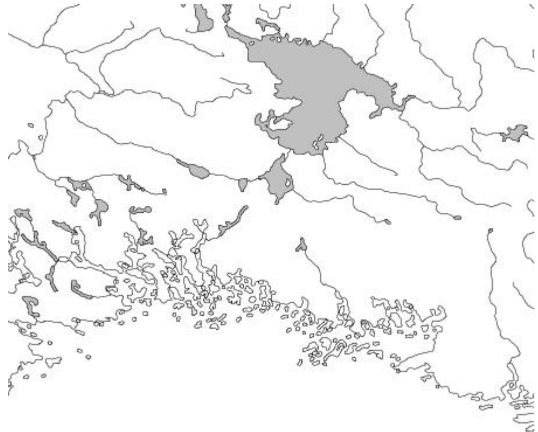

Figure 3: After automated generalisation - cartographic process

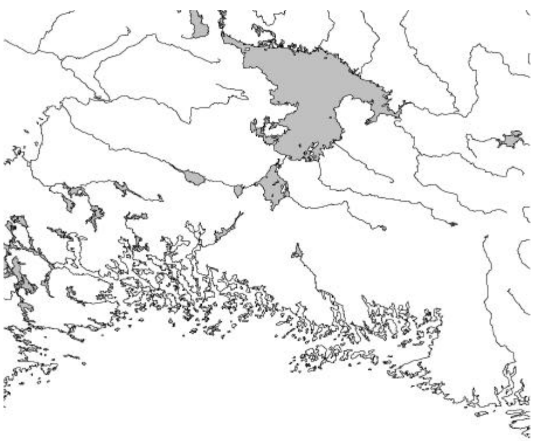

Figure 2: After automated generalisation – selection

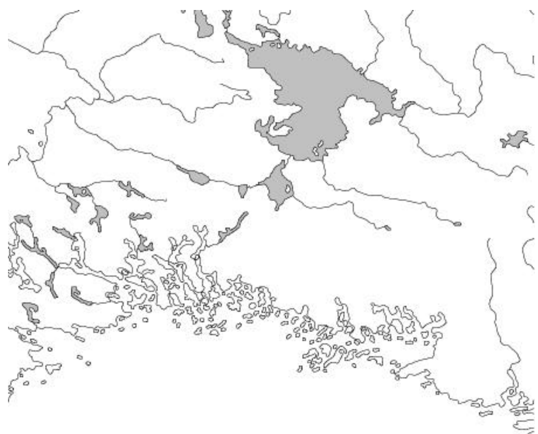

Figure 4: After cleanup by an experienced cartographer.