Model-based cartographic generalisation with uncertainty

Antoni Moore

Spatial Information Research Centre
University of Otago. Dunedin, New Zealand
Phone: +64 3 479-8138 Fax: +64 3 479-8311
Email: amoore@infoscience.otago.ac.nz

Presented at SIRC 2005 – The 17th Annual Colloquium of the Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
November 24th-25th 2005

ABSTRACT

The aim of this paper is to outline a proposed project to holistically generalise spatial data using agents. Cartographic generalisation is a process that is fraught with uncertainty – for a particular spatial scale there are an infinite amount of combinations for the display (or non-display) of data in the map space. Each map element (e.g. objects such as roads or buildings can be map elements) is an agent, with the ability to self-diagnose for cartographic conflict and reason with uncertainty (using Dempster-Shafer theory) to choose how to display itself in conjunction with neighbouring objects. Synoptically, a legible map will have been created through the intelligent interaction of agents at the local scale. This paper will explore issues associated with the above process.

Keywords and phrases**:** agents, holistic, ontology, Dempster-Shafer theory of evidence, automation

1.0 INTRODUCTION

Despite digital advances, the map that you buy at the bookshop is the product of manual intervention. The *generalisation* of map features ensures that maps are readable (by simplifying the map data as it decreases in scale), yet such common cartographic processes are not yet fully automated. Digital maps are created in Geographical Information Systems (GIS), the recent proliferation of which to community level (Weiner *et al*, 2002) raises questions: if non-cartographers have the power to produce maps that are widely available (via the Internet), will there be an abundance of poorly-designed maps? Is there any way in which cartographic expertise can be distributed to non-experts?

1.1 Aims

This paper outlines a project that aims to capture and apply cartographic knowledge digitally. The specific objective is to develop a system that models automated cartographic processes in the form of generalisation techniques, managing them in tandem to produce a well-designed map. Several issues - each of which translate into a secondary aim - present themselves:

- *how to make the system intelligent enough to deal with the complexity of a conventional map*

Generalisation processes are normally rendered digitally, as programming code. Therefore, a secondary aim is to imbue this “dumb” code with intelligence. This is part of the modelling task, and is achieved by individual map features acting as *agents*, working together to effect a *holistic generalisation*. A linear road object, and any other feature on the map, can be an agent, an autonomous software object that can interact with their environment, including other, similar objects. This is an example of a reactive agent (Gimblett, 2002), being able to change their spatial and attribute behaviour as a result of their negotiation.

A case for holistic generalisation is that certain features cannot be placed in their true position without causing conflict; some form of displacement is needed, but so doing may propagate conflict elsewhere. Essentially the synoptic approach monitors the state of the map at the global scale, making changes that have meaning at that scale. The alternative is the piecewise approach, which may produce a good map locally, but inevitably it will be at the expense of neighbouring map regions.

- *how to work within this system to automatically recognise cartographic conflict when it occurs*

The agents should be able to store generalisation functions as part of their make-up, so that they can perform any required processes on themselves. Establishing requirement is another, secondary aim, and could be achieved by self-diagnosis functions that detect cartographic "ailments". Such an ailment could be if a feature has too large a density of points at the current scale of display. This occurrence would be a prerequisite for line simplification to take place.

- *how the system can choose an ordered subset from several different methods of generalisation, each of which could reasonably be applied*

The choice of the "best" generalisation process(es) to apply may not be straightforward, with "conditions" such as illegibility as well as cartographic conflict and other constraint violations competing to be "cured". The third sub-aim is to use Dempster-Shafer theory of evidence (Shafer, 1976) to model this uncertainty and help resolve the confusion.

- *how specific feature types are prioritised, depending on the map purpose*

To reach a recommendation, the Dempster-Shafer technique will often refer to a feature's overall importance on the map. This will depend upon the feature's type, as arranged in a hierarchical ontology of the map type concerned (an ontology is an arrangement of what exists; Mark *et al*, 2005). Therefore, the next sub-aim is to build ontologies of selected map types, supported by metadata ("data about data"). For example, a road map ontology would prioritise road line features by promoting that node in the hierarchy, giving them precedence in the case of cartographic conflicts (see the ontology view in Figure 1). Associated with these feature nodes are a sequence of generalisation operators and recommended scale ranges for display. This is stored as metadata.

- *minimising the spatial error that occurs as a result of the generalisation (even though the practice of cartographic generalisation sanctions spatial error to ensure legibility)*

The fifth secondary aim is to provide facilities for the measured error of generalisation to be fed back and iteratively minimised.

Some of the background leading up to the ideas put forward in this paper is provided next, followed by a more in-depth examination of the aims above, reinforced by examples. Finally, associated issues and future work will be discussed.

2.0 BACKGROUND

Holistic generalisation was kick-started in the early 1990s with a project that employed a knowledge-based approach (Buttenfield and McMaster, 1991). This project was not totally successful (Kilpeläinen, 2000), though the impetus for maintaining the map's gestalt remains (Mackaness, 1994; Müller *et al*, 1995). Amongst the many attempts to formalise generalisation rules (Mackaness, 1994; Bundy *et al*, 1995; Müller *et al*, 1995; Weibel, 1997) were projects to assess whether generic graphical constraints have been violated after the generalisation of an entire map (Ruas and Plazanet, 1997), and the mapping of generalisation processes to specific features (Lee, 1995).

The use of agents in generalisation is a fledgling research area (Institut Geographique National –IGN, 2000; Li *et al*, 2002). The use of a method to holistically reason with the uncertainties (i.e. Dempster-Shafer theory of evidence) associated with model-based cartographic generalisation is novel, even outside of an agent context. Within geospatial science, Dempster-Shafer has been mainly used as a method of data fusion in remote sensing (e.g. agricultural classification – Kontoes *et al*, 1993; sedimentary geology – Ferrier and Wadge, 1997). A Dempster-Shafer engine (COAMES - Moore *et al*, 2001) has been developed to model and process metadata collected for coastal zone management in conjunction with expert rules. The core of that system will be adapted for use in the proposed project.

The use of ontologies and hierarchies to arrange generalisation processes and map features is not new (van Oosterom, 1993; Li *et al*, 2002), and the use of metadata has been suggested, but so far not followed up (Lam *et al*, 2005). The proposed monitoring and minimization of generalisation error effects is still a rare endeavour, despite an in-depth and convincing argument (João, 1998). Finally, the problem set out in this proposal, that of automated generalisation, is one of the major challenges in Geographic Information Science (Rhind, 1998) and has contemporary relevance, as it is set out in the University Consortium for Geographic Information Science research agenda (Lam *et al*, 2005).

3.0 PROCESS

The overall aim of the project is to build a generalisation system accessed via a cartographic language. The system would ultimately put cartographic knowledge at the disposal of non-experts. The mode of use of the system would be like this (refer also to Figure 1):

- Using the language to build an ontology of map features specific to a particular type of map (the ontology in the figure is based on van Oosterom's scheme - 1993). Elements of the language input will also be used to build a metadataset for each feature type.
- Take the source map, which has been designed for a specific scale, and decrease the scale of the map so that the features within are now crammed into a fraction of the original space. Generalisation is needed.
- Assign each feature on the map to an agent. Each agent will use various small functions to self-diagnose any cartographic "ailments". There is a generalisation solution for any of these ailments; in this way the system acts as a cartographic "doctor". There is a set order of self-diagnosis functions and generalisation solutions to perform (the order is defined by the ontology and the metadata). However, the degree of intervention will be dependent on the Dempster-Shafer belief value; the higher the value, the more action is needed.
- As an example, a line agent (representing a road) may perform a point density check on itself and recommend a line reduction solution. The greater the density, the more reduction occurs, therefore the greater the Dempster-Shafer value. This value may be modified by the road feature's position in the ontology. If roads are important to the map, simplification will occur (and a further self-diagnosis may recommend exaggeration too); if not important, then elimination is an option. In the case of conflict (i.e. overlapping features) communication with neighbouring agents will be needed to reach a solution.
- Each solution is assessed for generalisation error (the geometric difference between the generalised map as it stands, and the original map) – in the event of a choice, the solution that propagates the least error may be chosen.
- The self-diagnosis is performed again iteratively until no changes are required to the map.

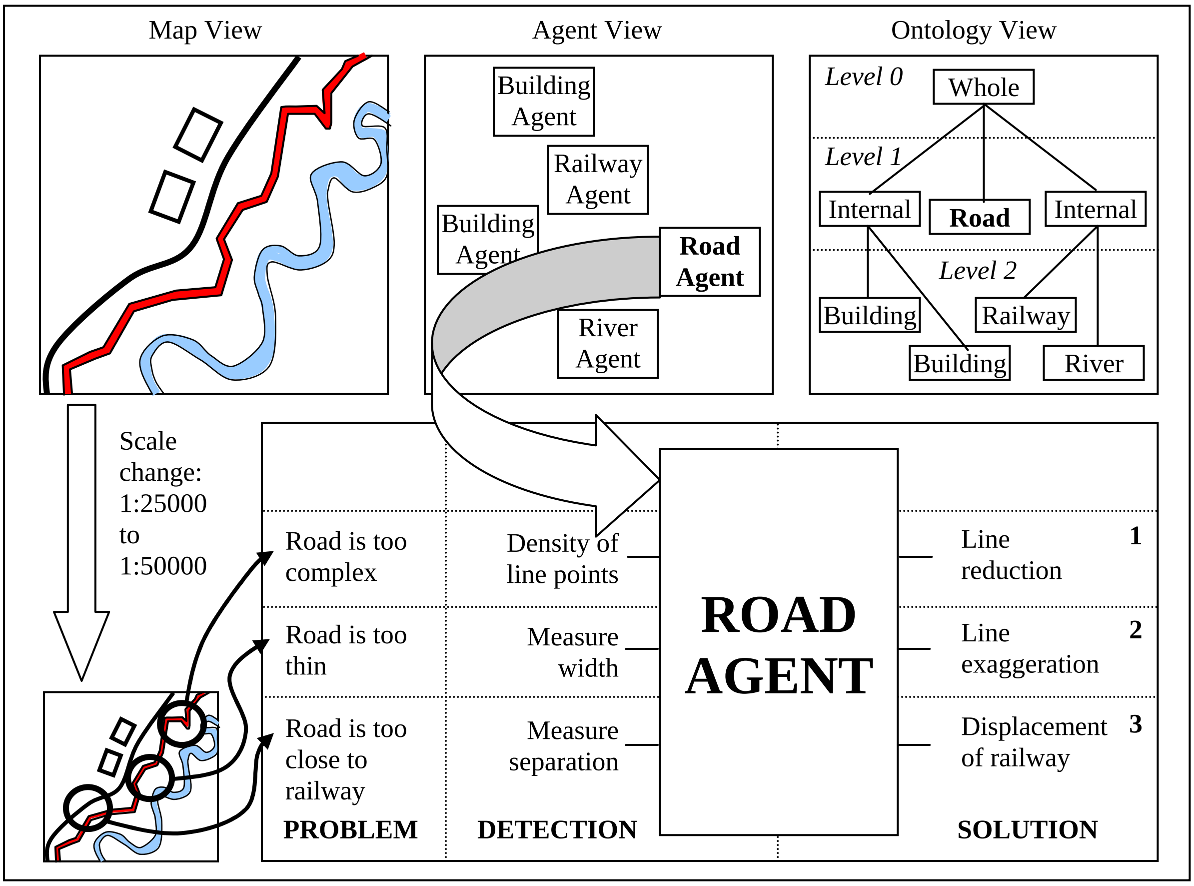

Figure 1: A scale change results in cartographic ailments (problems) for a road feature on a road map (therefore it is the most important feature – see ontology). These can be recognised by self-diagnosis functions stored with the road agent that contains the feature (detection). The agent can also implement the required generalisation function to cure the ailment (solution), but only if there is sufficient evidence, as calculated by Dempster-Shafer.

There will be four immediate tasks in this research project (the development of a language and eventual user testing of the system will be undertaken at a later stage – see next section):

a) Identifying generalisation processes and map types

The first stage in this task is to identify the processes of generalisation that exist. One such ontology may include simplification or reduction (of the complexity of the feature being treated), collapse (e.g. if the feature is a polygonal city, it may collapse into a point with scale reduction), elimination, exaggeration and displacement (of the feature) amongst others. Such conceptual frameworks of generalisation processes already exist (a good example is in McMaster and Shea, 1992; for a general overview of generalisation processes see Jones, 1997). Second, the types of map that exist (e.g. road map, general purpose atlas, topographic map for tourists etc.) will be identified and will focus on New Zealand cartography. Two or three representative types of map will be chosen (enough to address changes in scale within a continuum from local [Dunedin city] to global level). From the chosen maps, a selection of features will be extracted for use in the study, enough to demonstrate the interplay of various generalisation techniques as they perform a holistic task (driven by agents in concert).

b) Generalisation Processes and Rules

A selection of generalisation techniques will be made, enough to work with the selected features to demonstrate the concept behind this proposal. The methods will be represented by name in each feature types' metadataset (see next subsection). Programming code (collated from online sources or specifically written if needed) will be associated with the techniques as applicable. Techniques will also be written to measure the amount of error generated by the application of the chosen generalisation techniques.

c) Deriving Ontologies of Map Features

A hierarchical ontology of feature types will be derived for the chosen map types and recorded using suitable means (a scripting language such as XML - Li *et al*, 2002). A metadataset will be built in parallel and linked to the nodes in the hierarchy; this will store crucial map information such as the range of scales at which a feature can be displayed and the order in which generalisation functions should be applied to that feature. These resources will be "inherited" by the agents, and will be used by an agent to self-generalise, as applicable.

d) Map Rules to Feature Agents

Agents (programmed in Java) come in types – for example, any road is an instance of one type of agent (that representing all roads), as is any specific feature type. This stage will entail the linking of agent types to feature types, and therefore linking to the map type ontology hierarchies (through to the processes of generalisation represented in the metadata). Also in this stage is the development of self-diagnosis functions for each agent type (this is for use in cartometric evaluation – see McMaster and Shea [1992] for an example conceptual model). These functions will work on the data possessed by the agent, each passing a Dempster-Shafer belief value to a central engine (outside of the agents), which will combine all belief values for that agent and recommend a solution (belief values may also come from neighbouring agents, in the case of conflict). As stated earlier, the Dempster-Shafer engine has already been written as part of the Coastal Management Expert System (COAMES).

4.0 DISCUSSION AND LATER TASKS

A future stage in the development of the system is the definition of a cartographic language. This is the mode via which a user can control the system, and is therefore an essential component in the overall aim of disseminating cartographic knowledge to non-specialists. The cartographic language will drive map feature agents and is designed to control the workings of b, c and d in the previous section. Through this language, it is intended that the ontology of a map type can be established, and nodes in the ontology linked to sequences of generalisation techniques and other constraints, as applicable (and subsequently written to the metadataset, though this will be "behind the scenes"). The language will also be used to link ontology nodes with feature types, as represented in the map data. A good initial test for the language, once developed, will be to parameterise an as yet undeveloped map type. Previous attempts at deriving a cartographic language include a modelling language for spatial analysis (Tomlin, 1990; Pullar and Sun, 2001) and use of VRML modelling language for 3D maps (Fairbairn and Parsley, 1997), though none of these work holistically.

The major test for the system is whether the computer-generated maps are comparable with manually generated maps. An experiment is planned at this stage, where participants will be shown the map at original scale, a manually generalised map at smaller scale and a computer-generated map at smaller scale. Of the latter two maps, they will not be told which is which. They could be prompted to answer questions designed to highlight possible constraint violations (e.g. “is feature A too close to feature B?”; “does feature C have too much detail at this scale?”). Ultimately, they will be asked which one is the computer-generalised map. There will be 3 scale changes tested: local => regional (Dunedin => Otago); regional => national (Otago => New Zealand); and national => global.

5.0 SUMMARY

This paper has outlined a proposed project for the model-based generalisation of a map in a holistic manner. The overall purpose of such automated generalisation is to put cartographic expertise into the hands of the increasing number of everyday GIS users at the community level. To model the uncertainty associated with which generalisation technique to use for a particular situation, Dempster-Shafer theory of evidence is proposed. Based on a sufficiently high D-S belief value, intelligent map feature agents should be able to check themselves to establish need for generalisation, before self-applying the relevant technique. Later tasks include the development of a cartographic language as user interface, and the overall testing of the system.

ACKNOWLEDGEMENTS

Thanks to Peter Whigham for his input into this proposal. Part of the proposed research will be supported by a University of Otago Research Grant in 2006.

REFERENCES

Bundy, G L, Jones, C B and Furse, E. 1995. Holistic Generalization of Large-Scale Cartographic Data. In: J C Müller, J P Lagrange and R Weibel (eds.) *GIS and Generalization: Methodology and Practice*. London: Taylor and Francis, 106-119.

Buttenfield, B and McMaster, R B. 1991. *Map Generalization: Making Rules for Knowledge Representation*. Harlow, UK: Longman.

Fairbairn, D and Parsley, S. 1997. The Use of VRML for Cartographic Presentation. *Computers and Geosciences*, 23, 4, 475-481.

Ferrier, G. & Wadge G. (1997). An Integrated GIS and Knowledge-Based System as an Aid for the Geological Analysis of Sedimentary Basins. *International Journal of Geographical Information Systems*, 11 (3), 281-297.

Gimblett, H R. 2002. Integrating geographic information systems and agent-based technologies for modelling and simulating social and ecological phenomena. In H R Gimblett (ed.) *Integrating Geographic Information Systems and Agent-based Modelling Techniques for Simulating Social and Ecological Processes*. Oxford: Oxford University Press, pp. 1-20.

Institut Geographique National (IGN). 2000. *AGENT: Map Generalisation by Multi-Agent Technology*. http://agent.ign.fr [Accessed 25th July 2005]

João, E M. 1998. *Causes and Consequences of Map Generalisation*. London: Taylor and Francis.

Jones, C B. (1997). *Geographical information systems and computer cartography*. London: Longman

Kilpeläinen, T. 2000. Knowledge Acquisition for Generalization Rules. *Cartography and Geographic Information Science*, 27, 1, 41-50.

Kontoes, C., Wilkinson, G.G., Burrill, A., Goffredo, S. & Megier, J. (1993). An Experimental System for the Integration of GIS Data in Knowledge-Based Image Analysis for Remote Sensing of Agriculture. *International Journal of Geographical Information Systems*, 7 (3), 247-262.

Lam, N, Catts, D, Quattrochi, D, Brown, D and McMaster, R B. 2005. Scale. In: R B Mc Master and E L Usery (eds.) *A Research Agenda for Geographic Information Science*. Boca Raton: CRC Press, pp.93-128.

Lee, D. 1995. Experiment on formalizing the generalisation process. In: J C Müller, J P Lagrange and R Weibel (eds.) *GIS and Generalization: Methodology and Practice*. London: Taylor and Francis, pp. 219-234.

Li, M, Zhou, S and Jones, C B. 2002. Multi-agent Systems for Web-Based Map Information Retrieval. In: M J Egenhofer and D M Mark (eds.) *GIScience 2002, LNCS 2478*. Berlin Heidelberg: Springer-Verlag, pp.161-180.

Mackaness, W A. 1994. Issues in Resolving Visual Spatial Conflicts in Automated Map Design. In: T C Waugh and R G Healey (eds.) *Advances in GIS Research: Proceedings of the Sixth International Symposium on Spatial Data Handling.* London: Taylor and Francis, vol.1, pp.3-17.

Mark, D, Smith, B, Egenhofer, M J and Hirtle, S C. 2005. Ontological Foundations for Geographic Information Science. In: R B Mc Master and E L Usery (eds.) *A Research Agenda for Geographic Information Science*. Boca Raton: CRC Press, pp.335-350.

McMaster, R B and Shea, K S. 1992. *Generalisation in digital cartography*. Washington DC: AAG.

Moore, A B, Jones, A R, Sims, P C and Blackwell, G K. 2001. Intelligent Metadata Extraction for Integrated Coastal Zone Management. *Proc. GeoComputation 2001*, University of Queensland, Brisbane, Australia, CD-ROM: ISBN 1864995637.

Müller, J C, Weibel, R, Lagrange, J P and Salgé, F. 1995. Generalization: state of the art and issues. In: J C Müller, J P Lagrange and R Weibel (eds.) *GIS and Generalization: Methodology and Practice*. London: Taylor and Francis, pp.3-17.

Pullar, D and Sun, S. 2001. Using a Cartographic Modeling Language to Manipulate Spectral Satellite Imagery. *Photogrammetric Engineering and Remote Sensing*, 67, 6, pp.685-690.

Rhind, D. 1998. Foreword. In: E M João, *Causes and Consequences of Map Generalisation*. London: Taylor and Francis, pp. xi – xii.

Ruas, A and Plazanet, C. 1997. Strategies for Automated Generalization. In: M-J Kraak, M Molenaar and E M Fendel (eds.), *Advances in GIS Research II: Proceedings of the Seventh International Symposium on Spatial Data Handling.* London: Taylor and Francis, pp.319-335.

Shafer, G. 1976. *A Mathematical Theory of Evidence.* Princeton, NJ: Princeton University Press.

Tomlin, C D. 1990. *Geographic Information Systems and Cartographic Modeling*. Englewood Cliffs, NJ: Prentice Hall.

van Oosterom, P. 1993. Reactive Data Structures for Geographic Information Systems. Oxford, UK: Oxford University Press.

Weibel, R. 1997. A Typology of Constraints to Line Simplification. In: M-J Kraak, M Molenaar and E M Fendel (eds.), *Advances in GIS Research II: Proceedings of the Seventh International Symposium on Spatial Data Handling*. London: Taylor and Francis, pp.533-546.

Weiner, D, Harris, T M and Craig, W J. 2002. Community participation and geographic information systems. In: W J Craig, T M Harris and D Weiner (eds), *Community participation and geographic information systems*. London: Taylor and Francis, p. 3-16.