Automating geological mapping: A constraint-based approach

Azimjon Sayidov
University of Zurich
Winterthurerstrasse 190
Zurich, Switzerland
azimjon.sayidov@geo.uzh.ch

Robert Weibel
University of Zurich
Winterthurerstrasse 190
Zurich, Switzerland
robert.weibel@geo.uzh.ch

Abstract

Cartographic generalization in geological mapping is receiving increasing interest, though only few reliable automated generalization tools are available for this purpose today. Thus, improvements to methods for the generalization of categorical data, such as geological or soil maps are in demand. We advocate a constraint-based approach for geological map generalization, which could be implemented by integrating vector and raster based generalization methods. The research is divided into three parts: conceptual development, process modelling and data processing, and vector and raster based geological map generalization. In the first part, we develop the general methodology of the research, including identification and classification of constraints for geological map generalization, while the second part is dedicated to process modelling and its implementation. The third part of the research evaluates the results of generalization while comparing advantages and drawbacks of vector-based generalization against raster-based generalization. Below we give a short summary of the overall research idea highlighting the gaps found, methods used and some initial results.

Keywords: Geological mapping, map generalization, constraint-based.

1 Introduction

Map generalization is both a central and complex process if map-making. This process is responsible for producing legible and useful maps, by making choices about what to display, simplify, aggregate or even emphasize for specific map purpose. Due to the importance of map generalization, its automation has been an active area of research for several decades [4]. Most research on map generalization, however, has focused on topographic maps, which are the most common map type used (e.g. national maps, Google maps etc.). Specific thematic maps, such as geological map, which have specific geometrical and topological demands, have been largely neglected by generalization research [13]. Moreover, applying the same strategies and processes used for topographic map generalization to categorical mapping would not render a proper solution as requirements and procedures for geological map generalization are quite different from topographic mapping.

Geological maps are among the most complex thematic maps, with various elaborate shapes and structures, rendering the generalization process more demanding and require in-depth analysis of these structures prior to the generalization. One of the key properties of geological maps is that the entire map space is covered by polygons, with no overlaps or gaps.

Geological maps contains big, small, long and narrow, concave and convex, round and rectangular and etc. shapes of polygons and generalization of such complex fabrics requires making multiple interrelated and possibly conflicting generalization decisions. Such situations can be best formalized and controlled by using constraints.

The constraint-based approach to automating map generalization has emerged as the leading paradigm over the past two decades [3, 14]. In this approach, constraints are understood as design specifications and graphical condition that a valid map should adhere to. For instance, map objects should be sufficiently large to remain visible and legible on a reduced scale map; or map objects should be separated by sufficient space to remain visually separable when the map scale is reduced. In these two simple examples, a constraint would be defined for the minimum size, and a second one for the minimum separation distance. If any of these constraints are violated, a conflict resolution action is triggered, such as in the first case, when a map object becomes too small, it may be either removed or enlarged, depending on whether it is considered unimportant or important. The definition of constraints has the advantage of formulating the map generalization in a modular fashion, and formulating it as an optimization problem [3].

The overall objective of the research is to develop a methodology to automatically generalize geological maps using a constraint-based approach. The methodology considers the generalization of individual polygons as well as group of polygons. This papers presents a methodology that deals with the individual polygons in the geological maps. Next, step of the research however, is dedicated to a procedure to detect meaningful groups of polygons as a precursor to generalizing these polygon groups.

2 Background

Generalization of categorical maps can be carried out in raster as well as in vector environments, depending on the demand on the output. Thus, researches are divided in two parts. Early research aiming at generalization in a raster environment was carried out by [4] or [14]. In vector representations [7, 1, 2, 12, 6] provide examples. The integration of methods for both representations was addressed by [8, 11]. The approach of [11, 13] is confined to raster-based generalization, i.e. to maps that exist in raster form, where it works relatively well. In terms of available software tools for geological map generalization, the work by [11] still defines the state of the art. However, the approach is not able to explicitly consider cartographic properties of features such as the size of polygons or the distance between them.
Moreover, since most geological maps are stored in vector format, data will have to be converted to raster format in order to execute the generalization step, and subsequently back to vector format again. These two conversion steps cause a loss of data accuracy, which is a further drawback of the approach.
Thus, the conceptual approach used in this paper aims to improve existing methods for the generalization of geological maps by firstly identifying constraints for geological map generalization and modelling them for integrated vector and raster approaches, which are at the same time able to provide quality control for the target map.

3 Methodology and initial results

Our conceptual framework is based on defining constraints, defining corresponding measures, modelling the generalization process and finally executing the process, while monitoring quality evaluation. Moreover, it may also be regarded as a dynamic generalization model guided by constraints, where decisions depend on the semantic and geometrical characteristics of an object or set of objects, requiring the existence of procedural knowledge in order to appropriately select map generalization operators and algorithms.
In categorical maps typically the entire surface of the map is covered with contiguous polygons or areal features, with no holes nor overlaps. Such maps can equally be modelled as a vector or raster data representation, respectively.
Raster generalization is seen by some authors as the preferred choice and ideal for geological mapping at all scales [5], using classification, reclassification, majority filters, or low and high pass filters. However, it is generally not recommended to use raster generalization, unless there is a good reason, such as if the source map is in raster format or if only raster operators can handle a particular task. Otherwise, converting vector data to raster causes loss of information as well as positional accuracy of the features in the map.
The vector representation lends itself better to geometrical transformations of vertices, such as shifting the position of individual vertices, or removing vertices or polygons altogether. Also, since geological units are modelled as entire polygons rather than simply as a collection of pixels, spatial relations between polygons can be explicitly modelled, enabling better contextual operations, such as contextual aggregation of sub-categories to a unique category.
The next main steps of the framework consist in defining the generalization constraints, and in defining the measures that can implement the previously defined constraints and thus assess whether any constraints are violated.
Constraints dictate the decisions, limit the search space of the generalization process and reduce the content of the map, while generalizing it. They can be defined conceptually regardless of the spatial data model used, vector or raster, however their implementation may differ. For instance, if the pixel size of a raster is already larger than the minimum visual separation limit, the associated constraints (minimum size, minimum separation distance) will not apply. Similarly, the measures used to implement the constraints will differ between the two spatial data models. For instance, distances are measured differently in vector or raster data.
In the generalization process constraints have the following functions (Figure 1): conflict detection - to identify areas that have to be generalized, for example by evaluating the quantity and severity of constraint violations; and conflict resolution - to guide the choice of operators according to constraints priorities [2].

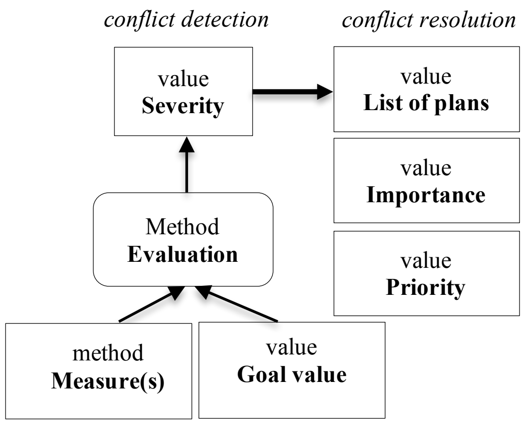

Figure 1. Modeling Constraints.

Graphical constraints, also referred to as *size constraints*, are related to the readability of the map features, such as size, width and differentiation of the objects. They are detected by graphical legibility limits and are handled in the first part of the research. Six size constraints as well as associated measures have been identified (Figure 2): 1. The number of polygons in the source and target scale should correspond to the number which identified by Radical Law [15, 16] (1).

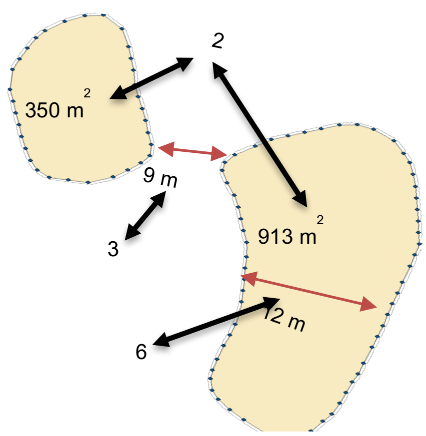

Figure 2. Size Constraints: 2. Minimum area; 3. Object separation; 6. Distance between boundaries

$$n_f = n_a\sqrt{N_a/N_f} \quad (1)$$

$n_{f-number\ of\ polygons\ in\ target\ map}$

$n_{a-number\ of\ polygons\ in\ source\ map}$

$M_{a-scale\ denaminator\ of\ the\ source\ map}$

$M_{a-scale\ denaminator\ of\ the\ derived\ map}$

2. The minimum area of polygons should not be less than 1250 m^2 (for the example of a transition from 1:25k to 1:50k); if there are polygons less then this limit they are either removed, enlarged based on their geological importance, or aggregated based on their similarities with neighbouring polygons. 3. The distance between polygons should not be less than 25 meters, and if so, they are either aggregated (again based on the geological properties) or displaced to the minimum distance. 4. and 5. The distance between consecutive vertices and the outline granularity may be handled by a bandwidth simplification algorithm and smoothing respectively, removing vertices that are very close and giving the shape a smoother look, respectively. 6. The distance between interior boundaries of a polygon should be larger than 15 meters. If not, the polygon is grown by a certain value, until its width reaches the corresponding graphical limit (Figure 2). We have recently developed a workflow-based methodology that implements the above size constraints (Sayidov & Weibel, in prep.). The methodology starts by detecting polygons that are too small. Depending on their geological importance, they are then either enlarged or removed. Proximity conflicts that may have been caused by the enlargement of polygons then trigger a series of aggregation and displacement operations, and finally the remaining size constraints are dealt with.

So far, in the first stage of this research, we have only considered constraints that deal mostly with single polygons or groups of polygons confined to their immediate neighbourhood. The next, second stage will deal with groups of polygons or polygon patterns, which could be regarded as constraints on the level of the entire map. These include e.g. 'number of categories', 'area ratios', 'group polygons proximity', 'maintenance of overall shape of patches'. On the other hand, these two stages, or levels, are closely connected and it seems fit to always link them and iterate between the two levels (i.e. individual polygons vs. groups of polygons). For instance, reducing the number of polygons in reaction to the minimum area constraint will directly affect the constraints 'maintenance of overall shape of patches', 'group polygons proximity', and 'area ratio between source and target map' which belong to the group level and map level constraints.
The final stage of this PhD research will cover the comparison of operators used in vector- and raster-based geological map generalization to assess their corresponding advantages and weaknesses in order to make further recommendations regarding the integration of these two approaches.

4 Conclusion

This PhD project departs from the hypothesis that automating the generalization of geological maps can be made more objective and flexible by integrating vector and raster-based generalization techniques and by guiding and monitoring the process with constraints that define cartographic requirements and legibility principles. Defining constraints, taking into account the properties and peculiarities of geological maps, however, is a key point accompanied by logical and structural integration of generalization algorithms. It does not only require generalization algorithms, but also algorithms that implement the measures needed to assess whether the constraints are maintained.

References

[1] Downs, T. C., and Mackaness, W. A. (2002). *An Integrated Approach to the Generalization of Geological Maps*. In *Cartographic Journal*, The, 39(2), 137–152. http://doi.org/10.1179/000870402786962489.

[2] Galanda, M. (2003). *Automated Polygon Generalization in a Multi Agent System*. *PhD thesis*, Department of Geography, University of Zurich, 188 pages. Retrieved from http://www.geo.uzh.ch/fileadmin/files/content/abteilungen/gis/research/phd_theses/thesis_MartinGalanda_2003.pdf.

[3] Harrie, L. & Weibel, R., 2007. *Modelling the Overall Process of Generalisation*. In: Mackaness, W.A., Ruas, A. & Sarjakoski, L.T. (eds.). *Generalisation of Geographic Information: Cartographic Modelling and Applications*. Elsevier Science, 67-87. http://doi.org/10.1016/B978-0-08-045374-3.X5000-5.

[4] Mackaness, W., Ruas, A., & Sarjakoski, L., 2007. *Generalisation of Geographic Information. In Cartographic Modelling and Applications*. Oxford: Elsevier. http://doi.org/10.1016/B978-0-08-045374-3.X5000-5.

[5] Marjoribanks, R. (2010). *Geological methods in mineral exploration and mining*. In *Geological Methods in Mineral Exploration and Mining*. http://doi.org/10.1007/978-3-540-74375-0.

[6] McCabe, C. (2008). *Vector Approaches to Generalizing Faults and Polygons in 1:24,000 Geologic Maps: Santa Rosa, California, Case Study*. https://www.geovista.psu.edu/publications/2008/CMcCabe_GeologicMapGeneralization.pdf.

[7] Muller, J. C. and Wang, Z. (1992). *Area-patch generalization for land use and land cover maps.* In *Cartographic Journal*, 29(2), 137-144, (1992).

[8] Peter, B., & Weibel, R. (1999). *Using vector and raster-based techniques in categorical map generalization*. In *Third ICA Workshop on Progress in Automated Map Generalization*.

[9] Sayidov, A.K. & Weibel, R. (in prep.): *Size constraints: Geological map generalization*. In *Cartography and Geographic Information Science*.

[10] Schylberg, L. (1993). *Computational Methods for Generalization of Cartographic Data in a Raster*

Environment. *Doctoral Thesis*, Royal Institute of Technology, Stockholm, Sweden.

[11] Smirnoff, A., Paradis, S. J., & Boivin, R. (2008). *Generalizing surficial geological maps for scale change: ArcGIS tools vs. cellular automata model*. In *Computers and Geosciences*, 34(11), 1550–1568. http://doi.org/10.1016/j.cageo.2007.10.013.

[12] Steiniger, S. and Weibel, R. (2005). *A Conceptual Framework For Automated Generalization and its Application to Geologic and Soil Map*s. In *Proceedings 22nd International Cartographic Conference*, La Coruña (Spain), 11-16 July 2005.

[13] Steiniger S., Weibel R., 2007. *Relations among map objects in cartographic generalization*. In *Cartography and Geographic Information Science*, 34(3): 175-197.

[14] Su, B. Li, Z. Lodwick, G. and Müller, J. C. (1997). *Algebraic models for the aggregation of area features based upon morphological operators*. In *International Journal of Geographical Information Science*, 11(3): 233–246.

[15] Töpfer, F. (1974). *Kartographishe Generalisierung*, VEB, Hermann Haack, Leipzig.

[16] Töpfer, F. and Pillewizer, W. (1966). “The principles of selection”. The Cartographic Journal, Vol. 3, 1966, pp. 10-16.

[17] Touya, G., Bucher, B., Falquet, G. & Jaara, K., 2014. *Modelling Geographic Relationships in Automated Environments*. In: *Burghardt, D., Duchêne, C., & Mackaness, W. (eds.). Abstracting Geographic Information in a Data Rich World. Springer International*. http://doi.org/10.1007/978-3-319-00203-3.