Combining Attribute with Geometry for Automated Generalization and Schematisation

Didier G. Leibovici, Jerry Swan, Suchith Anand, and Mike Jackson

ABSTRACT: Map Generalization is the process by which coarse scale maps are to be derived from fine scale maps, balancing the amount of real-world information with visual confusion. This requires the use of operations such as simplification, selection, displacement and amalgamation of features that are performed subsequent to scale change. Map schematization can be regarded as a specific type of map generalization, focusing more on the readability, the purpose and the user context, including the visualisation device, for example the London Tube map. Advanced generalization functionality derived from the literature is now being found in commercial GIS software, but many research challenges remain. Recently, focusing on the attribute values of the geometrical objects, some research has been on thematic maps, such as demographic maps, soil maps, land cover and land use maps. For these situations, algorithms need to consider ontology associated with the theme and/or statistical clustering methods. What is happening to the attributes when performing a geometrical generalization or schematization of a map with polygons in which some attributes describe their semantics? Conversely what is happening to the geometries when, some kind of generalization based on attribute values and their spatial distribution, is performed? For instance, these questions concern sequential approaches, implying that one generalization either geometrical or attribute based, will force another generalization on the other characteristic. Investigating how to use the two main generalization steps in a more integrated approach will be the ultimate goal of this paper. Examples from land cover dataset and census dataset are used for this study.

KEYWORDS: Automated generalization, map, schematization, classification, competing algorithm, entropy, census data, land cover.

Introduction

Traditionally the domain of cartographers, map generalization or schematization aims to suit the purpose of the map and the scale of the required output , Robinson et al. (1995); Ware and Jones (1998). When creating a map using traditional manual techniques, a cartographer aims to achieve a balance between the amount of real-world information required to make the map useful and avoiding confusion for the user. This is a time consuming and expensive process. GIS has led to the realization that the efficiency of the cartographer could be increased through the automation of some of the more time consuming techniques such as line and polygon simplification, Robinson el al (1995), Weibel (1995); Weibel and Jones (1998); Jones and Ware (2005). Current GIS software contains tools that allow basic generalization to be performed. Although these algorithms go some way to help in the automated production of smaller-scale maps, there is lot of work to be done to achieve full scale automation of this. The challenge of replacing an experienced cartographer with a computer that can make the same decisions to produce a map is significant. Map schematization, Waldorf (1979), Avelar (2002), Anand (2006), can be regarded as a specific type of map generalization that is concerned mainly with the simplified representation of line data. Considering both geometries and attribute values, Liu et al. (2003), Steiniger and Weibel (2005) and Revell (2007) approach generalizing thematic maps for an upscaled map delivery. So a general question is: what is happening to the attributes when schematizing a map with polygons in which some attributes are describing their semantics? The context can be either qualitative, where polygons for example will be associated to land cover classes, or be quantitative where polygons have one or more quantitative measurements attached to them such as in census data. We propose to examine the situation using three different approaches: the first two explore the sequential approaches where one would perform either a polygon schematization then an attribute “adaptation” for the new polygon delineations, or an attribute clustering and then a polygon schematization on modified/merged polygons. The third approach focuses on performing geometric and attributes transformations in a combined way, by integrating the two schematization steps in one. This can be done simply by conflating the different approaches or by integrating the two objective functions involved in attribute and geometric algorithms. We will use the terms sequential Attribute-Geometric Schematization (sAGS) and sequential Geometric-Attribute Schematization (sGAS) for sequential approaches, and the term combined Attribute-Geometric Schematization (cAGS) for the fully combined approach. On the Attribute Schematization side we are looking for classification and/or clustering algorithms helping “simplifying” the attribute description. The term classification implicitly means using non-spatial algorithms, while the term clustering may bring some spatial concept into the algorithm. On the geometric schematization side we are using optimisation principles, widely adopted for example in map generalization and described in Anand (2006), Sester (2005), Anand et al. (2007ab) within a schematization context. For the combined situation the idea is to integrate the two objective functions in the same algorithm. A general framework is described to allow competing optimization at each iteration step. A similar approach is done in Neun (2008) using web services and multi-criteria selection and collaboration filtering for orders of transformations, Burghardt (2008). The herein approach is nonetheless seen as operating at a micro-level or atomic level, therefore proposing a real integration between the two types of generalization.

The paper is illustrated for qualitative and quantitative attributes using real datasets- the very same dataset used by Revell (2007) which is an Ordnance Survey land cover map, Data-OS-LC (2000), and a census dataset with of the population in the Nottingham area, Data-Census-Nott (2001). Within the proposed general framework it is possible to take into account the use of ontology associated to a particular attribute, but also multivariate quantitative attribute considerations.

Sequential Approaches

Classification and/or clustering algorithms helping at “simplifying” the attribute description is used for Attribute Generalization. Notice the problem is not a segmentation problem as spatial objects have been already discovered and their geometry delineated, but aim at being a simplification process the attribute schematization algorithm may use a pixel-based clustering algorithm, or at least denying the rigid geometrical structure of the input dataset. The ontological description plays also an important role, especially when dealing with qualitative description where a hierarchy of classes for example may be part of the simplification classifying algorithm. We will show on a theoretical example and on the real datasets, that sequential approaches may emphasize the antagonism between the two characteristics.

The task of map generalization, is usually seen as geometrical, but papers dedicated to map generalization, Buttenfield and McMaster (1989); Muller et al.(1995); Weibel(1995); Weibel and Jones (1998) ; Jones and Ware (2005) usually exhibit sets of operators, in order to perform the generalization, which indeed may contain some attribute aspects. In summary these operators carry out tasks that include the following operators or transformations (Shea and McMaster, 1989):

- Simplification: Involves selection of a subset of the original coordinate pairs. The key objective of the simplification process is to reduce the amount of information to portray it legibly at the chosen reduced scale. Line simplification algorithms are good examples of this process. This type of operators appears purely geometrical.

- Smoothing: Refers to relocating coordinate pairs in order to plane away any small perturbations. Various smoothing procedures are applied to reduce the angularity of the source lines to fit curves which approximate the caricature of the source line. This type of operators appear also purely geometrical.
- Elimination: The removal of less important features based on given criteria, retaining the more important features. This type of operators can contain some attribute criterion, e.g. comparing the relative areas and attribute distances to the surroundings.
- Amalgamation: This process joins distinct features into a larger representative objects. It is usually applied to a set of adjacent objects of the same class. For example a group of adjacent buildings may be joined as a single building. This type of operators contain an attribute constraint, either equality of attribute of ontological proximities of them.
- Collapse: Refers to the loss of dimension by the feature during the reduction in scale process. Examples are changes from area features to line features and point features during scale reduction. This operator can be seen as elimination in the attribute space.
- Typification: This uses a representative pattern of features or symbols and is used in cases where there is a complex pattern consisting of individual features that need to be processed to retain the essence of the pattern on the smaller scale map. This operator seems very much attached to an ontological characteristic influencing the geometrical transformation.
- Exaggeration: Amplifies the shape or size of features to meet the specific requirements of a map. This is done when the mapmaker needs to deliberately highlight important features like roads on a scale reduced map. The width of the roads is in that case exaggerated compared to their real world width to highlight their importance. This transformation appears related to a constraint in the spatial attribute space induced by an ontological property, i.e. the width of the road.
- Displacement: Shifts the position of features to gain clarity.

On the attribute side, statisticians and geographers have provided ways of dealing with map productions, acknowledging some difficulties such as the MAUP issue (Modifiable Areal Unit Problem), Marceau et al. (1999). MAUP is related to first the choice of the spatial units at a given scale, and in a dependent fashion to the change of scale by aggregating these units. As here we suppose the initial representation legitimate, there is merely the second issue which could introduce a bias in the interpretation of the produced map. Putting a legend on a map with numerical attributes is already an attribute schematization; one usually chooses a number of intervals, e.g. equal breaks, percentiles. For qualitative attribute simplification can occur by reducing the number of categories to display; one usually uses an ontology associated to it or a hierarchical tree dependent on this ontology. Dealing now with spatial occurrences of the legend colors may lead to another simplification and/or a reassignment of some “regions” to specific classes. It is worth noticing that the geometric transformation by a series of operators is not going to change dramatically the units structure, on the other hand attribute clustering, already by merging polygons with the same attribute will modify the structure significantly depending on the clustering algorithms chosen. Constraints may be willing to be set in order to be conserve as much as possible some structures: for example for census data with administrative units, either by geometric schematization or by attribute schematization (and both) one usually wishes to preserve the number of administrative units. Before using real datasets it was interesting to imagine, according to a set operator respecting the rules given, what could give a sequential schematization? This is not an automated procedure but rather an interpretation of what the cartographer would do if he/she was performing these operators accordingly to the designed sequence. So we drew a simple sketch called the Before map of Figure 1, with polygons and an attribute (which could be taken as qualitative or as quantitative). This is theoretical exercise, as schematization was done manually, in order to investigate the added value of attribute schematization and to guide the choice of algorithms to perform the tasks (sGAS and sAGS).

Geometric Schematization First

Examining the sGAS (sequential Geometry schematization followed by an attribute schematization) there are, concerning the second step, two different cases to consider, *i.e.* the attribute schematization step. Once the geometric schematization has been done according to the series of operators given in the previous section, a simple and rather naïve attribute operation is the "follow-up" strategy. Respecting the changed vector structure the attribute associated to a particular geometry follows or in other words stays attached to the given geometry. The topology of the "database" ensures integrity of the evolving map.

On Figure 1, this result is illustrated by the *sGASnaive* map. A second possible approach is related to a true attribute "simplification" which could result from a clustering algorithm performed on the pixel representation of the geometrically schematized map. The map on the top right illustrates the ambiguity the attribute values may encounter, as the result of the geometrical schematization which is then adapted with an attribute "clustering", giving the *sGASAdapt* map. The values of the attribute may change, as for example with the value 1 which now 1+, meaning that it should be in between 1 and 2 (either qualitatively or quantitatively). For this very example one also may "decide" that the overlap is not enough between 1 and 2, to allow a change of values. Notice a change of value for the small polygon with the value 4 (on *BeforeMap*) which became for spatial coherence of the three plots" on the top right corner of the map. This is a kind of attribute amalgamation/typification based on closeness of values and spatial position (area and distance). Another amalgamation combined with a displacement happened for the small triangle at the bottom of the map, while an attribute collapse took place in the top right corner. This last operation can be seen as an approach already combining geometric-attribute in geometric generalization,as in Tinghua (2003) where additional semantic information were used with generalization.

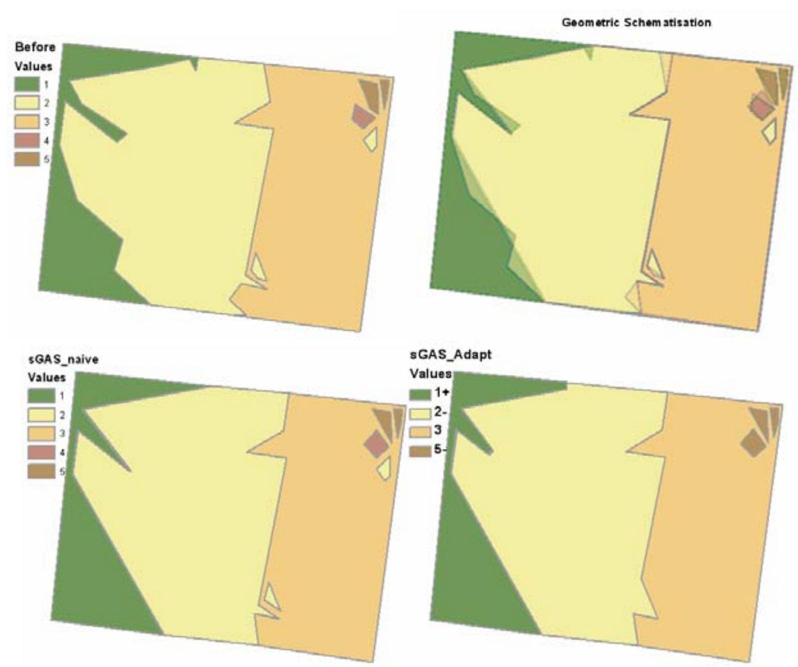

Figure 1: Sequential schematisation, geometrical first: hypothetical results with a naïve or adaptable approach for the attribute transformation; (top right is the Geometric Schematization superimposed with the Before map (top left) and the 1st step)

Based on pixel spatial neighborhoods, smoothing of the shapes occurred for the large polygon frontiers. To summarize, this sequence makes a geometric simplification preserving details that the attribute schematization either conserves or amplifies. In term of changes attribute values, we expressed a “semantic” change for example 1+ would mean something between 1 and 2. For a quantitative attribute this would mean a real increase in value by averaging according to areas concerned by the overlapping of the two values, for a qualitative attribute this could be a new created class, e.g. mainly 1 but occasionally 2 or if an existing class from the ontology of the classes. For either qualitative or quantitative, one may decide depending on an overlap threshold, if a new representing “value” has to be assigned or estimated, or if the dominant class or value can be retained. The tasks, described above for geometric schematization, can surprisingly be valuable for attribute schematization. It may be not that surprising if you consider the fact that an attribute will be “normally” attached to the geometry (polygon, line or point) in the map, unless we come back to the duality of model representation: vector or raster. Nevertheless in both model, these geometric tasks operators can be translated in relation to attribute values (range, distribution, thresholds, and series of intervals) but also to the contiguous areas and gradients.

Attribute Schematization First

The naïve and adaptable approaches can also be adopted for the *sAGS* sequence (sequential Attribute clustering followed by a geometric “schematisation”). The operations done for the attribute step in the second step of *sGASAdapt* can be “stronger” in term of spatial coherence as we do not try to preserve any geometric aspects, top right corner of Figure 2. Notice that the “groupings” due to clustering forces some potentially geometrical schematization as well in the sense of geometric smoothing and orientation (border effect). The naïve approach makes the geometry follows this forcing, the new delineation of categories (quantitative or qualitative) as a vectorization algorithm would do. Here, a vector based clustering algorithm or a high resolution pixel based algorithm may provide a less angular (not the case here) delineation, making then a fundamental difference between *sAGS* and the *sGAS* of the previous paragraph. On the *sAGSAdapt* second step, the little “triangle” (North-East) was eliminated due to surroundings of bigger areas polygons of different type (no extra typification). To summarize the *sAGS* sequence makes the attribute schematization smooth and simplify surfaces which are more likely line generalized at the second step. This is probably due to traditional statistical approach to perform an attribute schematization via attribute clustering is to consider the spatial content as pixels with values (univariate or multivariate). This is quite frequent on raster image data but also on vector data seen as a raster dataset. With this approach the clustering method will provide new classified values for each pixel which then will give rise to a zonation via a polygon representation of the raster.

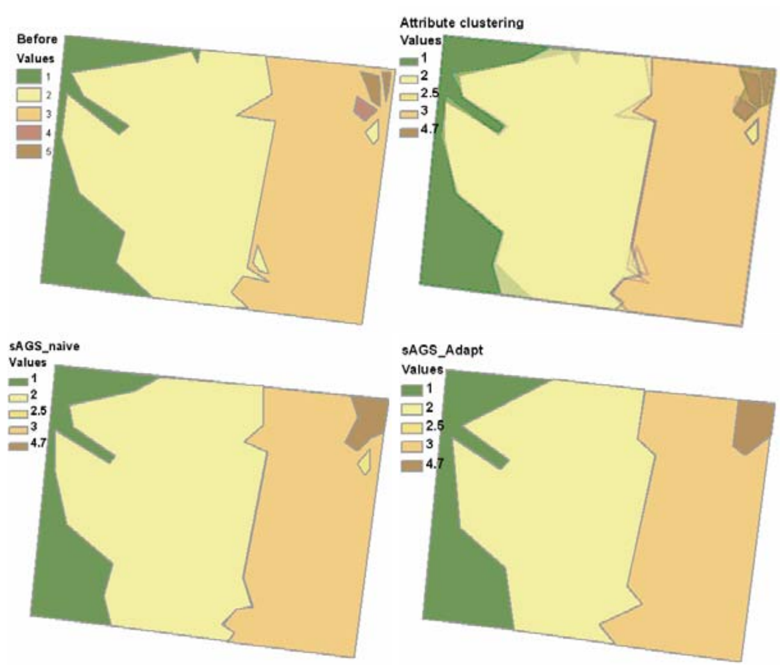

Figure 2: Sequential schematisation, attribute first: hypothetical results with a naïve (bottom left) or adaptable (bottom right) approach for the geometrical transformation; (top right is the Attribute clustering superimposed with the Before map (top left)

As a matter of fact the *sGAS* and *sAGS* sequences should not necessarily stop in two steps as after the second step of naïve or adaptable geometrical or attribute schematization there must be at least a naïve “follow the change” for the other characteristic. Also nothing stops us carrying on the sequence to improve schematization at a higher level.

Combining Optimization

The previous experiment convinced us to find common ground for combining the tasks, described in the first place as geometrical, but revealing themselves as *attributical* as well. Spatiality is in the attributes as well as semantic is in spatial pattern. This may be not that surprising if you consider that an attribute will be usually attached to a geometrical object (polygon, line) in the map, unless we come back to the duality of representations: vector or raster. It becomes interesting to investigate possible combinations of the two generalization issues.

Conflating Results

A consensus or an assimilation of the two schematised map can be a first solution. How do we conflate these two maps with schematization rules in mind? Carrying on the theoretical experiment we obtained in Figure 3. Conflation doesn’t ensure, as it happens here, a schematized final map result: just compare to the original (Figure 1 or 2). Conflation principle means also you would have more information in the conflated than in “the sum of the two”, somehow. The conflation of the end-results of the sequential approach can be also considered but would not necessarily reach a simplification level either. Though at first sight, it would be better than conflating just 1st steps of the two sequential approaches. The other combined way is doing a simultaneous optimization on geometry and attribute values.

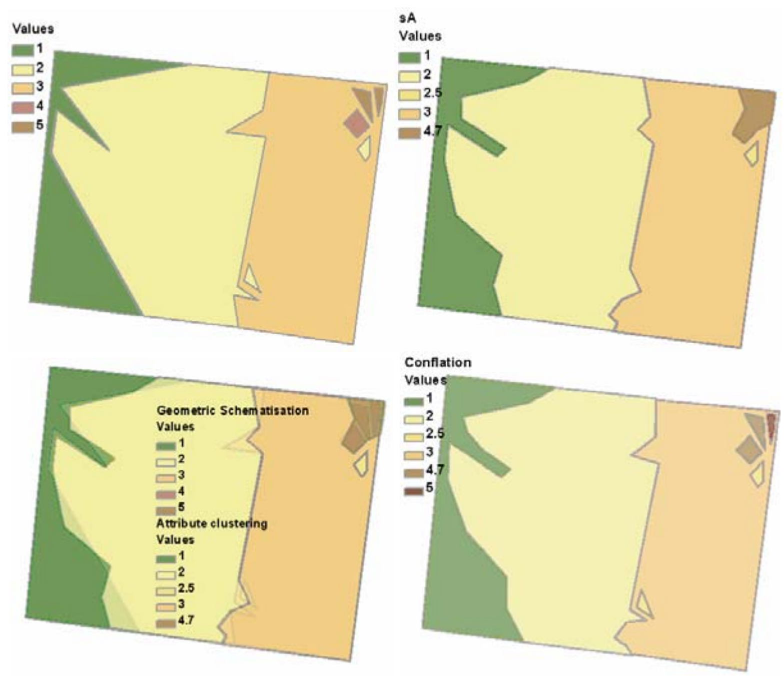

Figure 3: Combined schematization by (hypothetical) conflating (bottom right) of geometric schematization (top left) and the attribute clustering (top right)

Integrated Approach

A much more integrated approach consists of combining the two objective functions or heuristics associated to the two different characteristics of a map, *i.e.* attribute characteristic and geometric characteristic. For a chosen characteristic, we suppose that, a particular map ***M*** will provide an associated characteristic description c. Both the map instance and the specific characteristic description are the key elements to select a transformation operator to be applied: ***T*** belonging to a family of appropriate operators for the intrinsic properties of the characteristic, and the associated family type of maps, *e.g.* networks, land covers and economic map. ***T*** will operate by transforming the map instance and by setting up a possible new characteristic description.

$$T \in \mathfrak{F}\{O_p(\mathfrak{C},\mathfrak{M}))\,/T(c,M) = (\hat{c},\widehat{M})\} \tag{1}$$

At each step of the generalization optimization, the best transformation also depends on the available objective functions:

$$\widetilde{T} \quad = \quad arg \max_{\substack{i\in\{att,geo\}\\ T_{att}\in\mathfrak{F}_{att}\\ T_{geo}\in\mathfrak{F}_{geo}}} Of_i(T_i,(c_i,M)) \tag{2}$$

This can be seen as a sequential approach but, at an atomic level, within a competing algorithm in where the objective functions have to be normalized. Instead of doing an ad-hoc normalization, we used a functional able to cope with both: the entropy information framework has shown some success in the context of map generalization, Bjorke (1996):

$$Of_i(T_i,(c_i,M)) = -H(p_k) = \sum_{k=1}^{n_{c_i}} p_k \log(p_k) \tag{3}$$

where here p_k are proportions of equivalent classes in the mathematical sense, identifying categories semantically for the given characteristic description. For the attribute characteristic this can be defined by the grouping used for the legend, which for example would be the set of modalities for a qualitative attribute. For the geometric characteristic, classes of segments for example are defined by groups of angles values. Minimizing the entropy will therefore diminish the heterogeneity of the map. Because of different numbers of classes for different characteristics a simple normalization in comparison to uniform distribution is still necessary. This means simply dividing the entropy by $log(n_{ci})$. In terms of spatial heterogeneity, spatial co-occurrences of the categories represented can be used instead of occurrences as in the standard entropy. These spatial co-occurrences defined for a given distance give rise to a definition of spatial entropy, Leibovici et al. (2008a 2008b).

Optimization Framework for the Integrated Approach

Implementing the Integrated Approach can be done within a general optimisation framework. This framework generalises in some sense the optimisation schema developed by Swan et al. (2008) and its web service version as in Foerster et al. (2008). This framework allows considering various optimisation strategies such as *genetic* algorithm, *tabu* search, and even the *classical hill climbing*, as can be seen on Figure 4. In brief the framework considers an optimisation procedure as depending on three components: the states, a neighbourhood of a state and a heuristic function evaluating a state.

Here a state is defined by the couple (**c**,***M***) where ***M*** is a map and ***c*** is the characteristics considered. Here ***c*** is a list of length 2: the attribute and the geometric characteristics; each of these characteristics is also a set of the different classes associated. So ***c*** is a couple of vectors of characteristics. The neighbourhood of a state is created by applying the available operators associated to the characteristics. The heuristic function is our objective function described above. To illustrate this framework we used in the first place some simple operators and characteristics associated to the attribute or geometrical properties targeted in the generalization/schematisation.

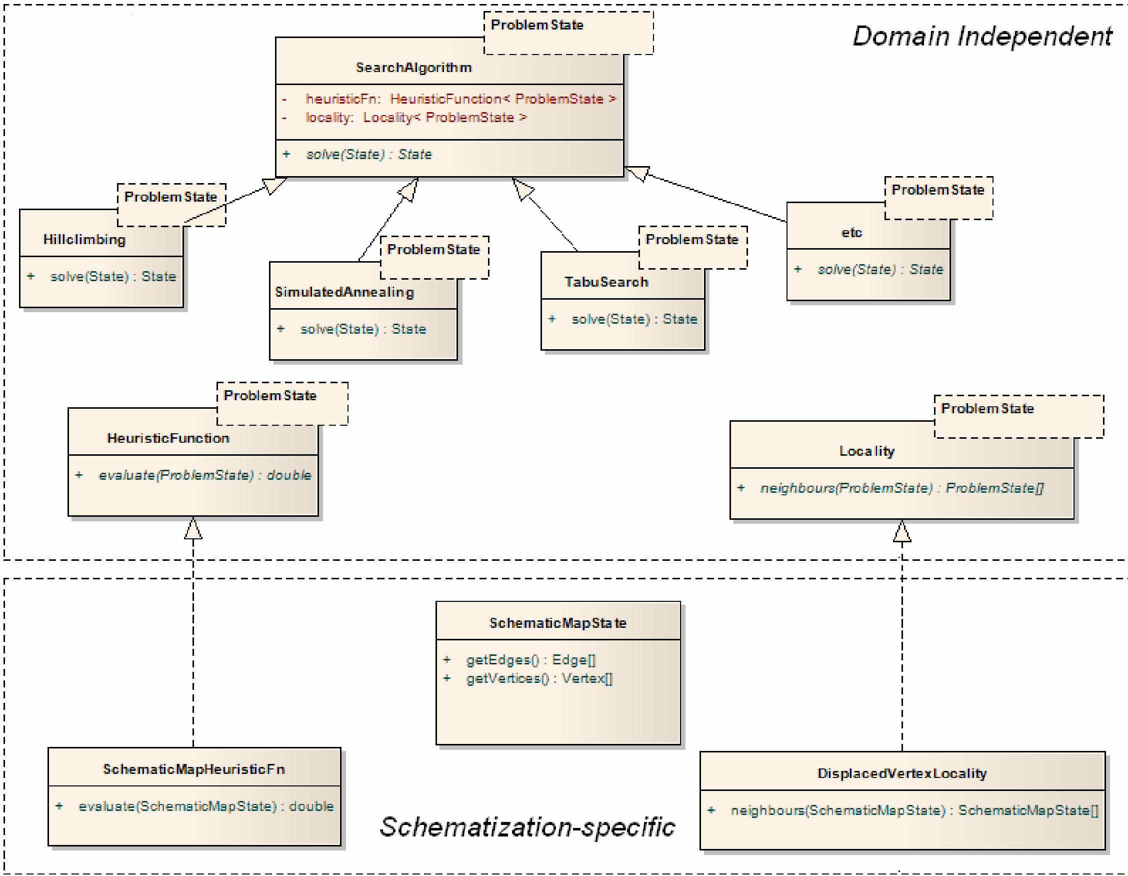

Figure 4: UML description of the optimization framework

The list of operators specific to each characteristic could be expanded with further development as discussed in the last section. One can notice that in equation (1), without regards to the characteristics used, the operator ***T*** operates in a separable manner:

$$T(c/M) = \widehat{c} \text{ and } T(M/c) = \widehat{M} \tag{4}$$

Each time ***T*** operates conditionally on the current state of either ***c*** or ***M***. The first operation is changing the structure and the second is changing the instance.

Simple Attribute Operators

Many different operators for the attribute transformations can be derived from classification clustering methods used in statistics and spatial statistics. Lots of these methods can work either in vector representation or in raster representation, though the latter is more often adopted especially for spatial clustering. For this first implementation we chose to stay in a vector representation system and use simple k-means ideas and hierarchical ascendant clustering based on Euclidian distances.

The ***k***-means attribute operator, $\boldsymbol{Oatt_{km}}$ can be summarised as follows. At the current state, a map is described from the attribute characteristic with all the classes represented by the available different attribute vectors. The distances of all the distinct objects can be computed using these vectors. ***k*** objects, called centres in attribute space, randomly chosen or selected from previous $\boldsymbol{Oatt_{km}}$ transformation, initiate the transformation. Then iterations between *assignment* and *update* of the centres take place: (i) The centres determine the assignments of all the others according to the minimum attribute distance. (ii) The attribute vectors of the ***k*** centres are updated as for example by the mean of all the assigned attribute vectors. The iterations stop when the assignments do not change anything. Then as this optimisation stops, the attribute vectors of each object take the values of the centre it belongs to. It is also possible to diminish ***k*** by 1 every time it is used. The hierarchical ascendant attribute operator, ***Oatt***$_{hier}$ can be summarised as follows. At the current state, a map is described from the attribute characteristic with the classes represented by all the available different attribute vectors. The metric distances (or any chosen ultrametric) between these objects in the attribute space are computed and the closest pair is aggregated. Like for the ***Oatt***$_{km}$ the aggregating function can be the mean of these 2 vectors. One can also keep the pair vectors and use a membership variable to indicate the aggregated status. This is convenient when using the minimum linkage criterion, ***i.e.*** the minimum distance, as the clustering ultrametric.

These operators can be used as such, but in order to introduce more interactions between the attribute and the geometrical characteristics, one can take into account some spatial properties of the objects within the operators and particularly for the aggregating function: weighting the objects by their areas when computing the mean, or using the inverse of the areas as weights when computing attribute distances for the clustering function. At the moment, as we do not use any spatial clustering based for example on pixels the geometrical properties of the objects are not modified when using a attribute operator ***Oatt***. As we implemented amalgamation in the geometrical attribute, it is not needed as an ***Oatt***.

Simple Geometrical Operators

Line simplification and amalgamation are the two basics operators we have chosen to implement. The amalgamation operator ***Ogen***$_{ama}$ will simply work on a constraint that the attribute of two adjacent objects (minimum distance chosen) have the same attribute vectors (minimum distance in attribute space chosen) and will merge the two objects by the envelope of their geometries. The line simplification algorithm ***Ogen*** line will remove one point for the geometry of an object (being represented by a polygon or a line) using the well known Douglas-Peucker algorithm. The classes' characteristic description will be given by the number of points in geometries. Here when modifying the geometry of an object by ***Ogen***$_{line}$ one could consider a change of attribute for this geometry: either no change which corresponds to the naive approach in the previous section, or a weighted transformation in order to conserve the volumes (attribute values are multiplied by the ratio from old to new area). For these two sets of operators, ***Oatt*** and ***Ogen***, some particularities occur for qualitative attributes. Distances and aggregations depend on the ontological structure or graph. If the ontological structure is not use for aggregation, rules such as majority linked to spatial coverage can be used.

Discussion about First implementation Results

Java programming using GeoTools has been used to implement the framework. Different parameters for the sets of operators and the way they operate can be chosen beforehand. The overall idea of the optimisation is to allow operators to be selected in a competing way. The tuning of the parameters has to be done for fair competing optimisation. For nearly each operator it is possible to investigate and use a version of it that is as atomic as possible or allowing a series of atomic transformation. For example the ***Ogen***$_{line}$ can be operating one polygon transformation or a given number, proportion of the objects (or classes) or until we get ***(k-1)*** classes in term of the geometric characteristic. ***Oatt***$_{hier}$ can be chosen to perform one attribute aggregation, equivalent here to ***(k-1)*** classes, or a given number, proportion of the objects (or classes). The behaviour of these two operators in terms of number of classes is very different. ***Oatt***$_{hier}$ will, after one iteration, reduce the number of classes by one, whilst ***Ogen***$_{line}$ will be unlikely to do so, but will have fewer classes in practice. In fact the ***Oatt***$_{km}$ transformation may be favoured too much, if allowed to reduce too drastically the number of classes, and, reducing just by one is equivalent to applying ***Oatt***$_{hier}$. The former problem is somehow reduced by the normalisation of the entropy relatively to uniformity.

For a simple implementation keeping the various operators, we allowed ***$Oatt_{km}$*** to do a ***(k-2)*** clustering and kept the other operators as defined earlier. Do the operators of one characteristic need to imply or use a transformation/regulation on the other characteristics? When using ***Oatt***, changes of the attribute values may take into account weighting of the areas used. In the same manner ***Ogen*** by suppressing one vertex of a polygon may change the attributes of the polygon (and the outside polygon) when focusing on conserving volumes on each particular attribute: area times attribute, or when focusing on conserving densities. These conditions appear to be sensible, in that attribute space maps are not just displays of numbers, the spatial context or extent is obviously meaningful for the attribute statistics; in the geometric space the relative comparison from geometry to geometry about their attribute should be conserved when modifying one geometry.
On an example small dataset, we were very pleased to see the series of operators used in the log history for the optimisation, coming sometimes from the attribute characteristic, and sometimes from the geometry characteristic. Unfortunately due to some issues with java within the GeoTools library used for the implementation, we are so far unable to produce a visualisation of the results. This happens for a complex object (a multiset of polygons) which is at the moment not interoperable with geoserver. The inconsistency of the identification of java objects made also the algorithm unreliable to date.

Conclusions

This paper described a general approach on using both attribute values and geometric properties to perform a schematization of a map. On a simple example we investigated the possible outcomes of chaining in sequencing classical algorithms for geometric schematization and clustering. A general framework to embed both schematization issues in order to perform a combined, simultaneous schematisation is also proposed. For a first implementation of this framework we concentrated only on quantitative attribute datasets, but we believe this framework sufficiently general to be able to adapt to datasets from different domains, but these may require specific constraints.
Constraints play an important role in schematization; they are related to the domain, the purpose and the desired level of schematization to provide. For example, in a map representing the land cover background of the *Tour de France* circuit, one would be interested more in the mountain differentiations rather than the forest and fields. The framework presented here with entropy based objective functions, can accommodate constraints by incorporating rules on some characteristic classes in terms of their list of allowed transformations. When using an entropy measure, the lack of spatiality has been notified by few authors, Li and Claramunt (2006). The last authors proposed with success to modify the entropy using spatial weights onto class k. The weights were linked to spatial coherence of the classes events using intra and inter-distances. We investigated here the direct use of co-occurrences of the events to describe the distributional property of the characteristic description classes, as in Leibovici et al. (2008ab).

Acknowledgments

We would like to thank Edina for providing Digimap Ordnance Survey data: © Crown Copyright/database right 2008, an Ordnance Survey/EDINA supplied service; and also MIMAS for providing the census data: “Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland.”

References

Anand, S Ware, J.M Taylor, G.E (2007) ***Generalization and schematization of large scale digital geographic datasets for MobileGIS applications***, Book chapter in Drummond,J., Billen, R., Forrest, D. and Jo˜ao, E. (eds***), Dynamic and Mobile GIS: Investigating Change in Space and Time***, Taylor & Francis publications, pp 161-174.

Burghardt, D. (2008) Collaborative filtering for an effective selection of generalization services. ***Geospatial Web Services workshop University of Nottingham 2008, http://www.nottingham.ac.uk/geography/geowebservices/.***

Bjorke, J.T (1996) Framework for Entropy-based Map Evaluation. ***Cartography and Geographic Information Science***, 23(2), p.78-95(18).

Data-Census-Nott (2001) Nottingham Unitary Authority Census. Office for National Statistics, 2001 Census: Standard Area Statistics (England and Wales).[england_ua_low_soa_2001.shp and Census2001ks001.dbf] ESRC/JISC Census Programme, Census Dissemination Unit, Mimas (University of Manchester)/Centre for Interaction Data Estimation and Research(University of Leeds) ***http://www.statistics.gov.uk/geography/soa.asp***

Data-OS-LC (2000) Land Cover dataset from Digimap bounding ***http://www.digimap.co.uk***

Douglas,D.H. and Peucker, T.K. (1973) Algorithms for the reduction of the number of points required to represent a line or its caricature.***The Canadion Cartographer***, 10(2):112-122.

Foerster, T Burghardt, D Neun, M Regnauld, N Swan, J and Weibel, R (2008): Towards and Interoperable Web Generalization Services Framework – Current work in progress. ***Proceedings of the 11th ICA workshop on Automated generalization and Multiple representation***.

Jones, C.B and Ware, J.M (2005) ***International Journal of Geographical Information Science: Map Generalization Special Issue*** (Taylor and Francis), Volume 19, Number13 8-9, pages 859-997.

Leibovici, D.G Bastin, L and Jackson, M (2008a) Discovering Spatially Multiway Collocations. ***GISRUK Conference 2008,*** Manchester, UK, 2-4 April, 2008.

Leibovici, D.G and Jackson, M (2008b) Spatial cooccurrences for spatial entropy. ***In preparation***

Liu, Y Molenaar, M Ai, T and Liu, Y (2003) Categorical Database generalization aided by Data Model***. Proceedings of the 21st International Cartographic Conference*** 201, 10-16.

Neun(2008) "Automated processing for map generalization using web services" ***GeospatialWeb Services workshop University of Nottingham 2008,http://www.nottingham.ac.uk/geography/geowebservices/***

Revell, P (2007) Generic Tools For Generalising Ordnance Survey Base Scale Landcover Data ***International Cartographic Association 10th conference on Generalization and Multiple Representation***, 2-3 August 2007, Moscow.

Robinson, A.H Morrison, J.L Muehrcke, P.C Kimerling, A.J and Guptill, S.C (1995***) Elements of Cartography, 6th Edition***, New York: John Wiley & Sons.

Swan, J Anand, S Hobona, G and Jackson, M (2008) Map schematization using the CGS Persistent Testbed. ***Geospatial Web Services workshop University of Nottingham 2008, http://www.nottingham.ac.uk/geography/geowebservices/.***

Sester, M (2005) Optimization approaches for generalization and data abstraction. ***International Journal of Geographical Information Science*** 19, 871.

Steiniger, S and Weibel, R (2005) A Conceptual Framework For Automated Generalization and its Application to Geologic and Soil Maps. ***Proceedings 22nd International Cartographic Conference***, La Coru-a (Spain), 11-16 July 2005.

Ware, J.M Anand, S Taylor, G.E and Thomas, N (2006) Automatic Generation of Schematic Maps for Mobile GIS Applications. Special Issue on location based services and MobileGIS, ***Transactions in GIS***, Blackwell Publishing, Vol 10, Issue 1 ,25-4

Ware, J.M and Jones, C B (1998) Conflict Reduction in Map Generalization Using Iterative Improvement, ***GeoInformatica2***:4, 383-407.

Weibel, R and Jones, C.B (1998) Special issue on automated map generalization, ***GeoInformatica***, Volume 2, Number 4.

Dr Didier G. Leibovici, Research Fellow in Geospatial Modelling, Centre for Geospatial Science, University of Nottingham, University Park. Nottingham, NG7 2RD, . E-mail:<didier.leibovici@nottingham.ac.uk>.
Dr Jerry Swan, Research Fellow in Geocomputation, Centre for Geospatial Science, University of Nottingham, University Park. Nottingham, NG7 2RD, . E-mail:<jerry.swan@nottingham.ac.uk>.
Dr Suchith Anand, Research Fellow in Geospatial Information, Centre for Geospatial Science, University of Nottingham, University Park. Nottingham, NG7 2RD, . E-mail:<suchith.anand@nottingham.ac.uk>.
Pr. Mike Jackson, Professor of Geospatial Science, Centre for Geospatial Science, University of Nottingham, University Park. Nottingham, NG7 2RD, . E-mail:<mike.jackson@nottingham.ac.uk>.