Quality Assessment of Model-Oriented Generalization

Tao Cheng
Department of Geography, University of Leicester,
University Road, Leicester LE1 7RH, United Kingdom
Email: tcheng@le.ac.uk

ABSTRACT

Multi-scale generalization changes the semantic and geometric resolutions of objects according to context defined by users. Although different strategies have been applied in the model-oriented generalization, the definition of the context is seldom crisp, which will cause uncertainties in the generalization results. Furthermore, the operations used in these strategies, such as aggregation/merging, will change the geometric, thematic and topological relationships between the objects. This will also create uncertainties in the generalization result. The objective of this research is to develop mechanisms for quality assessment in the model-oriented generalization. The questions to be answered in this research include what causes uncertainties in the model-oriented generalization, how they are propagated in the procedure and how to quantify and visualize them. A set of quality assessment indices is put forward to calibrate the uncertainties. A case study of a land use map is implemented to examine the geometry-driven and class-driven strategies for the model-oriented generalization.

Key words: uncertainty; multi-scale; model-oriented generalization; quality assessment;

1. Introduction

Map generalization is the selection and simplified representation of detail appropriate to the scale and/or the purpose of a map (ICA 1973:173). It is a highly complex process by which the presence of geographical features within a map is reduced or modified in terms of their size, shape or numbers. It is carried out either for display, for data reduction or for analysis. There has been progress in recent years for generalization in algorithms or entire software systems (Müller *et al.*, 1995a; Weibel & Jones, 1998; Richardson & Machaness, 1999).

Multi-scale generalization usually changes the semantic and geometric resolutions of objects according to context defined by users. The definition of the contexts is seldom crisp, which will cause uncertainties in the generalization. Furthermore, the operations of generalization, such as aggregation/merging of objects, will change the geometric, thematic and topological relationships of the objects. This will create uncertainties in generalization. However, the assessment of the quality of generalization results has received relatively little attention in research so far (Müller *et al.*, 1995b; Weibel & Dutton, 1999). Few exceptions can be found in the discussion of the effects of generalization on attribute accuracy in natural resource maps (João, 1995; Pairnho, 1995). Recently, quality assessment has been incorporated into the automated generalization design by some constrains (Brazile, 1998; Weibel & Dutton, 1998). Although the quality of generalization is controlled by the constraints, only a particular solution (operation or algorithm) for generalization is assessed (McMaster, 1987; Berg *et al*, 1995). In general, quantitative assessment methods are weak when multiple objects are involved or entire maps needs to be characterized. An integrated approach is needed to capture the more holistic elements of generalization (Weibel & Dutton, 1999, p. 150). Although Ehrliholzer (1995) proposed to integrate quantitative measurements with qualitative judgments by cartographic experts, a comprehensive palette of evaluation methods and strategies, however, is not available, which is indispensable for progress in generalization research.

Therefore, it is necessary to cast light on the study of uncertainties in generalization and provide a quality assessment of the generalization result. The objective of this research is to develop mechanisms for quality assessment in model-oriented generalization. We will study the uncertainties in generalization and provide the quality assessment of the result. The questions to be answered in this research include what causes uncertainties in generalization, how they are propagated in the procedure and how to quantify and visualize the quality of the generalization.

The paper is organized as follows. Following the introduction, Section 2 presents the four strategies of model-oriented generalization. Section 3 investigates the uncertainties in generalization. It analyzes the causes of uncertainties and how they are propagated in generalization operations. The approach for evaluating and representing uncertainties in generalization is discussed in Section 4. It is followed by a case study in Section 5. The generalization of a land use map is presented by using the geometry-driven and class-driven strategies. Section 6 discusses the results and provides the guidelines for quality control in these two strategies. Finally, the last section, Section 7, summarizes and concludes the paper. The direction for further research is also presented in this section.

2. Strategies Of Model-Oriented Generalization

There are two main branches in digital cartography and GIS: cartographic generalization and database (or model-oriented) generalization (Weibel & Jones, 1998). Cartographic generalization represents the process of deriving a graphic product or visualization from a source database. In cartographic generalization clarity and logical consistency of graphic expression are given priority over positional accuracy and completeness. Model-oriented generalization concentrates on the derivation of reduced database from source database. It may be carried out for various purposes, in order to control data reduction or to derive data sets of reduced accuracy and/or resolution; or as a pre-process step to cartographic generalization (Weibel & Dutton, 1999). It prioritizes spatial accuracy and completeness based on processes that can be modeled formally.

Four strategies are usually applied in the model-oriented generalization (Molenaar, 1998).

Geometry-driven generalization is a strategy where the geometric resolution is the driving factor of the aggregation process. Objects are aggregated to form new objects that are large enough to be represented. After generalization the adjacent objects might belong to the same type. Changing resolution of grid of raster images is a typical geometry-driven approach.

The *class-driven generalization* is a strategy where regions consisting of mutually adjacent objects that belong to the same class are aggregated to form larger spatial units with uniform thematic characteristics. A consequence of class-driven generalization is that after aggregation there are no two adjacent regions that are of the same type. The generalization of a land use database according to the land hierarchical classification scheme can be considered as a class-driven generalization.

Function-driven generalization aggregates spatial objects at a low aggregation level to form new objects at a higher level, according to the aggregation hierarchy based upon functional relations between objects. For example, the aggregation of area of residential houses and area of green land around the houses as a residential area is a functional approach.

Structure-driven generalization simplifies the description of a network system, while keeping the overall structure intact. The total functioning of the system after generalization can be understood at a less detailed level. For example, the branches of a stream are usually reduced based upon the structure of the stream.

These four strategies might be applied simultaneously to the generalization of a single database, though they are different in terms of selection criteria and alteration of resolution (see Table 1). For the geometry-driven strategy, the geometric criterion (e.g. size) has the first priority, and then the topologic criterion (e.g. adjacency) is applied. Under such strategy, the spatial resolution is changed as required, and the thematic resolution is also adjusted. For the class-driven strategy, however, the thematic criterion has the first priority, and then the topologic criterion is applied. Under such approach, the thematic resolution is reduced with the adjustment of spatial resolution. Other two strategies please refer to Table 1.

Table 1. Comparison of four strategies of model-oriented generalization.

Strategy	*Order of Selection Criteria*			*Resolution*	
	Thematic	Geometric	Topologic	Spatial	Thematic
Geometry driven		1	2	Change	Adjusted to new spatial resolution
Class driven	1		2	Adjusted to new thematic resolution	Change
Function driven	2		1	Adjusted to new function resolution	Adjusted to new function resolution
Structure driven		2	1	Adjusted to new structural resolution	Adjusted to new structural resolution

3. Uncertainties in Model-Oriented Generalization

The four strategies mentioned above apply several operations to achieve the aim of generalization, e.g. to derive secondary scale and/or theme specific datasets or compose special-purpose maps. The operation defines the transformation that is to be achieved; and a generation algorithm is then used to implement the particular transformation. There are many operations used in generalization and the basic ones are *simplification*, *selection/elimination/merging*, *aggregatio*n, *symbolization/collapse*, *exaggeration* and *displacement* (McMaster & Shea, 1992; Weibel & Dutton, 1999). In order to derive the reduced model (or representation) of reality, these operations essential modify the geometry, topology and/or semantics of the objects from high-scale to low-scale.

The operations, *simplification* and *exaggeration* only modify the metric aspects of the objects; *selection*, *symbolization* and *aggregation* essentially modify the topological aspects; and *displacement* is primary metric, but can involve topological changes in some cases (Dettori & Puppo, 1996, see Table 2). The changes in topology and geometry may result in semantic changes, and vice versa. The topological changes indirectly imply metric changes as well, but not vice versa. That's why uncertainties are created in generalization. For instance, where the real world distance between a lake and village is too small at a give scale to graphically display the railroad and the road between the edge of the lake and the village there is a representation conflict. Displacement is a solution, but leads to loss in positional accuracy (Harvey & Vauglin, 1998). Further, due to the missing of the (rail)roads, the semantics of the map also change, which leads to loss in the thematic accuracy.

Table 2. The modification of metric and topological aspects by operations (Dettori & Puppo, 1996, pp. 565).

	Model-Oriented	Cartographic
Only *metric* aspects	Simplification	Exaggeration Displacement
Topological Aspects	Selection Aggregation Symbolization	Displacement

As for model-oriented generalization, the basic operations are *merging* (*elimination*) and *aggregation* (Molenaar, 1998). Here *merging* is referred to a process to eliminate small areas or sub-polygons, i.e., objects are put together to build a composite object. After the merging process, the original objects cease to exist. Whereas *aggregation* refers to the process which deletes edges between similar objects and form a composite object. The semantics of the original objects are then transferred to the new composite object, but the original objects do not cease to exist. Since the merging is usually done based upon the neighbor that has the largest border or the largest area, it is actually a geometry-driven approach. Because the aggregation is usually implemented based upon a common thematic characteristic with its neighbors, it is actually a class-driven approach.

Here we analyze the uncertainties in these two operations. Although these two operations can be applied to point, line and area objects (Weibel & Dutton, 1999), we mainly consider the area objects since they are most complicated situations. Further, we will concentrate on the semantic changes after generalization because of the fundamental role of semantics for modeling and representation geographic methods (Harvey *et al*., 1998, p. 559), i.e., "the relationship between the degree of resolution and accuracy definitely cannot by determined solely by the scale or dimensions of a map, but rather by its semantics". It is also because the semantic change after generalization has been seldom studied, compared with the other two aspects, geometry and topology.

3.1 Uncertainty in Merging

In merging, the semantics of the original objects is changed to the new object. If the semantics of the original objects is different from the new one, uncertainty in semantics will be created. Let's analyze the uncertainties of the generalization in two situations of merging.

Case (1)

The first case is that a small area which is adjacent to several larger areas. We assume that D is a small area and is merged into A. We use A' to represent the area after generalization, although in the database it is still represented as A (see Figure 1).

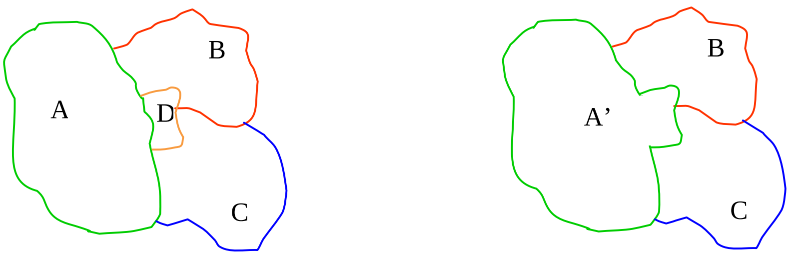

Figure 1. Case (1) - Small area D is merged into one of its adjacent area A.

If *A* and *D* are not thematically similar, the certainty (or uncertainty) of the new area A' can be expressed as:

$$u_{A'}^{C_A} = \frac{Area(A)}{Area(A) + Area(D)} \tag{1}$$

where $u_{A'}^{C_A}$ represent the uncertainty of area *A*' belongs to Class *A* (C_A).

If area *D* is merged into *B* or *C*, their uncertainty can be expressed as:

$$u_{B'}^{C_B} = \frac{Area(B)}{Area(B) + Area(D)} \tag{2}$$

$$u_{C'}^{C_c} = \frac{Area(C)}{Area(C) + Area(D)} \tag{3}$$

Case (2)

The second case is that a small area *D* is contained in (or isolated by) a large area A. After generalization, it is merged into *A*. We use *A'* to represent the area after generalization, although in the database it is still represented as *A* (See Figure 2).

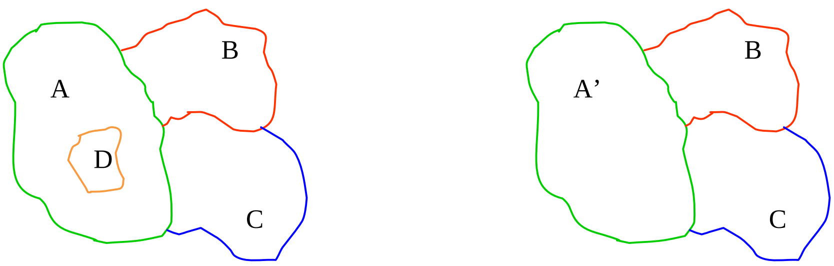

Figure 2. Case (2) - Small area D is merged into its adjacent Area A.

The uncertainty of the new area A' can be expressed as:

$$u_{A'}^{C_A} = \frac{Area(A)}{Area(A) + Area(D)} \tag{4}$$

If we compare Equation 1 with 2 and 3, we may unify them into a general case to represent the uncertainty of the new area A', as in Equation 5,

$$u_{A'}^{C_A} = \frac{Area(A)}{Area(A) + Area(D)} = \frac{Area(A)}{Area(A')} = \frac{Original\ Area}{New\ Area} \tag{5}$$

In Equation 5 we assume that Area *A* and *D* are crisp objects which belong to Class *A* and *D*, respectively and certainly.

If Area *A* and *D* are fuzzy objects (Cheng & Molenaar, 1999), i.e., they belong to Classes *A* and *D* with membership functions as $\{u_A^{C_A}, u_A^{C_D}\}^{\mathrm{T}}$ and $\{u_D^{C_A}, u_D^{C_D}\}^{\mathrm{T}}$, respectively, the thematic uncertainty of *A'* can be represented as:

$$u_{A'}^{C_A} = \frac{Area(A) \cdot u_A^{C_A} + Area(D) \cdot u_D^{C_A}}{Area(A) + Area(D)} \tag{6}$$

When Area *A* belongs to Class *A* certainly, i.e., $u_A^{C_A} = 1$, and Area *D* does not belongs to Class *A*, i.e., $u_D^{C_A} = 0$, Equation 6 becomes Equations 5. If $u_A^{C_A} = 1$ and $u_D^{C_A} = 1$, then

$u_{A'}^{C_A} = 1$, i.e., if *A* and *D* are thematically similar, there is no uncertainty after the merging operation. Therefore, Equation 6 is a general description of uncertainty of the merging operation.

3.2 Uncertainty in Aggregation

We assume there are five areas A_1, A_2, A_3, B_1, and B_2 as shown in Figure 3*a*. These five areas belong to Class *A* (C_A) and Class *B* (C_B) (c.f. Figure 3*b*). After aggregation two areas *A* and *B* are formed (c. f. Figure 3*c*).

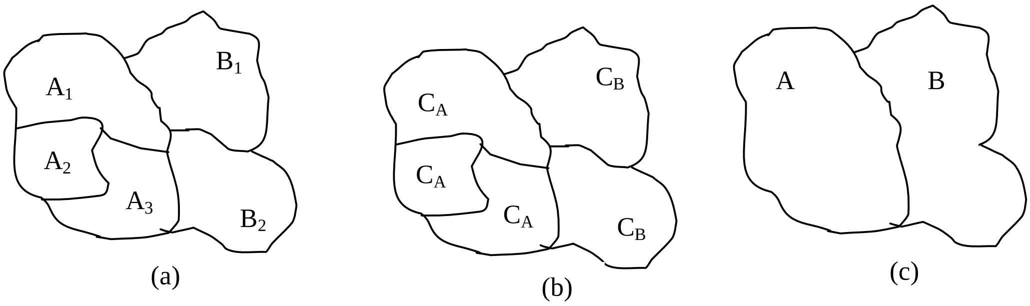

Figure 3. The case of Aggregation.

Similar to Equation 6, the uncertainty of Area *A* can be represented as:

$$u_A^{C_A} = \frac{Area(A_1)\cdot u_{A_1}^{C_A} + Area(A_2)\cdot u_{A_2}^{C_A} + Area(A_3)\cdot u_{A_3}^{C_A}}{Area(A_1) + Area(A_2) + Area(A_3)} \tag{7}$$

where $u_A^{C_A}$ represent the uncertainty that area *A* belongs to Class *A* (C_A).

If A_1, A_2, A_3 belong to Class *A* certainly, then

$$u_A^{C_A} = \frac{Area(A_1)\cdot 1 + Area(A_2)\cdot 1 + Area(A_3)\cdot 1}{Area(A_1) + Area(A_2) + Area(A_3)} = 1 \tag{8}$$

It means that when the aggregated areas belong to a same thematic class, the new area has no uncertainty.

As can be seen from Equations 6 and 7, thematic uncertainty will be created after generalization when the areas that are not thematically similar are merged or aggregated into a new area.

4. Quality Assessment and Visualization

As argued in the section of introduction, a holistic evaluation is required in order to provide the quality of the generalization. Since data quality of GIS contains components of accuracy, precision/resolution, consistency and completeness (Veregin & Hargitai, 1995), we will try to derive the quality indices of generalization according to these four components. Although space, time and theme are three essential dimensions of geographical data, the major concerns for generalization quality evaluation are the space and theme, i.e., geometry (metric & topology) and semantics. A matrix of geographical dimensions (columns) and data quality components (rows) of generalization is shown in Table 3.

Table 3. A matrix showing geographical dimensions (columns) and data quality components (row) of generalization (after Veregin & Hargitai, 1995).

	Metric	Topology	Semantics	Time
Accuracy	Positional accuracy		*Thematic uncertainties	/
Resolution	Minimum mapping units		Lowest classification	/
Consistency	Area difference per class	*Topological change	Redundant or contradictions	/
Completeness	Missing part of the area in the specification		Missing classes	/

Therefore, the following quality indices are proposed.

(1) Accuracy -- Certainty index

The accuracy of semantics can be described by the certainty index for the new area objects created after generalization. Equation 6 and 7 describe the certainty of the area object created from merging and aggregation, it can be adapted to all the objects created after generalization.

$$\begin{aligned} &\textit{certainty of new area object belonging to Class } K \\ &= \frac{\sum \textit{original area} * \textit{its certainty belonging to Class } K}{\sum \textit{original area}} \end{aligned} \tag{9}$$

(2) Consistency

(a) Object reduction index

In order to describe the topological change after the generalization, an object reduction index is defined as in Equation 8 (after Bregt & Bulers, 1996):

$$\textit{object reduction index} = \frac{\textit{reduced number of area objects after generalization}}{\textit{number of area objects before generalization}} \tag{10}$$

(b) Attribute change index

Daley *et al* (1997) compare the percentage area by class of original with generalization. It can be used to describe the consistency of semantics. But Daley's approach emphases the difference in area by per class. An overall indicator of consistency of semantics is needed. The attribute change index, proposed in (Bregt & Bulens, 1996), is adopted.

$$attribute\,change\,index = \frac{sum\ of\ \ absolute\ area\ differences\ \ per\ class}{total\ \ surface} \tag{11}$$

(c) Topology change index

The consistency in topology can be expressed as topology change index:

$$topo\log y\,change\,index = \frac{number\,of\ \ objects\,change\,topo\log y}{number\,of\ \ total\ \ objects} \tag{12}$$

(3) Resolution

The size of the minimum mapping unit and the lowest level of thematic classification can express the quality in resolution.

(4) Completeness

As for the geometric completeness, it can be detected by overlaying the generalization results with original data before generalization. As for the semantic completeness, we may check if any thematic class loses after generalization.

(5) Visualization

It is pointed out in Beard and Buttenfield (1999) that the error analysis method is associated with the graphic display modes and are necessarily bundled together. The current author believes that there are several ways to visualize the quality of generalization. In order to show the certainty of the semantics, the whole new area with its certainty index can be visualized as a fuzzy area (see Figure 4a, Figure 6c). Since the consistency in semantics and topology can be measured quantitatively, it can be calculated and reported as shown in Tables 4 and 5. They can also be visualized by indicating the objects that are merged or aggregated into other objects with their thematic values (as shown in Figures 4b, 7a & 7b). It can also be visualized by highlighting the original objects that change their thematic values after generalization (Figures 4c). The change in topology is more difficult to detect except the reduction of objects, but it can be detected by comparing the intersection of the original data with the generalization result. The objects that have different boundaries can be visualized as in Figures 11b & 12b. Other two indices, completeness and resolution can be described quantitatively and are quite straightforward.

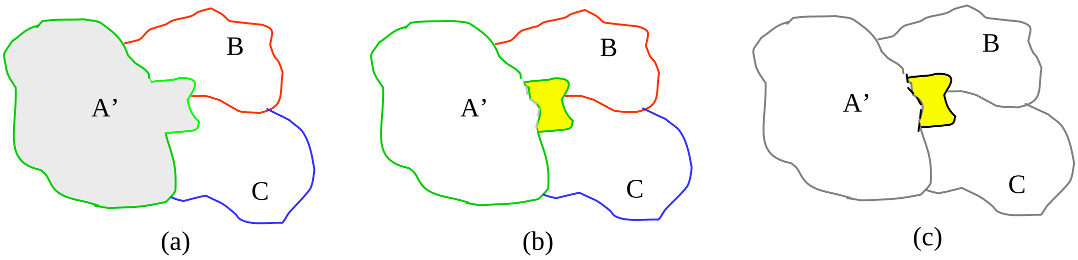

Figure 4. Visualizing uncertainties in generalization.

5. Case Study

We have a land use map with two level thematic classifications (which we assume to be free of uncertainty). Figure 5a shows the land use type at low level, and Figure 5b shows the types at high level. The objects identified at low level are shown in Figure 5c. This section uses this land use case to examine the influence of operation on generalization. We mainly study the uncertainties in the geometry and the class driven strategies.

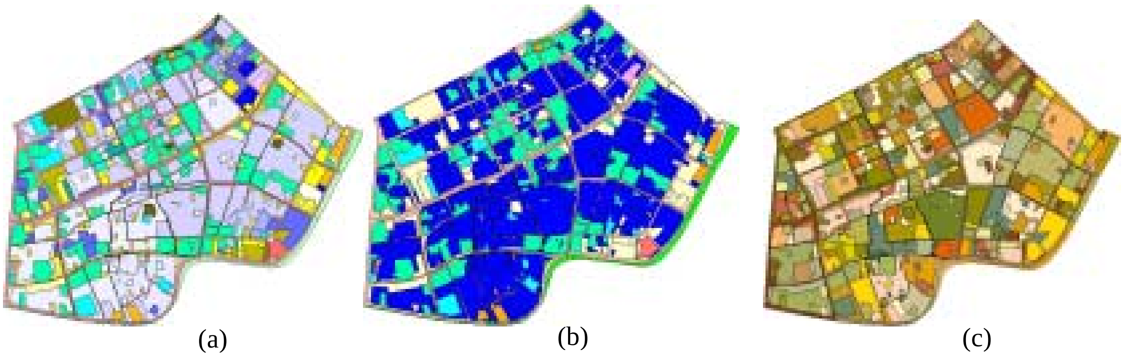

Figure 5. The land use map at two level thematic scales.

(1) the geometry-driven strategy

Since there are very small regions in the map, we use the merging operation to eliminate them. There are two algorithms to implement the merging. The first one is to merge the selected polygons with neighboring polygons that have the largest shared bored between them. The second one is to merge the polygons with neighboring polygons that have the largest area. Here we use these two algorithms and try to compare their difference. In order to check the influence of minimum-mapping unit (spatial resolution) - MMU, three threshold values (1000, 2000 and 4000) are applied. Figure 6a and 6b shows the merged results based upon largest area and border respectively, with the MMU= 2000. Figure 6c represent the certainties of Figure 6b.

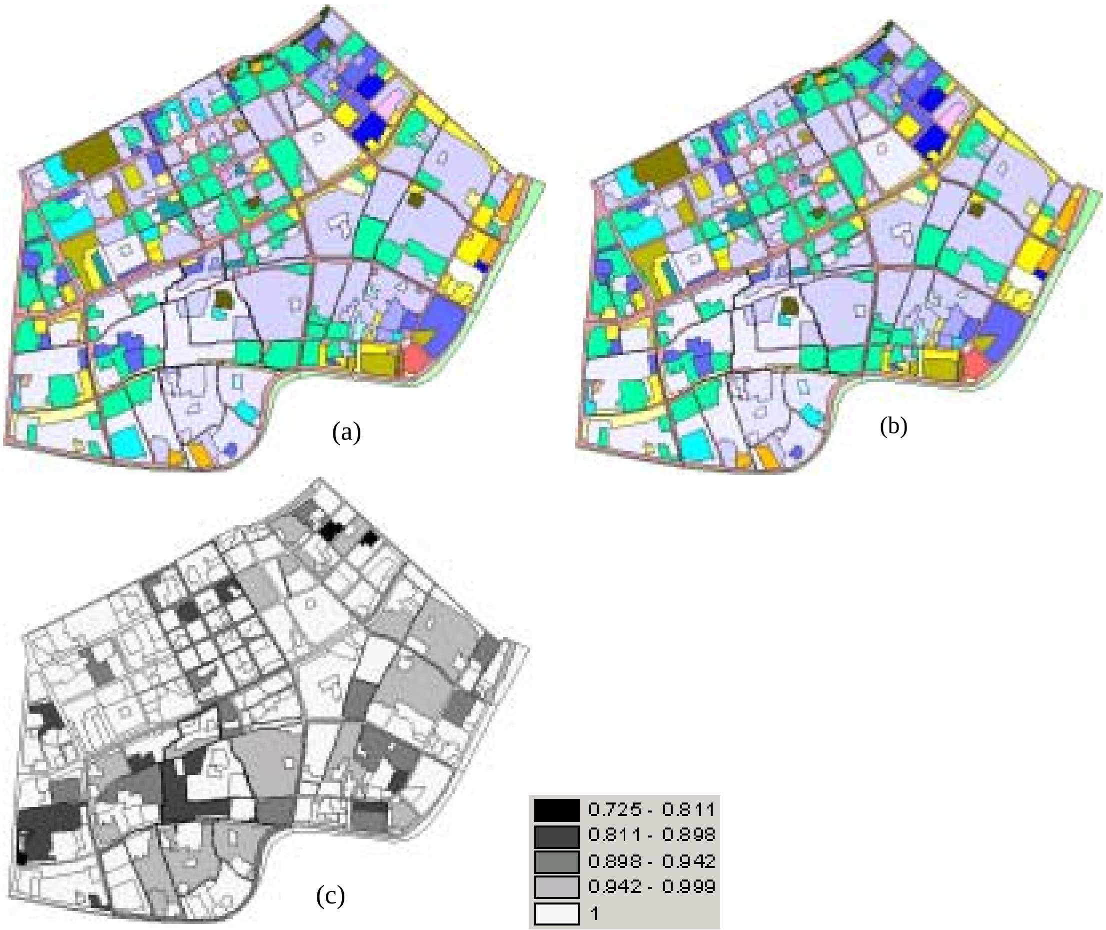

Figure 6. The merged results based upon the largest border and the largest area.

Table 4 summarizes the quality of generalization based upon these two algorithms with three MMUs. It can be seen from Table 4 that the number of objects in total and the areas of objects in each class are changing with the threshold values of the MMU. It can also be seen the attribute change is increasing with the threshold value. So does the reduction of object numbers. But the change degree of attribute is relatively lower than that of object reduction. Although the number of objects merged in two algorithms are the same, the indices of the area change are slightly different and the change based upon largest border is lower.

Table 4. Quality assessment of the merged results based upon geometry-driven strategy.

MMU	No. of Objects		Object Merged	*Object Reduction Index*	Sum of Area Difference per class		*Attribute Change Index*	
	Before	After			A	B	A	B
1000	425	386	39	0.092	30596.06	29931.74	0.0063	0.0061
2000	425	333	92	0.2165	149989.28	133887.30	0.0308	0.0275
4000	425	246	179	0.4212	599078.40	422151.82	0.1230	0.0867

(A and B represents the merging based upon the neighbor that has the largest area and has the largest shared border, respectively).

Table 4 provides the quality assessment for the whole map. We can also check the merged areas in detail; i.e. check their thematic values after generalization with their origins. The graphic visualization of the change is revealed in Figure 7. The difference between these results with original map is reported in Table 5. Averagely, 80% and 74% changed their thematic classes dramatically after merging. Also we found from Table 5 that generalization result is better based upon the largest border.

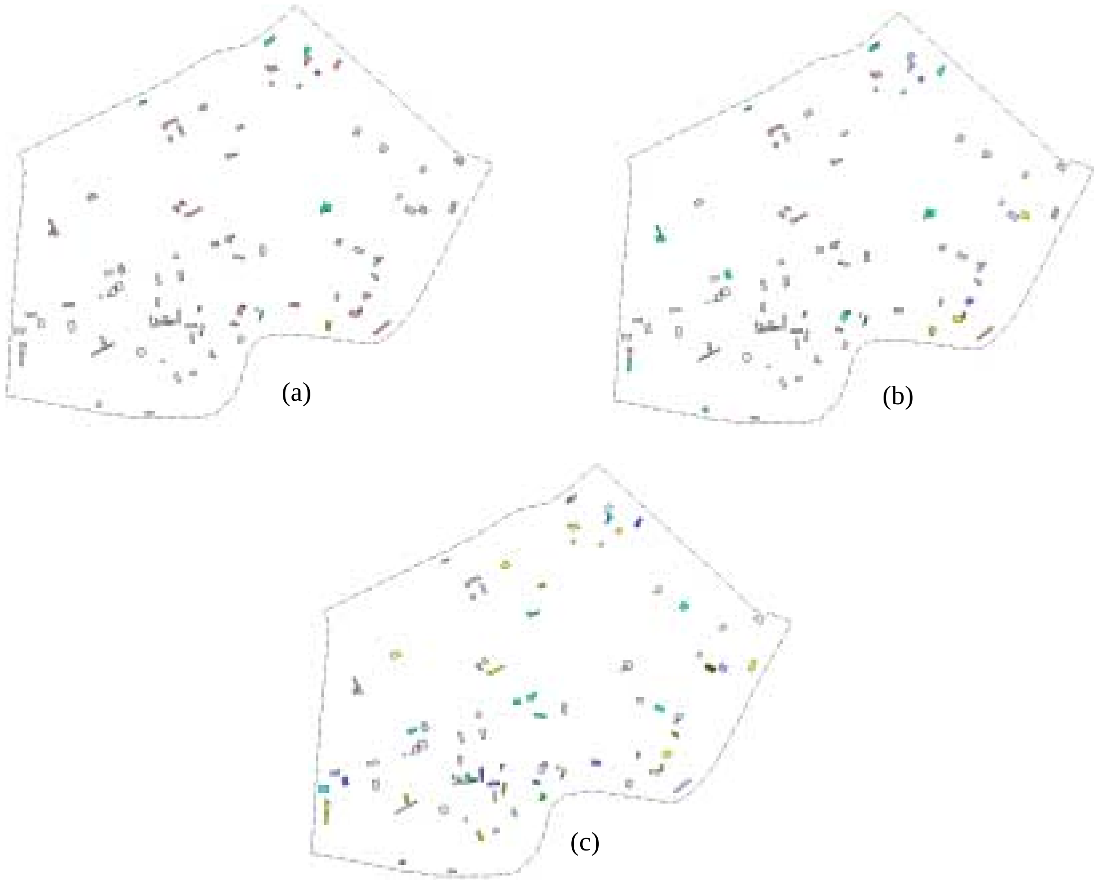

Figure 7. Visualizing the quality index of generalization.
Figure A represents their class type after generalization based upon largest area, Figure B represents their class type after generalization based upon largest border, and Figure C represents their original class type,

Table 5. Difference of the merged results with original map at the high thematic level.

Threshold Value	Object Merged	Difference (A&O)		Difference (B&O)		Difference (A&B)	
		Object	Ratio (%)	Object	Ratio (%)	Object	Ratio (%)
1000	39	32	82.1	30	76.9	9	23.1
2000	92	73| 91*	79.3	68 |90*	73.9	31| 31*	33.7
4000	179	145	81.0	130	72.6	69	38.5
Average			80%		74%		

(*Difference of the merged results with original map at the low thematic level. A and B represents the merging based upon the neighbor that has the largest area and has the largest shared border, respectively)

(2) the class-driven strategy

The class-driven strategy is implemented by dissolving the boundaries between the mutually adjacent objects of a same land use type, i.e. they are aggregated into a new object.

The aggregation results are shown in Figure 8. Figure 8a shows the result of aggregation based upon the low-level land use, while Figure 8b shows the results based upon the high-level land use.

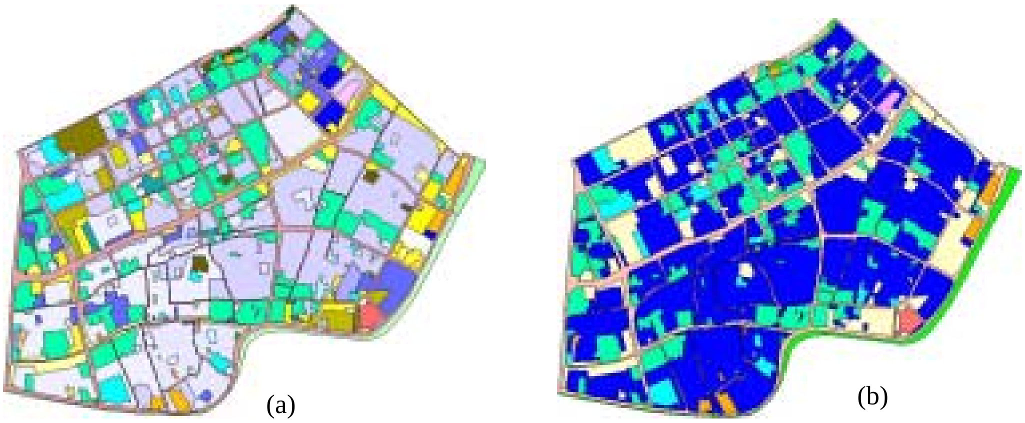

Figure 8. Generalization results based upon the class-driven strategy.

Table 6 reports the quality indices of generalization. It can be seen that the attribute change index is very low at two thematic scales, but the index of the object reduction is quite high, especially at high thematic scale, almost 30%.

Table 6. Quality assessment of the class-driven strategy at two thematic scales.

Scale	*Object No.*		*Sum of Surface Difference*	*Index*	
	Before	After		Attribute change	Object reduction
Low	425	393	0.0530	0	0.08
High	425	282	0.0403	0	0.34

6. DISCUSSION

This section analyzes the results in the case study of Section 5 in order to provide quality control of generalization.

(1) the geometry-driven strategy

For different minimum mapping units, the difference between attribute change and object reduction is obvious (see Table 4). The change of attribute is around 1% to 9%

and the reduction of object number is around 10% to 40%. The degrees of change in both aspects all increase with the size of the MMU. It implies that the uncertainty in semantics is increasing with the size of the MMU.

(2) the class-driven strategy

The class-driven strategy generalizes the map based up the thematic value of the adjacent polygons, i.e. the polygons that belong to same classes are aggregated into one. Therefore, the class-driven strategy only reduces the object numbers without changing their attributes. It implies the uncertainty in semantics by this strategy does not make much difference for different thematic scales, but the change in topology is increasing with the aggregation scale.

(3) comparison of two strategies

Since three threshold values of MMU are applied in the geometry-driven strategies, it is difficult to compare the results with those obtained by the class-driven strategies. But we find the result obtained at MMU=1000 is closer to the class-driven result at low thematic level (see Table 7). Still we can find that the numbers of the objects in these two strategies are slightly different. There are 386 objects (MMU=1000) in the geometry-driven generalization and 393 objects in the class-driven generalization. This is due to the fact that the adjacent objects belonging to a same class are aggregated into one object in the class-driven strategy, while in the geometry-driven case they are still separate objects. But there are smaller regions (less than 1000) in the class-driven results, which are merged out in the geometry-driven strategy.

Table 7. Comparison of the geometry-driven and the class-driven strategies.

Scale	*Object No.*		*Sum of Surface Difference*	*Index*	
	Before	After		Attribute change	Object reduction
MPU= 1000	425	386	29931.74	0.0061	0.92
Low	425	393	0.053	0	0.08

(4) comparison of the combination order

In order to check if there is any influence of the order of the operations, we generalize the same map (Figure 5 above) by two strategies in different combination orders. We apply the aggregation (class-driven) to Figure 6a at two thematic scales, i.e. dissolve the boundaries between adjacent polygons, which have same land use type at low level, and high level respectively. We obtained the results shown in Figures (9A, 9C) and (9B, 9D). Therefore, Figure 9 represent the results by combining firstly geometry-driven, then class-driven strategies.

Figure 10 is the generalization results of Figure8A and 8B after eliminate the polygons which are smaller than 2000. It represents the results obtained by first class-driven then geometry-driven strategies.

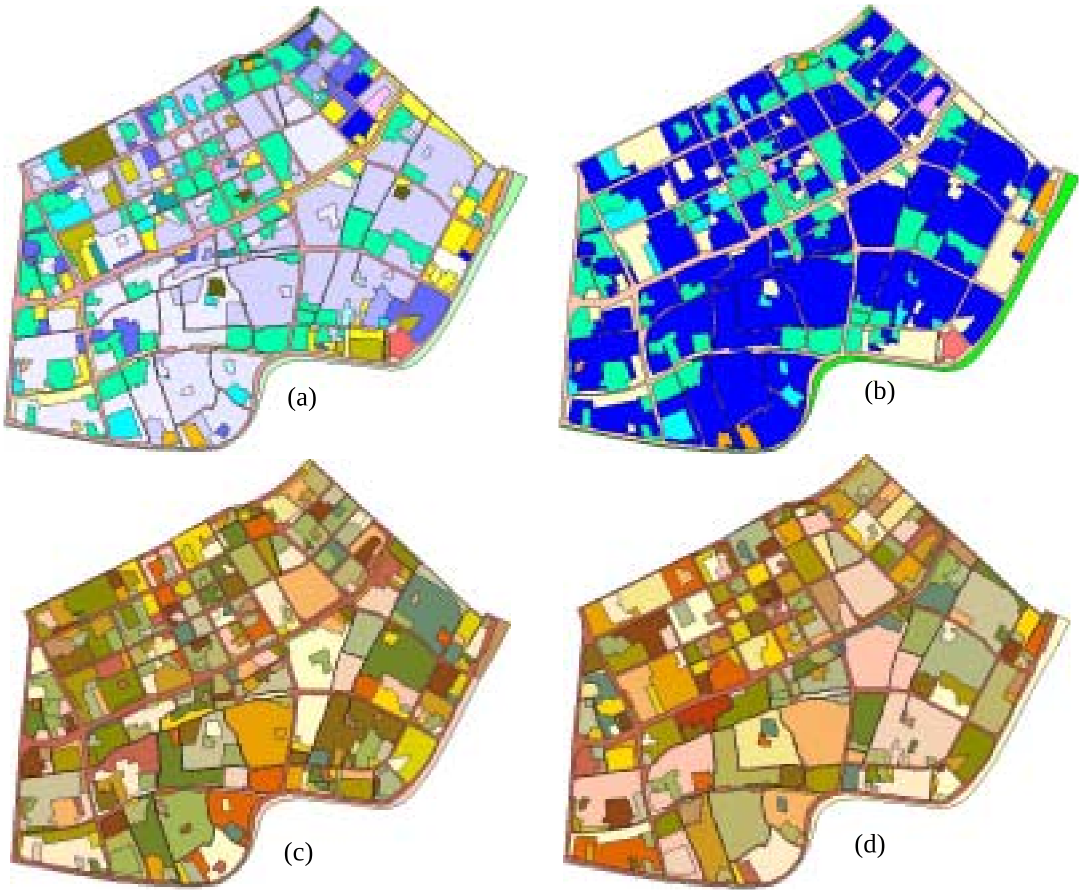

Figure 9 Results obtained by first geometry-driven then class-driven strategies.
(A) low level class type (B) high level class type
(C) low level objects (D) high level objects

Table 8 and Table 9 report the quality indices for two combinations at two thematic scales.

Table 8. Combination of the geometry-driven and class-driven strategies at low thematic scale.

	Number of Objects		*Difference* of Area	*Index of Attribute Change*	*Index of Object reduction*
	Before	After			
GC	425	303	133887.4	0.027	0.29
CG	425	304	127746.3	0.026	0.28

Table 9. Combination of the geometry-driven and class-driven strategies at high thematic scale.

	Number of Objects		*Difference* of Area	*Index of Attribute Change*	Object reduction
	Before	After			
GC	425	228	84275.9	0.017	0.463
CG	425	230	76486.4	0.016	0.459

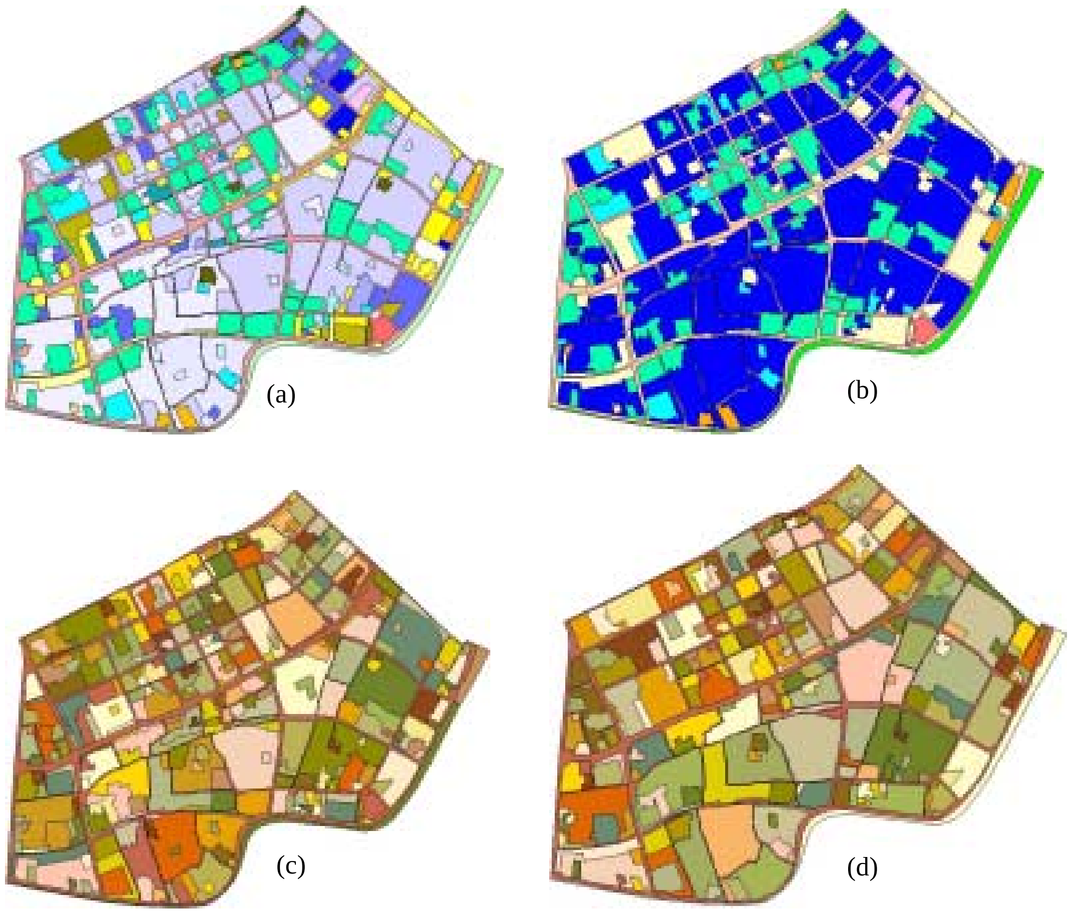

Figure 10. Results obtained by first class-driven then geometry-driven strategies
(A) low level class types (B) high level class type
(C) low level objects (D) high level objects

As shown in Tables 8 and 9, there is no obvious difference in the attribute change and object reduction between the two combinations, i.e. first apply the class-driven, then apply the geometry-driven; or firstly apply the geometry-driven, then apply the class-driven strategies. But the finial results between two thematic scales are quite different. It implies the combination of operations is quite sensitive to the thematic scale. But are there any differences between the finial results of the two combinations? We find they are lightly different in thematic values and much different in topologies, and the differences in these two aspects are larger at the high thematic scale than at the low thematic scale (see Figure 11 and 12).

(5) Comparison of Raster and Vector approaches

In a previous paper of the present author (Cheng & Lin, 2000), a raster model has been applied to study the uncertainty in the model-oriented generalization, because it is able to accommodate impression and uncertainty more easily than conventional vector (cartographic) models. Here we get a similar conclusion as we got before. As for geometry-driven strategy, the MMU influence the quality of generalization; as for the class-driven strategy, the thematic scale does not influence the quality of semantics but the change of topology. Our finding is in the same tune with Daley et al's (1997), i.e., there were no significant difference in area per class between the raster approach and the vector approach. Therefore, if the original data are raster-based, a raster-based generalization approach should be applied. Otherwise, a vector-based approach can be applied. But a raster-based approach does have advantage in certainty analysis (see also Beard & Buttenfield, 1999).

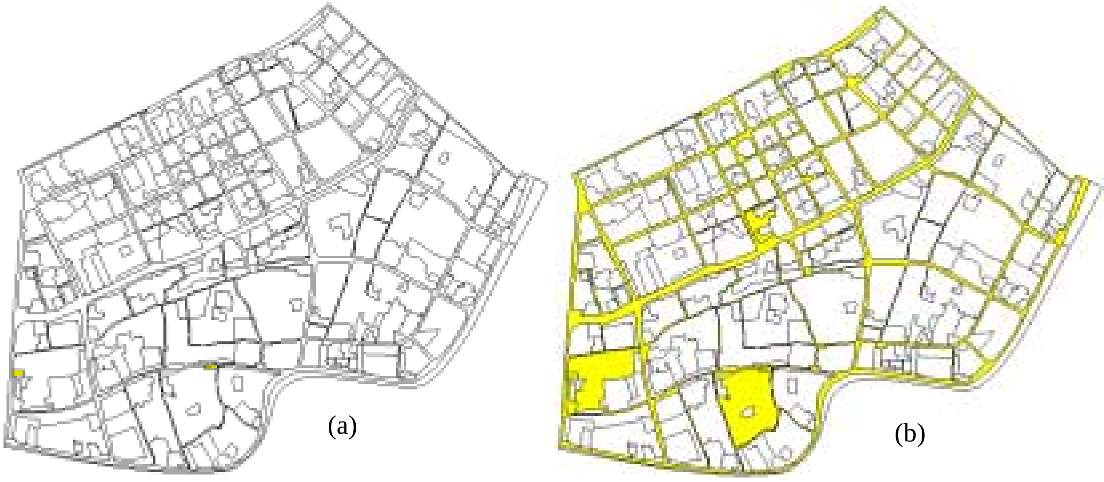

Figure 11. Difference between GC and CG at low thematic scale.
Figure A indicate 2 objects that have different thematic values; Figure B highlights 8 objects that have different topologies.

Figure 12. Difference between GC and CG at high thematic scale.
Figure A indicates that 10 objects have different thematic values; Figure B highlights 36 objects that have different topologies.

7. CONCLUSIONS

The objective of this paper was to develop mechanisms for quality assessment in the model-oriented generalization. To achieve that, we investigated the strategies of the model-oriented generalization, and analyzed the uncertainties in these strategies. Special attention is given to the uncertainty in semantics created in the merging and aggregation operations. We proposed the quantitative indices for holistic quality assessment of the generalization. The graphic visualizations of the quality were also discussed. Further, generalization of a land use map has been implemented to examine the geometry-driven and class-driven strategies. The generalization results have been thoroughly analyzed and discussed in order to check the influence of spatial resolution (the MMU) and the thematic resolution (thematic scale) on the qualities created by these two strategies.

It was shown in our case study that a proper scale and proper strategy should be chosen for a specific generalization case. In order to keep the thematic semantics of objects the class-driven strategy should be applied. However, a proper spatial scale (the MMU) should be selected for the geometry-driven strategy so as to preserve the geometric shape of objects and certainty of generalization. In order to achieve best visualization after generalization, the two strategies can be combined, e.g., the class-driven strategy is firstly used to derive the generalization result with low thematic uncertainty, and then the geometry-driven strategy is adopted to remove small regions.

As for further research, the other two strategies, i.e., the function-driven and structure-driven should be discussed to compare them with the existing results. What we may tell now is that the function-driven strategy is quite similar to the class-driven strategy. But the structure-driven strategy needs further effort because it deals with the structure, not the thematic attribute or pure metric in our case. Furthermore, we only discussed the situation in which the original source data have no uncertainty. It is necessary to study the situation where there is uncertainty in source data in further research. Moreover, we only discuss the operation of merging and aggregation. How to apply the quality assessment indices to other model-oriented operations needs more practice, although we believe they are general and applicable. In general, the study on quality assessment of generalization just begins; more efforts are required to open the fuzzy face of generalization.

Acknowledgements

The author wishes to thank Martien Molenaar for the stimulating discussions and Qinming Zhan for providing the data.

References

Beard, M. K. and Buttenfield, B. P., 1999, Detecting and evaluating errors by geographical methods, in *Geographical Information Systems,* Longley, P.A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (eds.), Second Edition, Vol. 1, 219–233.

Brazile, F., 1998, A generalization machine design that incorporates quality assessment**.** In *Proceedings 8th International Symposium on Spatial Data Handling*, Vancouver, pp. 349-360.

Berg M. De, Kreveld M van and Schirra, S., 1995, A new approach to sub-division simplification, *ACSM/ASPRS Annual Convention and Exposition 4* (*Proceedings AutoCarto 12*): 79-88.

Bregt, A. and Bulens, J., 1996, Application-oriented generalization of area objects, in Molenaar, M. (editor), *Methods for the Generalization of Geo-Databases*, Delft: Netherlands Geodetic Commission, New Series, Nr. 23, 57–64.

Cheng, T. and Lin, H., 2000, Uncertainties in multi-scale generalization, in *Proceedings of the 9th International Symposium on Spatial Data Handling (SDH2000),* Beijing, Aug. 10-12, edited by Forer, P., Yeh, A. G. O. and He, J., pp. 5a.18-31.

Cheng, T. and Molenaar, M., 1999, Objects with fuzzy spatial extent, *Photogrammetric Engineering and Remote Sensing*, 65(7), 797–801.

Daley, N., Goodenough, D.G., Bhogal, a. S., Bradley, Q., Grant, J. and Zin, Y., 1997, Comparing Raster and Object Generalization, In Proceedings IGARSS'97, Singapore, pp. 677-679.

Dettori, G. and Puppo, E., 1996, How Generalization interacts with the topological and metric structure of maps, *Proceedings of the* 7th *International Symposium on Spatial Data Handling,* edited by M. J. Kraak and M. Molenaar, P.559-570.

Ehrliholzer, R. 1995, Quality assessment in generalization: integrating quantitative and qualitative methods, *Cartography Cross Borders (Proceedings 17th International Cartographic Conference*), Barcelona, Institute Cartogràfic de Catalunya: 2241-50.

Harvey, F., Vauglin, F. and Bel Hadj Ali, A., 1998, Geometric Matching of Areas – Comparison Measures and Associated Links, *Proceedings of the 8th International Symposium on Spatial Data Handling*, T. K. Poiker and N. Chrisman (eds.), P.557-568.

Harvey, F. and Vauglin, F., 1996, Geometric match processing: applying multiple tolerances, *Proceedings of the* 7th *International Symposium on Spatial Data Handling,* edited by M. J. Kraak and M. Molenaar, P.155-171.

João, E. M., 1995, The importance of quantifying the effects of generalization, In Müller, J. C., Lagrange, J. P. and Weibel, R. (eds), *GIS* and *Generalization, Methodology and Practice,* London: Taylor & Francis, 183-193.

Molenaar, M., 1998, *An Introduction to the Theory of Spatial Object Modeling for GIS*, London: Taylor & Francis, 246p.

McMaster, R. B., 1987, Automated line generalization, Cartographica 24:74-111.

McMaster, R. B. and Shea, K. S., 1992, *Generalization in Digital Cartography*, Association of American Geographers, 134p.

Müller, J. C., Lagrange, J. P. and Weibel, R., 1995a, *GIS and Generalization, Methodology and Practice, GISDATA I*, London: Taylor & Francis, 257p.

Müller, J., Weibel, R, Lagrange, J. P. and Salge, F., 1995b, Generalization: state of the art and issues, in Müller, J. C., Lagrange, J. P. and Weibel, R. (eds.), *GIS* and *Generalization, Methodology and Practice*, London: Taylor & Francis, 3–18.

Painho, M., 1995, The effects of generalization on attribute accuracy in natural resource maps, In Müller, J. C., Lagrange, J. P. and Weibel, R. (eds.), *GIS and Generalization, Methodology and Practice*, London: Taylor & Francis, 194–206.

Richardson, D. E. and Mackaness, W., 1999, Guest Editors, Computational Methods for Map Generalization, *Special Issue on Map Generalization in collaboration with the ICA Working Group on Map Generalization, Cartography and Geographic Information Science, Journal of the ACSM*, Vol.26, No.1.

Veregin, H. and Hargitai, P., 1995, An evaluation matrix for geographical data quality. In *Elements of Spatial Data Quality*, S.C. Guptill and J. L. Morrison, (eds.), Oxford: Elsevier, 167-188.

Weibel, R., 1995, Guest Editor, Map Generalization, *Special issue*, *Cartography and Geographic Information Systems*, Journal of the ACSM, Vol.22, No.4, 1995.

Weibel, R. and Dutton, G. H., 1998, Constraint-based automated map generalization, *Proceedings of International Symposium on Spatial Data Handling*, T. K. Poiker and N. Chrisman (eds.), P.214-224.

Weibel, R. and Dutton, G. H., 1999, Generalising spatial data and dealing with multiple representations, in Longley, P.A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (eds.), *Geographical Information Systems*, Second Edition, Vol. 1, 125–155.

Weibel, R. and Jones, C. B., 1998, Computational Perspectives on Map Generalization, *GeoInformatica*, 2(4), 307–314.