CONSTRAINT-BASED GENERALIZATION OF SOIL MAP*

GAO Wenxiu, SONG Aihong, GONG Jianya

National Laboratory of Surveying, Mapping and Remote Sensing, Wuhan University
129, Luoyu RD. Wuhan, China, 400079
Email: wxgao@mail.liesmars.wtusm.edu.cn

Commission VI, WG VI/4

KEYWORDS: Constraint-based Generalization, Specific constraints, Generic constraints, Soil map generalization, Integration of constraints

ABSTRACT:
Soil map that shows the geographical distribution regularity of different soil types and the soil resources diversity is an important type of categorical maps. And soil map is applied in various professional studies such as vegetation studies, land use and environmental studies, and resources management. Therefore, soil map generalization not only follows the generic principles of cartographic generalization, but also is supervised by the professional principles of soil distribution. This paper explores the professional principles of soil distribution according to the regularity of soil distribution, the characteristics, representation and application of soil maps as specific constraints to soil map generalization. The specific constraints are indispensable for the comprehensive generalization of soil map. Then associating with the research of Peter and Weibel (1999) on generic constraints to generalization of categorical data, the specific constrains and generic constraints are integrated and applied in the different phrases of the generalization process to investigate the functionalities. The last component of this paper discusses two special issues of soil map generalization to demonstrate the importance and application of specific constraints to soil map generalization.

1. INTRODUCTION

Soil data is one of the important components to portray the natural environment and natural resource in agriculture, Soil map, which includes map of soil parent materials, soil category map, map of soil chemical types, map of soil texture and so on, is one of the important types of land resource map and agriculture map. Soil category map represents soil categories, soil geographic distribution regularity and diversified soil resources in natural environment. In many literatures, soil category map is just called soil map (Pan *et al.* 2002; Lu *et al.* 2001; Huang et al 1989). In the paper, we use soil map instead of soil category map to discuss the constraints of soil map generalization.

According to the purpose and application of soil map, soil map consists of generic soil map and specific soil map. The former one represents the generic characteristics of soil, while the latter one focuses on one or more specific characteristics of soil, such as the map of soil engineering, the map of forest soil, the map of soil erodibility (Huang *et al.* 1989). Chinese Soil Taxonomy (CST) (revised proposal, 1995) is employed as the base of soil mapping, which is a multiple category classification including six categories, namely soil order, soil suborder, soil group, soil subgroup, soil genus and soil series according to the order from higher category to lower one. Soil maps at different scale represent different level of soil category. The soil maps with the scale 1□50000 or more than 1:50000 called large-scale soil map describe the soil categories in soil genus level, which represent the detailed soil distribution in small-domain terrain environments (Lu *et al*. 2001). The middle-scale soil map, 1:50000~1:500000, and the small-scale soil map, smaller than 1:500000, could describe the soil categories in the soil group or soil subgroup (Huang *et al.* 1989), which represent the soil distribution regularity in larger geographic regions from the macro view of point (Lu *et al*. 2001). One of the functions of soil map generalization is to derive smaller-scale map from larger-scale map.

Furthermore, the purpose and application of soil map decide the level of detail representation of soil categories in the map, and they also decide the representation and completeness of a feature. For example, the categories of a soil map (called Map A), which is used to discern the adaptability of soil for the plantation of vegetations or the reasonable layout of crop plantation, must be more complete than that of a soil map (called Map B) which is 4as a reference map to partition the agricultural districts (Huang *et al.* 1989, Liao 1991). Map B could be derived from Map A, which could just generalize the content of the map without any change of the map scale. This is the another important function of soil map generalization.

In brief, soil map generalization could be employed in the two situations: one is to derive middle-scale or small-scale soil maps from large-scale soil maps; the other one is to derive soil maps with generalized contents from soil maps with complete contents. The derived maps in the first situation represent the soil distribution regularity in a larger geographic region than that of the original maps, whereas the derived maps in the second situation represent the coarser characteristics of soil distribution in a same geographic region as that of the original maps. Whatever kind of situations, soil map generalization should be implemented based on generalization constraints to steer generalization operations and derive valid products.

Our study focuses on the constraints involved in soil map generalization, especially specific constraints based on the

* Study supported by the Open Research Foundation of the Key Laboratory for Information Engineer in Surveying, Mapping and Remote Sensing (LIESMARS) (No. (02)0401) and 973 Program (No. 2003CB415205)

regularity of soil distribution, the characteristics, representation and application of soil maps. Section 2 designates specific constraints in the phases of the generalization process. To derive appropriate and correct maps, all kinds of generalization constraints should be integrated efficiently in the process of generalization. Therefore, in Section 3, we investigate the integration of different kinds of generalization constraints and their relationships. Then, two special cases in soil map generalization are illustrated in Section 4, and a workflow shows the whole process of soil map generalization with the application of generalization constraints.

2. CONSTRAINTS TO SOIL MAP GENERALIZATION

Constraint-based generalization of geographic map has been studied by many researchers (Beard 1991; Ruas 1998; Weibel and Dutton 1998; Peter and Weibel 1999; Gadland 2003). The concept of constraints used in the paper owes to previous research conducted by Peter and Weibel (1999) and Ruas (1998). Constraints designate the specification of final map or database based on properties of a geographic phenomena or feature. And constraints are used to detect conflicts, to control the sequence or strategy of generalization, to compare accomplished solution and evaluate the generalized result. Generally, constraints are defined as various specifications which control the specific aspects of an object, a group of objects or a whole map in the process of map generalization. Take the research in Peter and Weibel (1999) as an example to describe the constraints of map generalization. In their research, the generalization constraints are classified into: graphic constraints, topologic constraints, structural constraints and Gestalt constraints.

- **Graphical constraints** define graphic perceptibility thresholds of map objects based on human limits of perception, such as minimal area, minimal distance between two polygons.
- **Topological constraints** deal with basic topological relationships such as connectivity, adjacency and containment, which should be maintained when generalizing data.
- **Structural constraints** include spatial structural constraints and semantic structural constraints. Spatial structural constraints mainly are responsible for the preservation of typical shapes of individual map objects or patterns and alignments of a group of map objects. Semantic structural constraints deal with the preservation of the logical context of patches.
- **Gestalt constraints** are related to aesthetic aspects for the preservation of the patch characteristics as well as the retention of the overall visual balance when multiple patches or the whole dataset is considered.

Based on the classification system of constraint, Peter and Weibel (1999) investigate the constraints to categorical data generalization. Table 1 shows some titles of these constraints relate to soil map generalization.

Constraints Related to Patches	Constraints Related to Categories	Constraints Related to Partitions or Groups of Patches
Minimum size (graphical) *Minimum distance* (graphical) *Separation* (topological) *Self-intersection* (topological) *Amalgamation* (structural) *Shape/Angularity* (structural)	*Size ratio* (structural) *Shape/Angularity* (structural) *Alignment / Pattern* (Gestalt) *Aggregability* (structural)	*Neighborhood relations* (topological) *Spatial context* (structural) *Alignment/Pattern* (Gestalt) *Visual balance* (Gestalt)

Table 1 Constraints to Categorical Data (after Peter and Weibel (1999))

Soil data is one of the important types of categorical data. These constraints should therefore be taken into account in soil map generalization (it is called generic constraint in the paper). In our paper, we do not discuss these constraints in detail any more, but investigate the specific constraints of soil map generalization (it is also called thematic constraints) based on the regularity of soil distribution, the characteristics and application of soil maps. The following section discusses the thematic constraints based on the four main phrases in the process of soil map generalization: preparation and preprocess, database generalization, graphic generalization and evaluation of generalization results.

(1) Preparation and preprocess of soil map generalization

Obviously, the scale of the objective map should depend on the scale of the original soil map or the resolution of the present soil database. Furthermore, some more factors should be taken into account to confirm the scale such as the regularity of soil distribution in a specific geographic region or area, the purpose of the objective map.

- ***Constraint 1:*** *the scale of the objective map is appropriate to represent the dominating regularity of soil distribution in a specific geographic region.* In general, there are three kinds of main regularity of soil distribution: the regularity of horizontal distribution of soils, the regularity of vertical distribution of soils, the regularity of zonal distribution of soils. The first two regularities that are influenced by hydrothermal conditions and large geomorphology represent the geographic distribution of soil in large zones. The third regularity that is formed based on small or middle geomorphology, thermal condition, parent materials and human activities represents the geographic distribution of soil in small zones. Thus, the first two regularities are appropriate to be represented in middle-scale or small-scale maps (generally, the scale should be smaller than 1:300000), whereas the third one could be represented effectively in large-scale or middle-scale map (generally, the scale should be more than 1:100000).
- ***Constraint 2:*** *The scale of the objective map should meet the requirements of the purpose of the map.* If the objective map is used as the reference for the planning of anthropogenic soil and the planning of soil amelioration, the scale should be between 1:10000 and 1:50000. If the objective map is used in the statistics of soil resources, 1:500000 or smaller is appropriate for the whole nation or a province, and the scales between 1：100000 and 1：300000 are better for a county.
- ***Constraint 3:*** *the level of soil categories described in the objective soil map should be in harmony with the scale.* Based on CST, the large-scale soil maps describe the soil genus level, and the middle-scale and

small-scale soil maps could describe the soil group or soil subgroup (Huang *et al.* 1989).

- ***Constraint 4:*** *the level of soil categories described in the objective soil map should meet the requirements of the purpose of the objective map.* For example, a soil map used in the planning of anthropogenic soil should describe soil categories in detailed level such as soil genus. The soil map used as reference of the partition for the agricultural districts could just describe coarse soil categories such as the soil group or soil subgroup.

(2) Database Generalization

One of the main steps in database generalization is attribute transformation. For soil data generalization, the attribute transformation is to transform soil categories of soil parcels from the lower level to the higher level. A special representation, soil complex, is used to represent soil distribution in some special cases. For example, various soil genus intermix in a small geographical region makes their boundaries impossible to be identified even in large-scale map. Therefore, soil complex is used to represent the complex components in one single soil parcel with the associated soil types and their area percents are recorded too. The premise of the soil complex representation is the area percent of each soil type is not more than 15% of the total area of the soil parcel.

- ***Constraint 5:*** *emphasis the soil categories closely relate to the purpose of the objective map.* For example, if the objective map will be used in the planning of anthropogenic soil, the anthropogenic soil parcels could be described in more detailed categories than other kinds of soil parcels.
- ***Constraint 6:*** *one-way transformation*. The category of soil only can be transformed from the lower level to the higher one in CST.
- ***Constraint 7***: *semantic consistency between the new category and the old one of a soil parcel.* That is to say, the soil genera "Siallitic lateritic red earths" should be converted into the soil group "Lateritic red earths" rather than the soil group "latosols".
- ***Constraint 8***: *category transformation of soil complex*. A soil complex involves more than one soil category. Firstly, transform these soil categories to their corresponding higher categories. After transformation, the number of the soil categories in the soil complex maybe reduced because some categories may belong to a same higher category. Secondly, recalculate the area percent for each unique soil category. Thirdly, if the area percent of one category is more than 75%, the soil complex should be converted into a simple soil parcel with the soil category.

(3) Graphic Generalization

Graphic generalization solves graphical conflicts caused by the reduced space on a map or simplifies the graphic information by eliminating the unimportant or unnecessary details.

- ***Constraint 9:*** *Preserve distribution pattern of soil zonal discontinuity in large-scale or middle-scale soil map*.
- ***Constraint 10:*** *Preserve the characteristics of regularly continuous distribution in small-scale map*.
- ***Constraint 11***: *soil complex representation.* In our study, we employ soil complex representation to solve some graphical conflicts caused by small-area soil parcels which are adjacent each other and the sizes are smaller than the minimal area. The soil types of these soil parcels are associated in the soil complex and the area percent of each soil type is recorded too.
- ***Constraint 12***: *different simplification operators employed on parcels of natural soil and anthropogenic soil*. Anthropogenic soil, such as siltigic soil, athrostragic soil, paddy soil, is formed in the process of long term utilization and reclamation by human (Gong *et al.* 1997). Generally, the boundaries of natural soil parcels are more complex than that of anthropogenic soil. They should be simplified with different simplification operators to preserve the difference.

(4) Quality evaluation of the generalization results

Generally, constraints are involved in more than one phase in the whole process of generalization. And some constraints (e.g. minimum size, shape/angularity, alignment/pattern) are not only used to measure conflicts and select operators, but also evaluate the result after operators running. Most of the thematic constraints and generic constraints are also used as the standards to evaluate the validity of the generalization results (Figure 2).

3. INTEGRATION OF THEMATIC CONSTRAINTS WITH GENERIC CONSTRAINTS

Most constraints are contextually related and affect each other, and they should be integrated into the whole generalization process (Peter and Weibel 1999). Figure 1 demonstrates the integration of thematic constraints with generic constraints in the different phrases of the generalization process. The both kinds of constraints steer the generalization operations in different levels and views. Thematic constraints (e.g. Constrain 1, 2, 3, 4, 5, 9, 10) mainly portray the status or characteristics of soil map which should be reached or preserved during generalization, whereas generic constraints focus on some concrete graphical characteristics. As a whole, thematic constraint is the foundation and involved in every phase of the generalization process. Some generic constraints need to set the final value of parameters of relative operators based on thematic constraints. Meanwhile, the generic constraints support thematic constraints by adjusting appropriate operators and parameters.

(1) Preparation and preprocess

The objective scale and the levels of categories represented in the objective soil map are set according to Constraint 1, 2, 3, 4. Then the distribution pattern of soil which should be represented in the objective scale need to be explored and designated based on Constraint 9, 10. After that, the minimal size (e.g. minimal area), minimal distance can be defined and the size ratio of each category, typical shape/angularity and alignment/pattern need to be calculated and measured. Generally, the minimal area of soil parcels belonging to detailed categories is smaller than that belonging to coarse categories (Weng 1997). In large-scale maps, the minimal area should be smaller for the preservation of soil zonal discontinuity, but in small-scale maps, the minimal area is larger to represent the characteristics of regularly continuous distribution by eliminating the zonal small parcels.

(2) Graphic Generalization

Constraint 9 and 10 describe requirements and specifications of the objective soil map. They are involved in the phrase of preparation and preprocess, but also steer the solution of graphic generalization. Take the solution for solving conflicts caused by soil parcels smaller than the minimal area as example. For Constraint 9, the sequence of appropriate operations for the conflicts is enlargement, aggregation, amalgamation, collapse; but for constraint 10, amalgamation, collapse is more appropriate in terms of the efficiency of implementation. The generic constraints govern the implementation of these operators. For example, minimum size and self-intersection are used to avoid invalid or incorrect geometric conflicts; shape/ angularity, size ratio, alignment/pattern and neighborhood relations work for preserving the properties of soil distribution.

(3) Quality Evaluation

The generic and thematic constraints are used as the standards in the evaluation of generalization results. The thematic constraints mainly evaluate the quality of the result map from a macro view of point, whereas the generic constraints focus on the detailed aspects of geometric graphics.

Preparation and Preprocess

Thematic Constraints	Generic Constraints
Constraint 1, Constraint 2 Constraint 3, Constraint 4 Constraint 9, Constraint 10	Minimum size, Minimum distance, Size ratio, Shape/Angularity, Alignment/Pattern

Graphic Generalization

Thematic Constraints	Generic Constraints
Constraint 9 Constraint 10 Constraint 11 Constraint 12	Minimum size, Minimum distance, Separation, Self-intersection, Amalgamation, Shape/Angularity, Size ratio, Alignment/Pattern, Neighborhood relations, Spatial context, Aggregability, Visual balance

Quality Evaluation

Thematic Constraints	Generic Constraints
Constraint 3 Constraint 4 Constraint 5 Constraint 6 Constraint 7 Constraint 9 Constraint 10	Minimum size, Minimum distance, Separation, Self-intersection, Shape/Angularity, Size ratio, Alignment/Pattern, Neighborhood relations, Spatial context, Visual balance

Figure 1 The Integration of Thematic Constraints with Generic Constraints

4. IMPLEMENTATION OF SOIL MAP GENERALIZATION

Section 2 investigates the main thematic constraints in different phases of generalization process. Then section 3 analyzes the relationship and integration between thematic constraints and generic constraints. In this section, two special problems about soil map generalization will be discussed to further illustrate the importance of thematic constraints and the integration with generic constraints.

(1) Transformation of soil category

Transformation of soil category is to transform soil categories of soil parcels from the lower level to the higher one. We employ Chinese Soil Taxonomy (CST) as the base of soil mapping, in which the order from higher level to lower level is soil order, soil suborder, soil group, soil subgroup, soil genus and soil series, respectively.

- **Generalization based on paper map:** More than one level of soil category are represented in many paper maps to portray soil distribution (e.g. The Soil Atlas of China 1986, The Land-resources Atlas of Beijing 1990, The Atlas of The Changjiang River Basin 1999). For example, the soil map (1:600000) in the Land-resources Atlas of Beijing portrays soil group or subgroup by filling-color method and marks numbers on each colored soil parcel to distinguish different soil genus in a same soil group or subgroup. The basic soil parcel represents soil genus. When the map is used as the base map to produce a new map with smaller scale, the category level of soil genus is neglected and the soil parcels are merged with their neighbouring parcels which belong in the same soil group or subgroup. In the final map, the basic soil parcel represents soil group or subgroup.
- **Generalization based on soil database:** The strategy of generalization varies with the different database. Zhang *et al.* (2003) introduces the soil classification coding system of Chinese soil database at the scale of 1:1000000. The soil category can be encoded with 7 digitals. Soil order, soil group and subgroup are represented by the first, the second and the third digital respectively. And soil genus and the code of soil profile are represented by the combination of the fourth and fifth one, and the last two digitals respectively. The format of soil category code contains potentially the classification system of soil category. Therefore, the soil category can be easily transformed from higher category to lower one by omitting the latter digitals. If soil database does not involve any classification system of soil category, a classification system should be built first based on the database and the purpose of the objective map. Then the soil category can be transformed from lower level to higher level according to Constraint 5, 6 and 7. After generalization, the basic soil parcels in the derived database represents the lower categories.

(2) Generalization of soil complex

Soil complex is a special representation in soil map. Therefore, the generalization of soil complex is a special issue which is not involved in the generalization of other kinds of data or maps. The category transformation of soil complex has been discussed in Constraint 8 in detail. There are two situations relate to soil complex during the phrase of the graphic generalization. The first one mentioned in Constraint 11 associates several simple soil parcels into one new soil complex to solve graphic conflict. The second situation is to solve the conflicts caused by the present soil complex in the map or database. For example, one soil complex is smaller than the minimal area. There are some alternatives to deal with it. The first one is to associate the soil complex with its neighbouring soil parcels into one new soil complex with more soil categories if its neighbouring soil parcels are also smaller than the minimal area. The second one

is to collapse the soil complex and split its area to its neighbouring soil parcels. The second alterative is appropriate to the situation that none of its neighbouring soil parcels is smaller than the minimal area.
In brief, new generalization solutions need to be developed for soil complex in order to represent the special contextual characteristics of soil map.

5. CONCLUSION

In the generalization of categorical data or map, the special characteristics, requirements and applications should be considered to represent the professional principles of concrete categorical data. And they need to be converted into specific constraints and integrated with generic constraints imposed by cartographic principles to govern and steer the generalization process of concrete categorical data. This paper investigates beneficial attempts on the generalization of soil map, especially with the specific constraints. We aim to explore the specific constraints of soil data and map, analyze the integration of specific constraints and generic constraints. The integration provides the fundamentals for the development of comprehensive strategies for the generalization of soil data and map. However, further researches, such as the appropriate representation of specific constraints, generalization solution about soil complex, are necessary to develop a fully operational constraint-based generalization system to implement the soil data and map generalization.

REFERENCE:

Chinese Soil Taxonomy Research Group, Institute of Soil Science, the Chinese Academy of Science, 1995, Chinese Soil Taxonomy (Revised Proposal), Beijing: Chinese Agricultural Scientific Publishing House

Beard, M.K. 1991. Constraints on Rule Formation, in Buttenfield, B.P., and McMaster, R.B.,(Eds) Map Genralization: Making Rules for Knowledge Representation, London: Longman , pp. 121-135

Galanda, M.2003, Automated Polygon Generalization in a Multi Agent System, Ph.D. thesis, Department of Geography, University of Zurich, Switzerland

Gong Zitong, Chen Hongzhao, Wang Helin, 1997, The Distribution Regularity of Higher Categories in Chinese Soil Taxonomy, Chinese Geographical Science, Vol. 7, No. 1, pp.1-10, Science Press, Beijing, China

Huang Wanhua, Guo Yuxiao, 1989, The Application of Map in Agriculture, Surveying and Mapping Public House, Beijing, China

Liao Ke, Fu Xiaoxing, 1991, The Corpus of Agricultural Mapping, The Science Publish House, China

Liu, Y., 2002, Categorical Database Generalization in GIS, Ph.D. dissertation, ITC

Lu Chengwen, Gu Yeping, Liu Fucheng, Wei Xiang, Jia Hongjun, 2001, The application of classification system of soil in large-scale soil mapping, Soil (1)

Muller, J.C., Wang, Z., 1992. Area-Patch Generalization: A Competitive Approach. The Cartographic Journal, **29**(2), pp.137-144

Peter, B., and Weibel, R. 1999. Using Vector and Raster-Based Techniques in Categorical Map Generalization, Third ICA Workshop on Progress in Automated Map Generalization, Ottawa, 12-14 August

Pan Xianzhang, Shi Xuezheng, 2002, The Research on the Technology of Digital Mapping of Soil Quality, Soil, Vol. 3

Ruas, A.1998. OO-Constraint Modeling to Automate Urban Generalization Proceeds, Proceedings of the 8th International Symposium on Spatial Data Handling, Vancouver, BC, 6b, pp.225-235

The Soil Atlas of China, 1986, Cartographic Publish House of China

The Land-resources Atlas of Beijing, 1990, Planning Committee of Beijing, China

The Atlas of The Changjiang River Basin, 1999, Cartographic Publish House of China

Weibel, R., Dutton, G.H.1998. Constraints-Based Automated Map Generalization. Proceedings of the 8th International Symposium on Spatial Data Handling, Vancouver, BC, pp.214-224

Weng Rendi, 1997, Several Technological Issues on the Mapping of the Soil Atlas of Fujian, Journal of Fujian Agricultural University, Vol. 26, No. 6

Zhang Dingxiang, Pan Xianzhang, Shi Xuezheng, Yang Jinlin, Lin Jie, 2003, Several Remarks on Construction of Chinese Soil Database at the Scale of 1:1000000, Chinese Journal of Soil Science, Vol.34, No.2