WHAT IS SPATIAL CONTEXT IN CARTOGRAPHIC GENERALISATION?

Sébastien Mustière, Bernard Moulin

Computer Science Department and Geomatics Research Centre, Laval University, G1K 7P4 Ste-Foy, Québec, Canada
{sebastien.mustiere, bernard.moulin}@ift.ulaval.ca

KEY WORDS: Databases, Contextual, Analysis, Modelling, Cartography, GIS, Generalization.

ABSTRACT:

During the cartographic generalisation process, geographic objects cannot just be considered one by one. The way objects are processed clearly depends on their spatial context. In this paper, we first study the nature of spatial context encountered during map generalisation. We differentiate three kinds of relations that an object can have with its environment: being part of a significant group, being in a particular area, being in relation with 'same-level' surrounding objects. We also address the issue of scale-dependency related to spatial context. Then, we study how to represent context-related knowledge in a knowledge-based approach of cartographic generalisation. We discuss several aspects of the representation of objects' spatial context in a geographic data model dedicated to map generalisation: explicit representations of high-level objects, fuzzy representation of objects and relations, multi-scale representation of objects and relations, spatialisation of relations. We also study how context appears in the rules expressed by cartographers to describe the generalisation process: it can be used to express exceptions, to classify typical operations to be done, or to express constraints on the final result.

RÉSUMÉ:

Pendant le processus de généralisation, on ne peut pas considérer les objets géographiques un à un. Le traitement de ces objets dépend clairement de leur contexte spatial. Dans cet article, nous étudions tout d'abord la nature du contexte spatial qui influence la généralisation. Nous différencions trois types principaux de relations qu'un objet peut avoir avec son environnement: faire partie d'un groupe significatif, être dans une zone particulière, ou être en relation avec les objets aspects de même niveau. Nous abordons également un aspect important du contexte spatial: sa dépendance au niveau d'analyse. Nous étudions ensuite comment représenter les connaissances relatives au contexte spatial dans une approche à base de connaissances de la généralisation. Nous présentons diverses considérations sur ce que nécessiterait une représentation du contexte spatial dans un modèle de données adapté à la généralisation: représentation explicite d'objets de haut niveau., représentation floue d'objets et de relations, représentation multi-échelle d'objets et de relations, spatialisation des relations. Nous étudions également comment la notion de contexte apparaît dans les règles de généralisation exprimées par les cartographes: cela peut servir à exprimer des exceptions, à organiser différentes opérations typiques à réaliser dans différentes configurations, ou à exprimer des contraintes sur le résultat final.

1. THE IMPORTANCE OF SPATIAL CONTEXT

Geographic data models usually explicitly represent a set of basic objects, their geometry and their properties. But much of the geographic world's semantics appears in the relations linking objects [Worboys 96; Papadias and Theodoritis 97; Ruas 99; Mark 99]. Nevertheless, most of these relations are not explicitly represented in data models describing geographic databases. Usually, these relations only implicitly appear when one is looking at a *display* of a geographic database. This largely contradicts a principle in knowledge representation that a good representation should make the important things explicit [Winston 84].

This deficiency of geographic data models is a barrier to the analysis and derivation of geographic databases, and this is particularly the case for the cartographic generalisation process [Lagrange and Ruas 94]. Cartographers know that two similar objects located at two different places will not necessarily be generalised in the same way, even for the same target product. This is due to the (implicit) different relations that these objects have with other surrounding objects. "For example, an object can be preserved because it is a representative of a set of objects of the same nature, or because it allows a connection to objects which should be preserved […]" [Lagrange et Ruas 94]. In other words, the generalisation process is dependent on the objects' *spatial context*. Even more, "Context-related rules are probably the most significant of the 'know-how' for generalization" [Kilpeläinen 2000].

In the next section, we study the nature of spatial context encountered during map generalisation: what kinds of spatial context do exist; how does scale influence the notion of context? Then, we make different considerations on what is needed to represent objects' spatial context in a geographic data model dedicated to map generalisation.

2. DIFFERENT TYPES OF SPATIAL CONTEXT

When one reads a map, the geographic space is analysed according to different levels of analysis, from the identification of individual elements to an apprehension of the whole space, through the analysis of groups of objects. This multilevel aspect may be one of the most significant character of the geographic space [Scholl et al. 96].

This consideration is at the basis of our study of what is spatial context or, in other words, how an object is in relation to its environment. We first distinguish relations between objects of the same level of analysis, and relations between objects of different levels of analysis.

2.1 Hierarchical relations

We can find two main types of spatial relations implying geographic objects of different levels of analysis: the relations between a group and its members (e.g. an island is part of an archipelago), and the relations between a part of the space and its elements (e.g. a road is in a sparse area).

2.1.1 Being part of a significant group. A lot of geographic objects take a more precise meaning by being part of groups than on their own. For example, in certain studies an highway interchange may be more significant than the isolated road sections it contains. Certain generalisation frameworks explicitly represent significant groups and manage the interconnected generalisations of the groups and the parts. The identification of groups is seen as a "structure recognition" in the model of [Brassel and Weibel 88], and as a "meso analysis" in the AGENT model [Ruas 99; AGENT 00].

The identification of significant groups is very important during the generalisation process for several reasons [Ruas 00]. First, some generalisation operations cannot be done while looking at objects one by one and must be performed at the group level (such as merging and typifying). Second, the fact of being part of a significant group can influence the way the elements of the group are generalised (for example to insure a certain homogeneity of transformation in the group).

2.1.2 Being inside a particular area. This type of spatial context considers the property of being located inside an area that can be qualified by some global characteristics. For example a road can be inside a urban, rural, or mountainous area; or an island can be either in the middle of the see or in a coastal area.

The property of belonging to a particular area does influence the generalisation operations. It allows to better characterise the importance of objects. Consequently, it is possible to better determine which objects must be represented on the map, and how they must be represented. First, objects that are typical of an area must receive a particular attention during generalisation to efficiently reflect the global character of the area (e.g. the organisation of a street network must be well represented while selecting streets in a urban area). Second, another interest of characterising areas is to detect atypical objects which are considered as important and must be kept (e.g. a river in a desert area, an isolated village in a natural area, etc.).

2.1.3 Distinction between "member/group" and "element/part of space". The frontier between these two kinds of relations seems to be a bit fuzzy. In fact, being an element of a certain part of space can be thought of as being a member of the group made of all the elements located inside this part of the space. There is anyway two main differences between these relations, in terms of the nature of the relation, and in terms of its function.

First the "member/group" relation is a constituting relation (the group is *made of* the members). But the "element/part of space" relation is only a relation of localisation (the element is *inside* the part of the space). In addition, the "element/part of space" relation is usually more fuzzy than the "member/group" relation, because of the fuzziness of the limits of a given area (e.g. what are exactly the limits of a given urban area?).

Second, during the generalisation process, it may be important to know the exact constituents of a group (e.g. to merge individual islands into a unique object archipelago). But it is not important to know the exact limits of a part of the space, it is only important to know that a given element is inside a given part of the space (e.g. rules determining which individual houses to represent on the map may be different in rural and urban areas).

Anyway, let us notice that certain geographic phenomena may some time be viewed as a part of the space and as a group of objects during the generalisation process. For example, a urban area can be seen as the set of its houses when one wants to merge all the houses into a single polygon representing the urban area, or as a part of the space when one deals with the selection of the streets it contains.

The study of the relations between parts and wholes is a field of research in itself (namely mereology [Simons 87]), and we will not discuss it further on. Nevertheless, due to the multilevel nature of the geographic space, more studies should be done to better understand the semantics of the hierarchical relations in the geographic domain.

2.2 Non hierarchical relations

Another type of spatial context that influences the generalisation process corresponds to the local relations that a given object has with its surrounding objects at the same level of analysis (e.g. a house is *near* a road, a house is *aligned* with a road, a road *crosses* a river, a road is *parallel* to a railway, etc.). These relations can influence the generalisation process in different ways.

First, some of these relations must be kept. For example, if a road is *parallel* to a river or if a road is *crossing* a river, it may be important to keep these relations during the generalisation process.

Second, some local relations can be emphasised. For example, some close but distinct houses may be shifted from others to emphasise the *disjoint* relation. This example shows that generalisation can be seen as the process of *representing* an *abstracted* view of the world [Mustière et al. 00]. An abstraction is done when one considers that the disjoint relation is more important than the accurate location of the houses; and shifting houses is a way to represent this disjoint relation while respecting legibility constraints. Other examples where relations between objects are abstracted before being represented may be a house *nearly aligned* to a road that is moved to become *exactly aligned* with the road, or a house *close* to a road that is moved to *touch* it on the map.

It is then important to understand which types of relations may exist between objects in order to understand how they do influence the generalisation process : which relations must be kept, into which relations a given relation can be abstracted, etc.?

In the GIS community, the most studied relations are certainly topology-based relations (e.g. the 9-intersections model of [Egenhofer et al. 89]), followed by distance-based and direction-based relations. It is also often considered that "topology matters, metrics refines" [Egenhofer and Mark 95]. If topologic relation are undoubtedly important, one can moderate the predominance of topologic relations over other relations. First, a lot of relations rely on the notion of shapes and relative positions (e.g. surrounding, being parallel, being aligned, being a landmark) [Mathet 2000]. Quite few works do consider these notions, maybe because of the difficulty of identifying them contrary to topologic relations which are deterministically defined from the geometry of objects. Second, topologic relations such as adjacency must often be refined in order to better reflect significant spatial relations. "Defining connectivity solely on the basis of topological adjacency is inadequate for studying many types of natural and social processes […].For example, water basins (represented by polygons) are functionally connected only if they are adjacent and water flows from one into another" [Theobald 2001].

It is important to go further the analysis of topology to better understand the semantics of spatial local relations. The study of the way people naturally express these relations may be a good starting point for doing it. Several models have been developed in linguistic studies to represent the cognitive principles lying under the expression of spatial relations in the natural language. Certain authors develop models where topology is the central notion and may be extended with various concepts. For example, [Sablayrolles 95] considers a second "frontier" of the objects as a key notion: the limit of a "proximity area". Other authors go beyond notions related to topology. For example, [Mathet 00] considers that the notions of shape, distance, direction, topology and projection are all important. He also considers that the polymorph aspect of an object is a key issue to understand spatial relations: a given object is either seen as a line, a surface or a point according to the type of relation being considered.

In order to better understand how to manipulate spatial relations during the generalisation process, more attention could be paid to the development of cognitive models of spatial relations, such as the linguistic studies evocated above.

2.3 A scale-dependent notion

Of course, any model of the world depends on its intended use. Thus, the relevant spatial context of an object depends on the purpose of the considered geographic data. This purpose-dependency is by nature an important point for the generalisation process that deals with two different purposes: the purpose of the initial data and the purpose of the final data. The intended scale of analysis of the data is a key aspect of the purpose of the data. The notion of scale is important in cartography because it determines the display capabilities (e.g. minimal distance between objects on the paper map). Even more, the notion of scale is important because it is related to the level of analysis of geographic data (is the analysis done at the level of the town or at the country?). Thus, the influence of the scale on spatial context must be studied.

First, the notion of scale influences which objects are considered relevant in the spatial context of an object. For example, the trails crossing a road are interesting for hikers which have a local view of the geographic space (typically, they may use 1:50,000 scale maps). But these trails are out of interest for drivers who consider space at a larger level of analysis and who may be more interested by the towns accessible by the road (that may be found on 1:500,000 scale maps for example).

Second, the scale influences the definition of the high level objects (the groups and the areas) to be considered to represent spatial context well. For example, the coastal area may stop at a few hundred meters from the coast in a local analysis of marine plants, but it may stop a few miles from the coast in a global analysis of whales' displacement. As another example, let us consider a house which is located in an isolated village, the latter being located in a sparse country. At a large scale the house can be considered as being located in a dense built-up area (the village) comparatively to the houses located in the village's surroundings; but at a lower scale it may be considered as one of the buildings of a sparse built-up area (the countryside).

Finally, the semantics of the relation between different objects may change according to the scale of analysis. For example, a river and a road may be quite parallel from a global point of view but, from a more local point of view, may be not parallel at all because of the sinuous shape of the river. As another example, if a clearing is inside a forest, the meaning of the *inside* relation may vary according to the level of analysis: at a very local scale the clearing can be considered as a frontier of the forest (where the clearing is, there is no tree), and at a larger scale the clearing is considered to be part of the forest (the forest is made of the wooden areas and the clearings).

3. REPRESENTING SPATIAL CONTEXT AND CONTEXT RELATED RULES

3.1 Tool-box orientation vs. explicit representation of spatial context

Spatial context must be taken into account during the generalisation process. Two main approaches can be used to do that.

First, context can be seen as part of the process: it is analysed inside generalisation algorithms. This is the "tool-box metaphor" used in most of the works related to spatial analysis [Egenhofer and Kuhn 99]. In this case, spatial context of objects is not explicitly represented in data models, but it influences the processes used to generalise data. For example, let us consider an algorithm that takes a mechanical paradigm and considers geographic objects as producing forces pushing away other objects with a strength depending of the distance between them: objects will be moved one from the others without explicitly representing any proximity relationship.

Second, context can be seen as part of the data. In this case, the context is explicitly represented in the data model (as the objects are). In this approach, knowledge describing spatial context and knowledge describing how to use it are separated. It must be noticed that representing relations does not necessarily mean that all of them must be pre-computed and stored, it can be better to compute them on-line when needed [Theobald 2001].

Due to the complexity of the geographic world, it may be sometimes a hard work to explicitly represent all the relevant spatial relations useful for a certain purpose. In these cases, it may be better to consider certain parts of the context inside the algorithms transforming the data. But, representing context as an explicit part of the data may be better from a cognitive point of view if we agree that "spatial relations among spatial entities are as important as the entities themselves" [Papadias and Theodoritis 97]. This allows to develop knowledge-based approaches for generalisation, where the knowledge describing how to manipulate the data is separated from the knowledge describing the data (which can thus be easily reused for other applications). The next section explores how the spatial context and context-related rules could be represented in knowledge-based approaches.

3.2 Representing spatial context

In the second section we raised several points about what is spatial context. These considerations have some consequences on how spatial context could be modelled. Proposing a complete model of spatial context representation goes beyond our current work, but we mention in this section some requirements for spatial context modelling.

Representing high-level objects
The generalisation process requires the manipulation of high-level objects such as groups or areas (cf. section 2.1). These objects have relations with basic objects in the database, and are in relation to each other (two archipelagos can be close from each other, an archipelago can be part of a country). If one wants to consider these objects during the generalisation process, they should be explicitly represented [Lagrange and Ruas 94]. There can be several levels of high-level objects (groups, groups of groups, etc.) and these high-level objects dot not necessarily represent a hierarchical partition of the space (there can exist two overlapping groups of aligned houses, rural and urban areas have fuzzy boundaries that can overlap) [Ruas 99; Ruas 00].

Representing fuzzy concepts
One must be able to represent some relations between certain objects and the area they belong to (cf. section 2.1.2). The relationship between a geographic entity (e.g. a house) and an area it belongs to (e.g. a rural area) can be modelled in different ways in an object-oriented paradigm. It can be modelled as a relation between an object (the house) and another object which has a *fuzzy boundary* (the rural area). It can also be modelled as a *fuzzy relation* between an object (the house) and another object which has no explicit localisation in space (the rural area). Then, it can be modelled as an object (the house) with a *fuzzy attribute* (the attribute "is in" with the value "rural area" possibly with a fuzzy characterisation such as "certainly"). Whatever the choice that is made, an adequate representation of the "element/part of space" relation should model some fuzzy concepts: either objects with fuzzy boundaries, fuzzy relations between objects, or fuzzy attributes [Winter 2000].

Dealing with multi-scale representations
Spatial context is a scale dependent notion (cf. section 2.3). As the generalisation process considers at least two levels of analysis (the final and initial scale), it is necessary to be able to have a multi-scale representation of spatial context. For example, if the relation *is parallel to* is characterised by different properties (e.g. degree of parallelism, place where the two objects are parallel), there should be different instances of this relation at different scales.

Characterising the relations
Spatial relations between objects at the same level of analysis have to be represented (cf. section 2.2). Here are some requirements to represent these relations well: we should be able to characterise them by some quantitative or qualitative attributes (the house is *almost* aligned with the road); we should be able to spatialise them (the road is parallel to the river *in this area*) like it is proposed in the MADS [Parent et al. 98] or Perceptory [Brodeur et al. 99] models; we should be able to represent more than topologic relations.

3.3 Representing context-related rules in knowledge-based approaches

Spatial context is taken into consideration in numerous generalisation algorithms. For example, an algorithm dedicated to the selection of roads in a road network may be guided by the analysis of the other spatial objects accessed through this network (e.g. the towns or the houses depending on the level of analysis) [Richardson et Thomson 96 ; Morrisset et Ruas 97]. These algorithms usually implement some cartographic operations while respecting a set of cartographic rules (e.g. keep highways, keep roads necessary to access to main towns, etc.).

Context-related rules can also be explicitly represented in knowledge-based approaches used to combine numerous generalisation operations. Several knowledge-based models do exist (e.g. expert systems, case base reasoning). Identically, cartographers can express the influence of context on map generalisation using different types of cartographic rules, and each of these rules can be more directly represented using a particular model.

Context can be used as a mean to express exceptions (e.g. "remove all houses except the isolated one"). This use of context is particularly adapted to a rule-based representation of cartographic knowledge, where exceptions to the rules contain context-related terms (e.g. [McMaster and Shea 92]).

Context can also be used to classify the different actions to be performed (e.g. "if the island is in the middle of sea, transform it using *operation1*; if it is on a river, transform it using *operation2*; if it is on a lake…"). This use of context is adapted to a rule-based or case-based (e.g. [Keller 94]) representation of cartographic knowledge, where each rule or each case describes a certain context.

Context can also be used to express constraints on the expected result (e.g. "the possible parallelism relation between the road and the river must be kept" or "the river must remain in the valley"). These cartographic rules are more easily expressed in a constraint-based formalism (e.g. [Beard 91; Ruas 00]).

4. CONCLUSION

During the cartographic generalisation process, geographic objects cannot just be considered one by one. The way objects are processed clearly depends on their spatial context. In this paper, we studied the nature of spatial context encountered during map generalisation. We differentiated three types of relations that an object can have with its environment: being part of a significant group, being in a particular area, being in relation with 'same-level' surrounding objects. We also addressed the issue of scale-dependency of spatial context. This allowed us to draw certain requirements necessary to represent the spatial context of objects.

This work is a step towards a better definition of what is spatial context. The interest of better understanding the spatial context of geographic objects is not limited to cartographic generalisation. Most geographic analyses do consider the context of objects (why did these phenomena appeared in certain context? what would happen to this phenomenon if the context is changed? etc.). Objects of the geographic world are highly interrelated. It is thus impossible to explicitly represent in data models all the significant relationships between geographic objects. Anyway, we believe that a better modelling of spatial context could pave the way to the creation of richer and more effective geographic data models.

ACKNOWLEDGEMENT

Thanks to the network of centres of excellence, GEOIDE, to National Defence and to Ministère des Ressources Naturelles du Québec for supporting this project.

REFERENCES

AGENT 00. European research and development project December 1997- December 2000; ESPRIT 24939. Cf. http://agent.ign.fr

Beard K. 1991. Constraints on rule formation. In *Map Generalization : Making Rules for Knowledge Representation*, Buttenfield et McMaster (eds), Longman Scientific & Technical, Harlow, Essex, pp.32-58.

Brassel K. and Weibel R. 1988. A review and conceptual framework of automated map generalization. *International Journal of Geographical Information Systems*. Vol. 2, no. 3, pp.229-244.

Brodeur, J., Bédard, Y. and Proulx, M.-J.., 2000, Modelling Geospatial Application Database using UML-based Repositories Aligned with International Standards in Geomatics, In *Proc. of ACMGIS 2000*, November 10-11, Washington DC, USA

Egenhofer M.J., Frank A. and Jackson J.P. 1989. A Topological Data Model for Spatial Databases. In *1st Int. Symp. on the Design and Implementation of Large Spatial Databases*, LNCS 409, pp.271-286. Springer-Verlag.

Egenhofer M.J. and Mark D.M., 1995. Naive geography. Spatial Information Theory. *Proc. Of COSIT'95, Lecture Notes in Computer Science 98,* Springer Verlag, Berlin, p.1-15.

Egenhofer M.J. and Kuhn W. 1999. Interacting with GIS. In *Geographical Information Systems: Principles and Applications*, *2nd edition*, Maguire D.J., Goodchild M.F., Rhind D.W. and Longley P. (eds), Wiley, New York, pp. 401-412.

Keller S. 1994. On the use of case-based reasoning in generalization. *Proc. of 6th International Symposium on Spatial Data Handling*, Edimbourg, Vol. 2, pp. 1118-1132.

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

Lagrange J.-P. and Ruas A. 1994. Geographic information modelling: GIS and generalisation. In *Proc. of Int. Symposium on Spatial Data Handling*, Edinburgh, pp1099-1117.

Mark, D. M., 1999. Spatial Representation: A Cognitive View. In *Geographical Information Systems: Principles and Applications*, *2nd edition*, Maguire D.J., Goodchild M.F., Rhind D.W. and Longley P. (eds), v. 1, pp. 81-89.

Mathet Y., 2000. New paradigms in space and motion: A model and an experiment. In *Proc. of ECAI Workshop on spatio-temporal reasoning*, Berlin, 2000.

McMaster R.B. and Shea K.S. 1992. *Generalization in Digital Cartography*. Assoc. of American Geographers, Washington.

Morrisset B. and Ruas A. 1997. Simulation and Agent Modeling for Road Selection on Generalisation. In *Proc. of Int. Cartographic Conference*, Vol.3, pp.1376-1380.

Mustière S., Zucker J.-D. and Saitta L. 2000. An Abstraction-Based Machine Learning Approach to Cartographic Generalisation. *Proc. of 9th International Symposium on Spatial Data Handling*, Pékin, sec.1a, pp.50-63.

Papadias D. and Theodoridis Y. 1997. Spatial relations, minimum bounding rectangles, and spatial data structures. *International Journal of Geographical Information Science*, vol.11, n.2, pp.111-138

Parent C., Spaccapietra S., Zimanyi E., Donini P., Plazanet C., and Vangenot C.. Modeling Spatial Data in the MADS Conceptual Model. In *Proc. of the 8th Int. Symp. on Spatial Data Handling,* Vancouver, Canada, July 1998.

Richardson D.E. and Thomson R.C. 1996. Integrating Thematic, Geometric and Topologic Information in the Generalization of Road Networks. *Cartographica*, Vol.33, n.1, pp.75-83.

Ruas A. 1999. *Modèle de généralisation de données géographiques à base de contraintes et d'autonomie*. Ph.D. Thesis, university of Marne-la-Vallée, France (in French).

Ruas A. 2000. The role of meso objects for generalisation. *Proc. of the International Symposium on Spatial Data Handling*, Beijing, China, sec.3b.

Sablayrolles P. 1995. *Sémantique formelle de l'expression du mouvement. De la sémantique lexicale au calcul de la structure du discours en français.* PhDhesis, Univeristy Paul Sabatier, Toulouse (in French).

Scholl M., Voisard A., Peloux J.-P., Raynal L., Rigaux R. 1996. *SGBD Géographiques - Spécificités*. International Thomson Publishing, France (in French).

Simons P. 1987. *Parts, A Study in Ontology*, chapter 1-3, pages 5-128. Oxford, Clarendon Press.

Theobald D.M. 2001. Topology revisited: representing spatial relations. *International Journal of Geographical Information Science*, 2001, vol.15, no.8, pp.689-705

Winston P.H., 1984. The Key Role of Representation. In *Artificial Intelligence, second edition*.

Winter S., 2000. Uncertain topological relations between imprecise regions. *International Journal of Geographical Information Science*, vol.14, n.5, pp.411-430

Worboys M.F., 1996. Metrics and topologies for geographic space. In *Advances in GIS research II : proc. of 7th International Symposium on Spatial Data Handling,* Kraak and Molenaar (eds), Taylor and Francis, p.365-375.