UC Santa Barbara
NCGIA Technical Reports

Title
Report of the ICA Workshop on Map Generalization (97-3)

Permalink
https://escholarship.org/uc/item/6hh6380h

Authors
Mackaness, William A.
Weibel, Robert
Buttenfield, Barbara P.

Publication Date
1997-10-01

NCGIA

National Center for
Geographic Information and Analysis

Report of the ICA Workshop
on Map Generalization

Gävle, Sweden

19-21 June 1997

by

W.A. Mackaness, University of Edinburgh
R. Weibel, University of Zurich
and
B.P. Buttenfield, University of Colorado-Boulder

Technical Report 97-3

October 1997

Simonett Center for Spatial Analysis
University of California
35 10 Phelps Hall
Santa Barbara, CA 93106-4060
Office (805) 893-8224
Fax   (805) 893-8617
ncgia@ncgia.ucsb.edu

State University of New York
301 Wilkeson Quad, Box 610023
Buffalo NY 14261-0001
Office (716) 645-2545
Fax   (716) 645-5957
ncgia@ubvms.cc.buffalo.edu

University of Maine
348 Boardman Hall
Orono ME 04469-5711
Office   (207) 581-2149
Fax   (207) 581-2206
ncgia@spatial.maine.edu

Report of the ICA Workshop on Map Generalization
Gävle, Sweden
19-21 June 1997

Mackaness, W A, U. Edinburgh (Scotland)
Weibel R, U. Zurich (Switzerland)
Buttenfield B P, U. Colorado-Boulder (USA)

1 Introduction

This paper reports the discussions and key findings of a three day workshop held in Gävle, Sweden immediately prior to the 18th ICA/ACI International Cartographic Conference in Stockholm. The workshop mirrored the considerable success of the first workshop help in Barcelona in 1995 and comprised demonstrations, presentations, and parallel sessions. Thirty nine delegates came from seventeen countries comprising 20% from National Mapping Agencies (NMAs), 10% from the private sector and 70% from research institutions (Geography, Cartography, Surveying, Computer Science, Physics, Planning and Linguistics). The objective of the workshop was to hold a series of discussions focused on impediments to automated map generalization, the current state of knowledge, and progress on specific problem areas. Whenever possible, discussion pressed to define specific areas for research in the coming two years. These will be noted throughout the report.

The aims of the workshop were:
- to present results and assess the progress made over the past two years;
- to share and critique ideas, techniques and methodologies;
- to identify impediments to further progress; and
- to foster collaboration among researchers in the field.

Topics for discussion were distilled from submission of abstracts in response to an open call for white papers. The intense interest in participation required parallel sessions over two of the three days, with papers presented in the following sessions:
- Map Production Systems: Critique of the State of the Art
- Knowledge Acquisition and Encoding
- Modelling Semantics and Non-Spatial Structures
- Modelling Geometric and Spatial Structure
- Quality Assessment and Constraints
- Conflict Detection and Resolution
- Synthesis: Integration, Strategies and Control
- Conclusion: Priorities for Research

Each session consisted of three to five brief presentations followed by 'break-out' discussions. Summary reports of each group then provided a focus for collective discussion. In addition, live demonstrations illustrated generalization algorithms and experimental platforms and international collaborations currently under development. The precise structure of the workshop,

attendance lists, abstracts and other information on the research of the working group can be found at http://www.geo.unizh.ch/ICA/. This report is structured around the plenary session headings, includes a summary of one session held jointly with the ICA Visualization group meeting concurrently in Gävle, and concludes by listing the intended themes for discussion at the next workshop in 1999. First, however, a summary is given of the last workshop held in Barcelona in 1995, in order to place the Gävle meeting into context.

2 Setting a context: Summary of the Workshop in Barcelona 1995

The First Workshop on Progress in Automated Map Generalization was held on 1-9 September 1995 prior to the 17th International Cartographic Conference in Barcelona, Spain. Thirty researchers and practitioners (spanning eleven countries) with proven track records in generalization theory and practice presented their research. This was the first workshop held under the auspices of the ICA Working Group on Map Generalization, but already the fourth in a series of international meetings on generalization. The goal of the Barcelona meeting was to advance the state of knowledge on cartographic automation, to report demonstrated progress in areas of map generalization and design, and to introduce junior researchers from Europe and North America thus fostering an international collaborative environment.

The meeting was co-sponsored by the International Cartographic Association working group on Map Generalization, the National Center for Geographic Information and Analysis (NCGIA), the European Science Foundation GISDATA Programme, the Institut Géographique National (IGN, France), Institut Cartogràfic de Catalunya (ICC, Spain), Organisation Européenne d'Etudes Photogrammétriques Experimentales (OEEPE), and the Canada Centre for Remote Sensing (CCRS). A full day was devoted to each of three topics, alternating small group discussion with plenary sessions. The three topics were
> 1) quality assessment of design alternatives;
> 2) formalizing knowledge in the generalization process; and
> 3) feature conflict detection and resolution.

2.1 Quality in Map Generalization

During generalization, changes in detail may alter the map in terms of what relationships are shown, and how they are displayed. Progress has been made in developing a number of mechanisms for measuring these changes, and for measuring pattern overall. While it is now relatively simple to describe the quality of individual (primitive) cartographic objects, it is still challenging to assess the quality of complex or compound objects found at higher levels of abstraction, and to track polymorphism when features merge and re-emerge across map scales. Some researchers are adopting an operational research approach, with some success: to the extent that map design can be defined as a set of objectives, with the map generalization process being viewed as a compromise between those objectives. Knowing how map task influences this compromise remains an area in need of further study. Whatever approaches are adopted, the costs associated both with quality assessment and quality maintenance in the underlying database must be acknowledged.

2.2 Formalizing Knowledge in Map Generalization

Knowledge-based techniques can be used to automate selection of algorithms and modification of tolerance values. Progress continues to stumble over definitions, over the operational difference between information and knowledge, and over the degree to which implicit map information must be made explicit prior to decision-making. Several approaches for the derivation and use of knowledge were presented at the semantic, metric, and topological levels. At the time of the Barcelona meeting (September 1995), research priorities were to identify the conditional information that triggers actions in the decision making process (for example being able to detect clusters of objects (say buildings) in order to justify their aggregation). Machine learning techniques were under study by some researchers. One approach extracts both object descriptions and object relations from given examples, to acquire both spatial and semantic

relations that exist between objects. In other research groups, transaction logging of operator actions during interactive map design was proposed to recover heuristics (rules of thumb) for dealing with types of conflict and for establishing suitable default values for control parameters. A recurrent theme during the workshop was that knowledge-based systems would require either databases that are rich in semantic and geometric information, or cartometric analysis tools that can make such information explicit.

2.3 Conflict Detection and Resolution

Conflicts in map production arise from a combination of design factors such as map symbology (color or symbol size), output resolution, and proximity of objects. Conflict detection is central to automating the map design process; without detection it is not possible to diagnose the design problem, which in turn is a prerequisite to any set of possible solutions. The most progress has been made in this area, and of the three topics, this one generated the most heated discussion. Both vector- and raster-based approaches were suggested that afforded means of detecting and resolving conflict. It remains difficult to classify conflict types which may occur variously at both semantic and geometric levels and this taxonomy was cited as a high research priority. Effective strategies to determine the order in which conflicts are resolved are also in need of research, as are experiments for integrating a mix of generalization techniques to resolve various types of conflicts. It is also apparent that prior operations will affect objects geometrically and semantically; thus sequences of generalization may become meaningless. Backtracking requires a chronology of single and complex object transformations, however metadata description mechanisms are much too rudimentary at present to handle this complexity. Herein lay another high priority for research, to develop efficient mechanisms by which to track the generalization chronology and document its parameters and integration sequence.

2.4 Summary of Barcelona Workshop

Key research tasks identified at the close of the Barcelona meeting included:
- develop and integrate criteria, measures, and methods for quality assessment;
- identify 'missing' algorithms for map generalization processes;
- investigate the sequence and synergy between generalization operators;
- determine the domain of applicability of generalization algorithms with respect to scale, map purpose, etc.;
- develop hybrid data structures integrating semantic and geometric operations on simple/compound objects;
- develop strategies for 'database enrichment': methods for encoding primary (explicit) knowledge into the database as well as algorithms to derive secondary (implicit) knowledge from stored data;
- study knowledge utilization in the design process;
- perform user requirements and cost benefit analysis on automated map generalization; and
- establish a common shared database on which to carry out empirical studies.

3 Map Production Systems: Critique of the state of the art

3.1 Presentations

There is an increasing demand that generalization tools should be integrated into commercial GIS to enable the derivation of multi-purpose database and map products from a detailed master database. [Lee] presented latest developments at ESRI to integrate generalization functionality

into ARC/INFO, developments which are in many ways typical of the situation many vendors of GIS or cartography systems are facing today. The approach taken by ESRI focuses on the provision of a suite of tools which implement generalization operators and can be combined into customized workflows by the user.

[Pla] discussed the use of commercial systems (Intergraph's Map Generalizer and CHANGE developed at the University of Hannover) in map production at the Institut Cartogràfic de Catalunya (ICC). The implementation and evaluation of algorithms in the production workflows has highlighted weaknesses and strengths of current algorithms, as well as the evaluation of time saving resulting from the use of map generalization systems. The ICC has made a 20% time saving as a result of its adoption of digital generalization techniques. Discussion of the efficiencies and time saved through automated techniques highlighted the need for effective methods for measuring such benefits.

[Bakker] presented the work of the Dutch National Mapping Agency (Topografische Dienst) who should complete four of five defined vector databases in 1997. Current work is focusing on 1:100,000 scale source data. The testing of Intergraph's MGMG software revealed a 20% time saving over manual practice through interactive generalization.

3.2 Discussion
Definite progress in automated generalization has been made since the last workshop in Barcelona in 1995. Developments in generalization, however, continue to be constrained by the poverty of underlying database models. Required extensions include mechanisms for modelling topological generalization and recording of non-spatial relationships between objects. The NMA representatives present at the meeting asserted that implementations of algorithms and testing using real datasets could be extremely useful. Note that the OEEPE Working Group on Generalization is presently conducting a series of such software tests (for further information, contact Anne.Ruas@ign.fr).

What the group felt to be missing from current generalization software includes:
• richer sets of explicit topological information for implementing new generalization algorithms,
• mechanisms for automatic conflict detection and resolution,
• criteria for assessing and controlling the workflow (sequence), and
• mechanisms for choosing between alternate solutions.

The following ideas for research and development were proposed:
• development of comprehensive data structures that model the interdependence of geographical entities and their importance in governing the map generalization process,
• research into rules and constraints governing operators according to scale range and purpose.

4 Knowledge Acquisition and Encoding

4.1 Presentations
[Edwards] explored the parallels between cartographic generalization and natural language descriptions. "Can the study of how humans conceptualize and abstract space help us to understand how best to represent space?" For example there are interesting parallels between cartographic

generalization and route description -- a topic covered by [Gryl] in her presentation -- in the selection of scope and information content, on form, on scale and on the use of symbols. Both tasks are driven by context and intended use, and both have definable, albeit different, constraints. Generalization is constrained by graphical and visualization requirements. Route descriptions are constrained by linguistic and memorability requirement. The two tasks have in common the need to move from more global to more local scale in terms of context and understanding of the salient message.

[Buttenfield] explored ways of 'de-centralizing' the generalization process in the context of client/server relationships, in such a way as to minimize the need to download large volumes of data in support of interactive systems. The presentation gave emphasis to the idea that while research issues initially arose in the context of map production, it is increasingly the case that map generalization lies at the core of problems in data integration, exploration and contextualization. [Buttenfield] gave examples of encapsulating data and generalization operators in object-based and agent-based structures which are sent to a server to process large datasets remotely. The return subsets called "data kernels" of the original archives are encapsulated with the parameters used in the generalization tasks which additionally captures a chronology of the applied processes. Demarcating a chronology of processing steps was cited as a priority for research at the Barcelona meeting.

In his presentation [Neumann] pointed out that significant 're-engineering' of the way we have approached digital generalization in previous years is now necessary, in order to overcome what he calls the 'cognitive function crisis' of map generalization. He observed that more often than not, map products derived from master databases using automated techniques do not represent the cognitive value of the 'nominal ground', often making them incomprehensible, obstructing the message the map wants to convey. As a result, he called for research that formalizes the cognitive knowledge that goes into the production of maps and develops methods for extracting the deep structure of maps.

This was followed by [Tian] who discussed the generalization of nautical charts and the constraints of design that had to be closely adhered to in the interest of ship safety. Specifically, he pointed out that generalization of coastlines on hydrographic charts must be constrained to err on the deeper side of linear features, to preotect as much as possible against ships running aground. [Tian] used the example of a modified Douglas algorithm to show how such constraints can be incorporated. The work illustrated the need to integrate the spatial and semantic components of generalization, as well as the purpose.

4.2 Discussion
A range of techniques have variously been applied in order to extract and encode the knowledge used in the map design process. The discussions in this session explored potential contributions from the field of natural language processing, philosophical issues such as an holistic approach to design and ways of embedding design constraints in the behaviour of generalization algorithms. Embedding knowledge in GIS requires that we examine the efficiencies in workloads, and balance decision making between the user and the machine. We must also explore design efficiencies in the transfer of knowledge from the server to the client. The success of this approach depends on 1) database enrichment, 2) creation of hierarchical structures to support map generalization, and 3) minimizing interaction workload.

The justification for utilizing a rule based approach has long been known. Rules can be established from a broad range of sources and can be used to modify the application of geometric operators as well as to integrate tools. For example experimental work in aggregation techniques has been implemented in ARC/INFO as part of a Hong Kong mapping project. Similar work has been accomplished for other commercial systems.

Computer based systems offer a different paradigm for map creation. The necessity for mimicking human solutions in automated map design is debatable, though some would argue that it is important to do so in order to develop an intelligent generalization system.

To the extent that map generalization systems are applicable to a broad range of purposes, it is important to emphasize flexibility in design. It is apparent that differing degrees of freedom exist according to map or database purposes (Figure 1).

Nautical charts   Few degrees of freedom
Aeronautical charts   Standardised products
                             Concrete products
                             Traditional design conventions

Topographic maps

Thematic maps

**Cartographic
models in GIS**   Many degrees of freedom
                             Abstract models
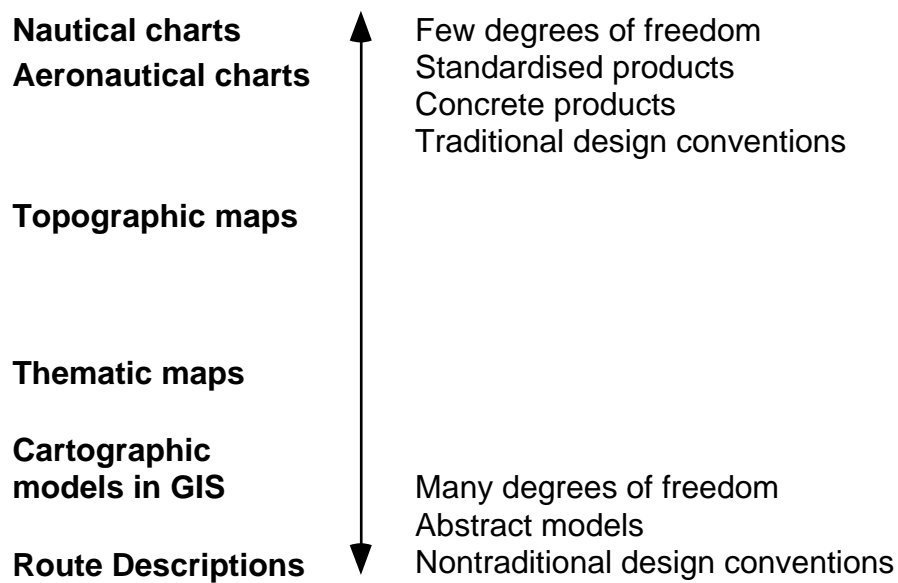**Route Descriptions**   Nontraditional design conventions

Figure 1
Degrees of freedom in map generalization vary according to purpose.

Given this continuum, one may ask what kind of knowledge is needed for what kind of product? Is it possible to extend knowledge gained about more standardized products to less standardized abstract products? A range of sources have been proposed (textbook, NMA guidelines, etc.) as well as using reverse engineering through the analysis of map series and directly from map experts. Within the context of actual implementations, it is possible to track the process through interaction logging. Various authors have explored methods for machine learning (using genetic algorithms and neural nets) but the issue still remains as to what type of knowledge (geometric, semantic) is used, where and how it constrains the degrees of freedom along the continuum of map products.

5 Modelling semantics and non-spatial structure

5.1 Presentation
The session began with a presentation by [van Smaalen] on mechanisms for combining attributes through aggregation. The presentation explored issues of meaning in aggregation and again emphasised the importance of semantics. [Bjørke] considered entropy-based modelling for map generalization and illustrated its use in homogenous point removal by adding constraints to a simple entropy model. In effect these constraints model the semantic requirements of map design.

[Harvey] proposed a framework for considering generalization in its broadest sense, basing his ideas on the work of Alfred Hettner. [Harvey] examined the weaknesses of the reductionist approach to map generalization and emphasised the need for a semantic framework for geometric matching of geographic information. The paper highlighted the complexities of information integration and explored the changing role of generalization as 'semantic emphasis through design', with the consequent need to understand the purpose and intentions of the user.

The surveying and mapping agencies of Germany are developing ATKIS, a system that records information at three DLM (digital landscape model) levels of varying resolution, accuracy and content [Schürer]. To reduce the significant manual effort required to maintain these datasets, a system for the automatic derivation from finer to coarser resolution DLMs is being developed. The generalization of map semantics includes the steps of classification, selection and modification of attributes. The generalization of geometry includes changes in geometry type, and adjustment of resolution. The parameters of the algorithms are set according to the minimum legibility constraints and the number of points required to describe the geometry at lower levels of resolution in the DLM.

5.2 Discussion
Though it has long been known that 'purpose' defines the core constraints of design, the link between 'intended emphasis' and a collection of operators applied to a variety of objects continues to prove elusive. Semantics are crucial for qualifying constraints in map design. The presentation of map semantics comes from the transformation and arrangement of a set of primitives. Map purpose underpins a thematic approach to design, and assumes increasing importance as we steer away from Fordist views in topographic map production towards a broader range of cartographic data products.

To formalize generalization effectively, we must acknowledge the intrinsic link between the geometric and the semantic. We can derive some of the meaning through the analysis of an object's spatial characteristics. Yet the problem of incorporating meaning in the map design process remains occluded by a focus on geometric solutions. Meaning requires context. Balance in representation between what is merely contextual and what is salient is not easy to model. An element of redundancy is required to reinforce the context while at the same time giving emphasis to what is most salient.

The message of the map is manifest in symbols in Cartesian space. Symbol geometry is subservient to the semantics of geographic pattern, flow, relative density, connectivity, and neighbourhood relationships. In past attempts to automate generalizaton, we have largely solved geometric problems in the absence of semantic preservation, and in this regard the generalization

community has missed its mandate. The need to acknowledge semantics in research and in production surfaced over and over again during the Gavle discussions (the same had already happened at the Barcelona meeting). Map purpose governs the types of semantics to retain or emphasize. From purpose it should be possible to identify for each object a set of relationships that are deemed worthy of preserving during the generalization process. In the absence of techniques for modelling such semantics, the generalization process will remain ill-defined.

Discussion revolved around modelling constraints through semantic nets. It was argued that semantics make generalization specific and can probably be used as a basis for constraining the map generalization process. It is important that the semantic net models the relationships conveyed in the map space, and that the semantic net is used to translate in an explicit manner the relationships between object classes. The idea of preserving defining characteristics of geographic space leads intuitively to the concept of pattern recognition. Indeed we can view the interpretation of semantic nets as but one special case of pattern recognition. But patterns manifest differently at different scales. Thus the question becomes 'which patterns to measure' and 'which patterns to preserve', and at which scales are operations appropriate. It was also proposed that a possible alternative to semantic nets is the use of case based reasoning and the potential of prototypical modelling techniques.

As researchers explore interdependencies between operators and objects, it remains to be seen how necessary it will be to make explicit a holistic view of the map design process. It is intriguing to observe how the language in our discussions was precisely that of the discipline of geography - modelling and visualizing essential/pertinent manifestations of various patterns across scales.

The discussion of semantics expanded to link context with mechanisms for constraining the solution space, and to the phenomenological view of map generalization. This discussion led to the idea that the emphasis should be on 'semantic generalization': the transformation of abstraction of space which preserves specifically defined relationships. Relationships that may be well-formalized or semi-formalized. These discussions on meaning and purpose provided valuable input to later discussions on strategy and orchestration of generalization operators.

6 Modelling geometric spatial structure

6.1 Presentations
[Brown] presented a discussion on raster and vector approaches in map generalization highlighting the advantages and weaknesses of each. He explored the appropriateness of each with respect to the map use cube proposed by MacEachren, suggesting that grid cell approaches may offer advantages over vector in the data exploration and analysis stage. This provided a comprehensive backdrop to work by [Zhilin Li] concerning algorithms for elimination of objects and displacement using a raster based model (work that was carried out jointly with [Bo Su]). He presented ideas on some of the missing algorithms in generalization exploiting operators of mathematical morphology in the context of a raster based model. [Edwards] presented research on behalf of [Yang] who was unable to attend. [Yang's] work concerns dynamic object generalization underpinned by a Voronoi model that allows explicit statements to be made relating to an object's local neighbourhood. The hierarchical structure implicit in maps of varying resolution underpins the idea of a hierarchical coordinate system (Quaternary Triangular

Mesh - QTM) proposed by [Dutton]. Feature locations are replaced with QTM coordinates that encode the required degree of accuracy, and the efficient retrieval of information at small scale (coarser resolution). The tessellation of space is based on an octahedron representation of the earth, with recursively finer meshes linked through a numbering scheme that enables the tessellation to be traversed in extent and resolution. The model has the potential to support 'structuration' of map features, and to assess levels of content by region. [Richardson] presented work done at the Canadian Centre for Remote Sensing (CCRS) aiming at the ingestion of provincial or other large scale data into databases at the federal scale. The methods presented support automatic feature checking and structuring for topological quality, and provide a mechanism to automatically classify, code and build an enriched and object oriented database using only topologic and metric relationships. The paper highlighted questions such as 'can we define aggregation hierarchies for generalization?' and 'how can feature classifications and codings be automatically derived from a topologically structured database prior to generalization?'.

6.2 Discussion
It is interesting to observe that all but one of the presentations were based on exhaustive tessellations of space (tiling methods), thus highlighting the need for detailed knowledge regarding context and neighbourhood information. This was a recurring theme among the demonstrations and discussions (and something that was proposed as a concept at the Barcelona workshop). Although algorithms are being developed that consider the context of their operation, it was noted that at the precise point of generalization of any one object, the topological and proximity information is effectively lost in that the object alone is considered. This makes contextual generalization difficult to model. Local neighbourhood information is not well incorporated into algorithms; neither is uncertainty and error resulting in the need for post application evaluation. Obviously, tessellated data models are advantageous in representing proximity and topological relations (particularly between disjoint features). All commonly used tessellations were represented at the workshop. The raster model was praised by its proponents for its simplicity and the fact that it does relate to the predominant output format (CRT, raster plotters), the Voronoi model is capable of comprehensively modelling all relevant topological relations of points, lines, and polygons as well as proximity and shape properties, and triangulated models have often been used as an auxiliary data structure for contextual operations.

The debate continues over the suitability of raster and vector techniques in both detection and resolution. It remains the view of most researchers that benefits can be derived from both approaches but that the conversion between raster and vector models may not be without problems. Others argue that the debate is sterile, advocating instead the use of Voronoi modelling techniques (which represents a hybrid model) and the use of object oriented approaches. It was argued that the appropriateness of various models (raster, vector, Voronoi, QTM) varies with scale, purpose, format of source data and whether there is a need to model fuzziness. Though various hybrid models have been proposed, it is not clear if this is due to the poverty of techniques developed in vector data structures or whether there are enough operations unique to the raster environment to warrant the cost of conversion from one to another.

The debate over these geometric modelling requirements led to discussions of database enrichment in support of complex model generalization. Such a need is apparent from the difficulties encountered in categorical mapping and the creation of composite features that

minimize internal variation. It is evident that generalization requires us to examine a range of lexical and geometric information in support of such processes.

7 Quality - assessment and constraints

7.1 Presentations
The first presentation examined mechanisms for assessing the quality of terrain generalization and argued the case for a range of morphometric measures in support of this [Wilke]. Current evaluation tends to be visually based. [Wilke] proposed objective measures based on statistical, geometric and morphometric criteria emphasising the needs and benefits of evaluation methods that are reproducible. The second paper discussed the use of fractal techniques for the analysis and enhancement of both linear and DEM information [Nakos]. The fractal generalization methodology was based on three steps - verification of fractal character, determination of fractal dimension, and application of a fractal simplification algorithm. The third paper [Jaakkola] presented work on the automatic generalization of categorical maps using raster based techniques, the overall aim being the creation of multi scale datasets and maps from a single detailed dataset (focusing around CORINE data). The paper generated discussion of assessment techniques that went beyond a visual assessment. It was shown that where standards do exist for manually derived maps, humans are not consistent in their approach [Jaakkola].

[McMaster] presented work on the visualization of quality in the application of generalization algorithms (such as line simplification). The images illustrated the complexity of task in displaying meaningful quality indicators as well as understanding better the behaviour of algorithms. The final paper in the session [Lagrange] emphasised the need to better understand the confines of applying various generalization algorithms and their relationship with the overall management of the generalization process. Control assessment and understanding process (e.g. by identifying the metric, topologcial and semantic constraints governing generalization) are critical to the development of self evaluating algorithms.

7.2 Discussion
The group first bounded the discussion by defining what is meant by the term Quality? Several definitions could be used, incorporating not only data quality (accuracy, precision and fitness for use) but also product quality (such as the aesthetic quality of paper maps) and quality of use (whether the data or maps are put to appropriate use, for their intended purpose). In the remaining discussion, which became quite heated at times, the group tackled five questions relating to Quality Assessment.

The first question considered what should be assessed, and what measures and methods for assessment were appropriate. A table which was modified from the previous talk by [Lagrange] was compiled, cross-tabulating assessment of individual objects, 'situations' (groups of objects that make up a meaningful pattern, such as an alignment of buildings along a street) and entire map products in terms of geometry, topology, semantics, and aesthetics. Assessment, according to this view, consists of evaluating to what extent the contraints (which are defined for the generalization process) are satisfied. For example, assessing topological constraints at the individual feature level might involve checking for self-intersection. At the local or situation level assessing geometric constrainst might involve measuring and preserving proximity relations between members of a group of objects.

The second question the group addressed related to specifying what cartographic content should be assessed. In a map production environment, assessment is "expensive" in terms of both time and labour. Assumptions that quality (geometric, topologic, semantic, aesthetic) is uniform across a map surface differ greatly from a thematic or GIS mapping situation, where one must assume just the opposite (that quality varies across a map surface). Proper assessment techniques must focus on constraints associated with the content. For example if we compare a Michelin road map with a topographic map, we find the road map the road map emphasises topological and labeling qualities aat the expense of absolute positional accuracy. For the topographic map, geometric constraints on content take highest priority, in that (for example) rivers must maintain accurate positioning in order to lie correctly on the terrain. This emphasises the need to define constraints with respect and relative to, other layers of geographic data.

The third question the group addressed related to assessment, with emphasis on relative and meaningful measures of quality. [Jaakkola] presented research showing that when assessments are referred against manual generalizations, a lack of consistency often appeared; his group in Finland concluded this is due to inconsistencies in the manual solutions, not necessarily in the generalization being assessed. The group discussed other referents, concluding that assessments against non-generalized data will prioritize the degree of generalization in a test, and that assessments against other automated methods will emphasize such aspects as the number of categories, proportion of data falling within categories, and so forth.

Fourth, the group considered differences in assessment for specific mapping contexts, comparing topographic map production environments with GIS data analysis environments. In these two situations, the need for assessment and methods of assessment were seen as quite distinct. In topographic map production where comparisons are made in reference to a "nominal ground", operators are most often trained cartographers familiar with production specifications and (sometimes) with the area to be mapped. Visual inspection was deemed sufficient for identifying cartographic blunders and errors of logic or omission. It was also pointed out that for many clients who do not fully understand the nature of cartographic data quality, notification that quality assessments are being performed on a product could be viewed as indicating that the product is inaccurate or otherwise suspect. In GIS data analysis where operators are not often trained in cartography, and where data are produced by map overlay and modelling, the introduction of spatial error is expected to vary across a map surface. Some operations will reduce errors, others will compensate, and the effect of other operators on map quality is difficult to predict. Development of tools for assessing and visualizing variations in data quality will help hydrologists, foresters, geologists, and other GIS users to understand the map data they generate, and to recognize map locations where their GIS models are more or less reliable. This fourth issue (assessment in topographic versus GIS mapping contexts) generated perhaps the most heated debate of the entire workshop, as it brought forward two distinct cartographic "cultures" who face similar generalization problems within different working environments. Issues of quality assessment, liability for data, and legal mandates surfaced in this discussion.

The final question addressed by the group revolved around identifying at what stages in the cartographic process assessment should occur. The group felt that assessment is appropriate at any stage where costs can be reduced as a consequence, and this led to a call for a cost-benefit analysis of generalization and assessment systems (another of the key research tasks identified at the Barcelona meeting). Stages seen as benefiting most from cost benefit analysis include system

development, algorithm testing, and integration of sequential generalization operators. The idea that cost factors can take equal priority to validating final map products in quality assessment activities was one of the important outcomes of this session's discussion.

8 Techniques in conflict detection and resolution

8.1 Presentations
The session began with a paper that explored the use of triangulation based techniques and finite element analysis (an example of lateral application of ideas borrowed from structural engineering) [Højholt]. The idea being that deformation of the grid allows the displacement to be propagated across the map space. The model is built on a constrained Delaunay triangulation of vector data and has been applied to buildings and streets. Different deformation properties are associated with different features (such as rigid cased buildings and plastic roads) in order to create solutions in which objects retain their essential defining characteristics. The technique is holistic, simple, homogeneous and avoids conflicts.

[Burghardt] presented work on automated displacement by energy minimization using snakes (active splines). Using line displacement as an example, he was able to illustrate how the conflict can be described in terms of external energy, and the internal energy of the line can be used to maintain shape during line displacement. Shape measures were formulated from derivations of the first and second derivatives of the arc length. This research is again illustrative of the lateral application of some exciting ideas from other disciplines.

[Lecordix] presented Plage, an experimental vector based system developed at the Institut Géographique National (IGN) for conflict detection and resolution. He illustrated how conflicts can be detected in roads and information provided can be used to select appropriate algorithms to resolve the conflicts. New algorithms were developed and implemented for resolving conflicts (such as coalescent bends) by line caricature, including 'ballooning' tight bends, offsetting sets of bends ('accordion'), and merging of consecutive bends ('plâtre'). The implementation has been shown to be very reliable in resolving bend coalescence (i.e. bends of a single line) while a solution for dealing with coalescent lines is being developed based on Nickerson's (1988) algorithm. This solution is not 100% reliable and requires user interaction.

[Ware and Jones] presented their work on conflict resolution strategies based on object displacement and trial positions, borrowing from the extensive work in automated point feature label placement, in particular use of triangulated simplicial data structures developed in a prototype system called MAGE. MAGE operations include amalgamation, displacement, simplification of building outlines, and local search procedures for identifying nearest neighbours. In some situations predefined trial positions might find an adequate solution to conflicts via displacement. The success of the system will depend on efficient search, and the ability to recognize a good solution. The work illustrates the lateral application of ideas developed in text placement. The development of such algorithms are always seen as but one in a range of techniques for solving map generalization problems [Ware and Jones].

[Lifan Fei] illustrated a method for building symbol displacement in the context of buildings and street networks, using both raster and vector techniques. The approach is similar to that of the IGN [Ruas] in that it uses roads to define regions in which to displace objects. Buildings are

classified according to their relationship with the road. The algorithm works by interative displacement of buildings with conflict detection taking place in raster mode in each iteration [Lifan Fei].

8.2 Discussion
It is important to acknowledge the different types of constraints and conflicts that govern the creation of candidate solutions. For example any solution needs to consider the maintenance of:
• semantic (inappropriate change in classification or juxtaposition of classes)
• relative orientation
• relative position
• patterns / gestalt (themselves difficult to measure)
• topological relationships (containment, connectivity, association.)
• a minimum size threshold and separation between objects.

The classification of such requirements leads directly to issues of detection. A range of techniques can be used to detect conflict (vector, raster, triangulation, Voronoi, QTM). This bag of possible techniques raises the issue of hybrid methods as well as the cost of vector to raster conversion. In terms of preserving the inherent qualities of a map, it was suggested that sufficient shape measures already exist to support the creation of hierarchical structures as part of the map generalization process. It is certainly the case that pattern recognition is seen as an important part of conflict resolution both initially and in evaluation. But identifying potential spatial conflicts is not enough in the formulation of a solution. It is additionally necessary to analyse content, explore the topological properties of the conflict as well as the inter-spatial relationships (such as similarity, and disparity between qualities such as orientation, size and proximity).

In terms of conflict resolution, a range of operations can be efficiently modelled using either vector or raster based methods for example with respect to categorical mapping [Jaakkola]. A range of techniques have been implemented in raster format using mathematical morphology operations such as dilation, erosion and displacement [Li], the cell size being an important parameter in governing behaviour and resulting patterns from the generalization process. For all approaches the interdependencies of objects and their effect on generalization are well documented. For example when the process of displacement is applied to urban buildings it is important to avoid new conflicts with road features. Since some of the discussion criticized the Douglas algorithm for self intersection, it was suggested that perhaps we should develop algorithms that are based on conflict avoidance!

The complexity of identifying and verifying the application of generalization algorithms led to a discussion on data enrichment and the poverty of data models in some commercial GIS (such as layer-based GIS). There was also discussion of the variance in data models within and among different national mapping agencies. The question remains whether such modelling can be done on the fly or as part of the pre-processing data enrichment process.

The development and experimentation of techniques for conflict detection and resolution led to a discussion of the shared use of experimental datasets such as is being currently undertaken by the OEEPE (chaired by Anne.Ruas@ign.fr).

9 Synthesis: Integration, strategies and control

9.1 Presentations
Though we have the techniques for altering the geometry of objects (enlargement, displacement, network simplification, etc.) the challenge for research in map generalization has increasingly become 'to what, how much, and when'. In the first paper [Ruas] discussed strategies for design that address the issue of 'orchestration' between the various map generalization operators and examined the cartometric analytical requirements for automated strategies. Such strategies need to be capable of generating multiple solutions and not be deterministic in their approach. These ideas were illustrated by examining mechanisms in urban generalization for 1) removal of less important streets according to semantic and contextual information and 2) improving the clarity of groups of urban buildings through displacement. The work has been implemented on Stratège, an Object Oriented prototype devoted to contextual generalisation. Methods are automatically triggered by means of a rule based language whereby the process is dynamically changed according to spatial analysis.

[Mackaness] discussed the need to develop techniques of automated cartographic generalisation which are generic enough to be widely applicable, yet which are also sensitive to the phenomena of both object type and map task involved. A framework was presented that incorporates geometric properties, semantic properties and spatial relationships in determining appropriate generalisation procedures for objects. This approach was illustrated through an implementation of the phenomenological generalisation of area features. It is only with the integration of geometric properties, semantic properties and spatial relationships that generalisation can be truly sensitive to the phenomena of object types and map tasks [Mackaness].

In the third talk [Plazanet] presented her work on characterization of linear features, illustrating the need to tune generalization algorithms according to the defining characteristics of the line. The characterization algorithms are able to divide the line according to relative convolution, and other shape characteristics, and further able to sub-refine the metric qualities in order to support generalization techniques such as 'accordion' and ballooning (cf. [Lecordix] above).

The fourth paper moved beyond the two dimensions to which generalization is normally applied, and examined issues of integration in three dimensions. Such work highlights the need for logical consistency that goes beyond just geometric precision. For example the model proposed by [Mioc] provides a mechanism of tracking each event and change in the map state as well as a record of the sequence of operators applied in the map generalization process. The important issue being a system that supports exploration of geographic phenomena across scales through both space and time. Extending generalization techniques into the spatio-temporal domain represents an exciting development in the field.

9.2 Discussion
Generalization is still without robust frameworks for the integration and control of generalization operators. The emphasis being on the geometric issues rather than strategies. A phenomenological approach is advocated by several researchers, one that considers 1) the geometry of the object, 2) its attributes and 3) the inter-object relationships.

Though there was general agreement as to the importance of strategy, the question arose as to how far the semantic can be derived from the geometric (metric and topological) qualities of the

map and the degree of pre-processing required to support strategies that can generate a range of candidate solutions.

It was felt that preservation of pattern was a driving goal in the development of strategies in design. This led onto a discussion of types of patterns and methods for ensuring their preservation according to purpose and context. If the map is a series of patterns of similarity, differences, exceptions and associations that become apparent at varying scales, how do we identify, generalize and model such changes in pattern?

10 Live Demonstrations

A video was shown by [Ruas]. The video (accompanied by a great soundtrack) featured a demonstration of PlaGe and Stratège, two prototype systems developed at the COGIT research laboratory of the IGN by various authors. PlaGe focuses on 'independent generalization' (generalization operators which are independent of context such as line simplification and smoothing). Using the example of road generalization (1:50,000 to 1:250,000), a rich variety of algorithms was shown, some implemented from the literature, some developed at COGIT. In particular, three algorithms for line caricature ('ballon', 'accordion' and 'plaster') were shown in operation (cf. also the description of the presentations by [Lecordix] and [Plazanet]).

The second part of the video presented Stratège. The system has focused on research in urban contextual generalization, involving operations such as street and building selection, aggregation and displacement requiring a high level description of spatial and semantic object relationships. Stratège is based on two assumptions: 1) each generalization requires a specific analysis of geographical information, 2) the generalization process has to be as dynamic as possible in order to create alternate designs. Stratège uses rules and task trees to control design. A task tree is a predefined sequence of operations but includes rules that can be modified according to spatial analysis. The system further uses complex spatial data structures such as Delaunay triangulation and minimum spanning trees to evaluate spatial relationships between objects (neighbourhood and proximity).

[Jaakkola] presented a series of large format landuse maps that had been produced from medium scale raster landuse data to generate CORINE landcover data at smaller scale. The main purpose of this demonstration was to show outputs illustrating the points made in [Jaakkola's] paper presentation (see above). The raster-based generalization process involved re-classification to reduce the number of classes as well as geometric operators which were based on image processing operations (e.g. mode filters). The system was implemented using functions of the GRID module in ARC/INFO.

The main purpose of the demonstration by [Burghardt] was also to support his previous paper presentation (see above). A prototype system in C on a UNIX platform implements the energy minimizing approach for line feature displacement using active splines (snakes). The demonstration showed the operation of the system and results obtained using different parameter settings.

[Brazile] gave a presentation of the Working Group home page which contains:
• reports (of business meetings, workshops, etc.), Terms of Reference, announcements, etc.
• mailing list of the working group on map generalization (Address, email)

- a bibliography of recent publications (under construction)
- an image gallery of sample projects (under construction)
- software tools and data (under construction)
- links to related sites

The web page can be viewed at: http://www.geo.unizh.ch/ICA/

11 Joint Session of the Map Generalization and Visualization Working Groups

The workshop of the ICA Commission on Visualization also took place at this time. On the last day a joint session was held in which summaries of the findings of the two groups were presented and potential collaborative research themes were identified. Six points of discussion arose from the meeting:
- what methods can be developed for visualizing the effects of applying generalization algorithms (quality assessment and understanding)?
- what are the issues in the design of interface and metaphor in support of generalization operations?
- how can generalization techniques be applied to the presentation of information at varying levels of detail (LOD; generalization that is dependent on the viewing depth in a 3-D displays)?
- how can generalization techniques be extended to multiple dimensions in both three and four dimensions (spatio-temporal generalization)?
- can ideas in generalization be laterally applied to development of intelligent hyperlinks that connect and disconnect as the map detail changes?
- how can generalization techniques support dynamic display in virtual reality?

It is hoped that future meetings will report on collaborative efforts in these areas.

12 Conclusion

Some exciting developments have taken place over the past two years since the last meeting in Barcelona. Beyond some important implementations of specific algorithms, researchers are also beginning to model (and implement) strategies in design. The need for evaluation of outcomes has given impetus to research into quality indicators in generalization. There is an increasing sophistication of generalization operators, both in their effect and their ability to consider the context of their operation. There is an increasing acknowledgment of the importance of semantics in influencing geometric generalization, but precisely how 'purpose' translates to a set of object transformations is still not clear.

The framework necessary to support the process of identification, resolution and evaluation continues to clarify data model requirements and the need for data enrichment. Researchers continue to explore new paradigms and models for generalization, acknowledging the importance of structures that support deep knowledge of the neighbourhoods, associations and distributions of phenomena (for example through the use of tessellation techniques). The various models used raised the issue of interoperability - how to transfer data among different models whilst

minimizing cost in terms of quality and processing time. The importance of structuration as part of the intelligent application of generalization algorithms was again clearly made.

The context in which generalization research is undertaken continues to change. The creation and maintenance, by NMAs, of a number of databases at varying scales does not obviate the need for continued research in generalization. Increasingly issues of integration and Internet dissemination are refocusing and giving renewed emphasis to work in map generalization. Interest in this research is clearly illustrated through award of European funded research, interest and participation by a number of NMAs and from major GIS vendors. The deeper understanding of design and cartographic modelling over the past two years is complemented by a broadening of its application domain to a range of issues fundamental to GIS and geographic modelling more broadly. This is acknowledged in the growing number of researchers in the field. It is clear that a broadening range of disciplines are contributing to developments in map generalization and the emphasis is less on rule specification and more on the logic of space, tools for manipulating space, and modelling of purpose.

12.1 Trends
There is a growing awareness of the importance of generalization and collaboration between public sector, data providers, value added re-sellers, industry and academics. There is continuing growth in international funded projects.

More generally there is a recognition of the changing priorities in generalization: beyond map production to include issues of exploration and visualization, data interaction and analysis, and in database construction in the integration of multi-scale datasets.

On the technical side, it is certainly the case that more sophisticated data models are now being used, and semantics are being modelled within the database, for example the rule based methods within the IGN's Stratège system. More difficult problems are being addressed (contextual generalization), and decentralizing operators are being developed for generalization across distributed databases. Finally researchers continue to experiment with paradigms such as object orientation, multiple agents, and AI techniques (rule bases, semantic nets, neural nets).

12.2 Pointers to future research
Perhaps more than most, the following six issues seemed to run through the various group discussions and to a large degree reflect the intended focus of the 1999 Ottawa meeting.

Quality measures
A range of geometrically based measures exist for measuring change resulting from map generalization (cluster analysis, Voronoi, spatial statistics, graph theory, etc.). But how meaningful are they in the context of assessing candidate solutions? Beyond measures of gestalt and pattern analysis, are there summarative techniques that can be presented to the user that convey the objective of minimum loss of meaning and optimal reduction in data content?

Translating purpose into a set of generalization operations
It is well known that 'purpose' dictates content, scale and semantic emphasis (the balance between representing information that is merely contextual as a backdrop to the salient information and relationships you wish to emphasize. But precisely how do we model this

process such that it is possible to derive a sequence and degree of application of generalization operators to a range of objects?

Candidate designs
Invariably a range of solutions exist to any one generalization problem. A map generalization system should require minimum user input (beyond that of purpose definition) yet be capable of generating a number of potential candidate designs and to present these to the user. How can we develop strategies capable of generating a range of candidate solutions that are semantically driven in terms of map purpose, and geometrically constrained in terms of meeting the criteria defined by the intended use.

The interface and the balance in decision making
In the above context it is clear that decisions are being made on both sides of the computer screen between the system and the user. The issue then becomes what is the balance in the decision making process, how are decisions and quality indicators portrayed to the user? How are the responses that arise from this information fed back into the decision making process?

Structures and real time generalization for Internet access
The context and applicability of research in map generalization now extends well beyond map production, with an increasing focus on interactive map query, information retrieval and data integration that goes beyond a mere composite. There is need to provide timely response to user query and meaningful answers that include contextual information. Geographic data is typically complex and files are large. How can we minimize Internet traffic to reduce wait time during on line query yet generate solutions that are applicable to the entire region of interest?

Experimental platforms and empirical analysis
It is clear that a range of generalization operators exist coupled with a range of techniques for classifying objects according to their characteristic components. A variety of algorithms are now being implemented and tested on real data sets (such as network generalization, aggregation, displacement) for example through the work of the OEEPE. Such work provides a basis for development and refinement of these techniques and to identify 'missing' algorithms. It is hoped that this work can be extended among a broader community of researchers.

12.3 What next?
At the conclusion of the meeting, an attempt was made to identify and prioritise the impediments to advances in automated generalization systems. The top five items in the list will form the thematic focus of the next workshop meeting in Ottawa in 1999. The workshop will take place between August 11-13 immediately prior to ICA meeting. Those five items are:

- Methods for structure recognition (developing 'preprocessing' capacities for cartometric analysis, pattern recognition and map segmentation in order to characterize and classify object relations and representation)
- database enrichment techniques (encoding additional information in support of context dependent map generalization)
- Identification and development of missing generalization operators (from analysis and empirical testing of current generalization algorithms on meaningful test datasets)

- Knowledge acquisition and formalization (development of inductive/deductive strategies for establishing and formalizing knowledge from cartographic products and databases)
- Quality assessment techniques (development of measures, methods, and strategies for evaluating the quality of generalization solutions and algorithms)

It was also felt that live demonstrations provided an excellent basis for pragmatic discussion and that these and poster sessions will also be a feature of the meeting.

13 Acknowledgment

Citation for this paper:
Mackaness, W A, Weibel R, Buttenfield B P, 1997 Report of the 1997 ICA Workshop on Map Generalization. Gävle, Sweden, 19-21 June. Publication of the ICA Working Group on Map Generalization, http://www.geo.unizh.ch/ICA/