PRACTICAL RESEARCH IN GENERALIZATION OF EUROPEAN NATIONAL FRAMEWORK DATA FROM 1:10K TO 1:50K, EXERCISING AND EXTENDING AN INDUSTRY-STANDARD GIS

Paul Hardy, Dan Lee, & John van Smaalen
ESRI Europe, Cambridge UK; ESRI Inc, Redlands USA; and ESRI NL, Rotterdam NL
phardy@esri.com, dlee@esri.com, j.vansmaalen@esrinl.com

KEYWORDS: Generalization, Generalisation, GIS, Optimization, Constraint, Contextual analysis

ABSTRACT

Many nations have captured a digital master spatial dataset at a detailed scale, but need to derive smaller scale or lower resolution products. Where this has been attempted in the past, it has usually involved the creation of specialized software, or substantial bespoke development. This paper covers recent industry experience gained in making and assessing research prototypes for generalization in the 1:10K to 1:50K range, on European SDI (Spatial Data Infrastructure) national framework data. The trials involved a commodity industry-standard GIS (ArcGIS from ESRI). As well as making processing workflow models built with the standard geoprocessing tools, the projects involved writing some tailored processes taking advantage of the underlying ArcObjects functionality. Stretching technology further, the projects developed and exercised use cases and constraints for 'Optimizer' functionality currently in the research stage of software development, providing valuable inputs into development strategies for possible future capabilities. The paper summarizes the use cases investigated, the capabilities established, and the lessons learned.

1 INTRODUCTION

1.1 Netherlands RGI Generalization Project

A collaborative group of research partners are currently working together to investigate automated generalization of the Netherlands TOP10NL dataset (the primary SDI for the Netherlands) to derive TOP50NL data. These datasets are described in section 2. The partners are Netherlands Kadaster, ESRI Netherlands with ESRI Inc, and ITC Enschede. The project is partly subsidized by the Dutch government through the program 'Ruimte voor Geo-Informatie' project 'DURP ondergronden' (RGI002).

1.2 ArcGIS

The environment used in this research was ArcGIS 9.2, from ESRI [ESRI 2006]. The system offers an integrated set of tools for authoring, serving, and using geographic information. The principal ArcGIS component involved was ArcGIS Desktop, at the ArcInfo license level, which is a professional desktop application for data management, analysis, transformation, and visualization. Also involved were the ArcObjects toolkit which provides interfaces for tailoring and extending the system. The procedures developed can then be deployed in ArcGIS Server for non-interactive bulk processing.

1.3 EuroSDR State-Of-The-Art Study

EuroSDR (formerly OEEPE) is a European spatial data research organization that undertakes collaborative applied research projects. Membership consists of organizations representing national GI (Geographical Information) production and/or research throughout Europe with seventeen countries currently represented. EuroSDR is currently undertaking a 'state-of-the-art' study on generalization in commercial software systems, using test data provided by several European national mapping agencies (NMAs), including Netherlands Kadaster TOP10NL.

Testing of generalization capabilities is primarily being carried out by other NMAs and academic organizations using latest software releases. ESRI ArcGIS 9.2 was one of the software systems tested. However GI software suppliers (such as ESRI) also have the opportunity to carry out and report on parallel testing using versions of software beyond those publically available. This paper touches on such testing using a research prototype Optimizer to tackle urban building displacement, deletion and orientation.

2 NETHERLANDS DATA AND SCALES

2.1 Netherlands TOP10NL Data and Mapping

TOP10NL is a Dutch object-based topographic dataset aimed at visual representation at scales from 1:5K to 1:25K. TOP10NL is the successor of the non-object-based TOP10vector. TOP10NL consists of 14 feature classes (Table 1). The combination of land use, road and water polygon feature classes forms a geometric tessellation (full partition of the plane). Buildings are in a separate polygon feature class and overlay the land use layer (they do not cut holes in the terrain). Figure 1 shows an example of these TOP10NL polygon feature classes. As an exception, very large buildings which fill a complete ‘parcel’ are also represented in the land use feature class as built-up areas, with their geometry being identical to that used in the building feature class.

TOP10NL further includes road centerlines, linear water features (i.e. all water courses under 6 meters wide), and some additional point and line feature classes that do not play a large role in the generalization process. The latter are therefore not included in the EuroSDR dataset (see Table 1). Important buildings (post offices, police stations, churches etc.) are put into a separate feature class in the EuroSDR data. The feature class names, field names and attribute values in the EuroSDR data are translated from Dutch to English.

Table 1: TOP10NL feature classes and the EuroSDR selection.

TOP10NL	Geometry type	EuroSDR selection
FUNCTIONEEL_GEBIED_PUNT	point	-
GEBOUW_VLAK	polygon	building_area important_building_area
GEOGRAFISCH_GEBIED_PUNT	point	-
INRICHTINGSELEMENT_LIJN	line	-
INRICHTINGSELEMENT_PUNT	point	-
RELIEF_LIJN	line	-
RELIEF_PUNT	point	-
SPOORBAANDEEL_LIJN	line	railway_line railway_point
TERREIN_VLAK	polygon	landuse_area
WATERDEEL_LIJN	line	water_line
WATERDEEL_VLAK	polygon	water_area
WEGDEEL_LIJN	line	road_line
WEGDEEL_PUNT	point	road_point
WEGDEEL_VLAK	polygon	road_area

2.2 Netherlands 1:50K Data and Mapping

TOP50vector is the current Dutch topographic dataset aimed at representation at scales from 1:25K to 1:75K. Unlike TOP10NL, TOP50vector is not yet object-based but a product predominantly aimed at and developed for cartographic display. Two of the main differences between TOP10NL and TOP50vector are the amount of built-up area - in TOP50vector most of the urban extent is represented as built-up area instead of individual buildings – and the fact that roads in TOP50vector are stored as lines and symbolized for representation (see Figure 1).

2.3 Netherlands Constraints for Generalization

Generalization constraints were derived by Kadaster from the “Handboek generalisatievoorschriften TOP50vector” [Kadaster, 2006], the current guidelines for cartographers working on TOP50vector. Results are checked against the current topographic map at scale 1:50K to determine whether the automated procedures achieve what is intended by the generalization guidelines. When necessary, additional constraints are specified by Kadaster in consultation with ESRI.

One example where the existing formal constraints were incomplete is the treatment of thin areas of woodland or water. The illustrations below show the results of the traditional manual generalization producing the 50K product, in comparison with the base-scale 10K data. Of interest are the green woodland areas, which here are bigger than the minimum area constraint and hence should be kept, but are less than the minimum width constraint so must be enlarged. The original documented constraints however did not cover what should be moved (with what priorities) to make room for this enlargement, nor what should have priority if there is a conflict (e.g. if both narrow water and narrow woodland need enlarging, but lie between two immovable roads).

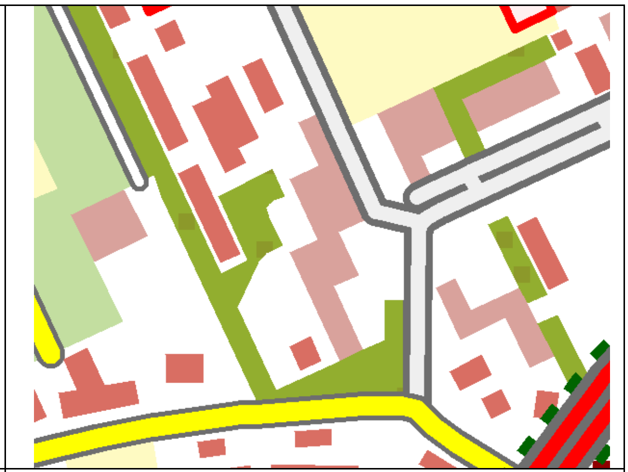


Fig 1a - 10K data, enlarged for 50K (viewed at 200%) | *Fig 1b - 50K data, manually generalized (at 200%)*

2.4 Need for Additional Information

The automation of constraints occasionally requires additional information, beyond that readily available in the TOP10NL data model. Nearly all land use features within the urban extent that have buildings on them are of the land use type 'other'. Constraint TDK3-5 states that land use features within the urban extent with land use type 'other' that are covered by buildings for more than 10%, should be converted to land use type 'built-up'. There are two exceptions to this constraint: terrain features in industrial zones are not converted and terrains with detached houses are not converted. The problem is that none of these properties: 'urban extent', 'industrial zones' and 'detached houses' are part of the TOP10NL data. In these cases additional information is required. In the project we were able for the test areas to retrieve the information needed from other Kadaster datasets.

3 GEOPROCESSING MODELS

3.1 Geoprocessing Framework, and Standard Tools

Managing, manipulating, deriving, and representing geographic data are essential parts of a GIS and mapping project. Geoprocessing in ArcGIS provides a rich environment to automate simple or complex tasks and to support the creation and sharing of workflows. A large and growing set of tools in ArcToolbox are available to perform operations on geographic data, for example to generate buffer zones around cities or to calculate values for an attribute field in a table. A sequence of operations can then be chained together using the ModelBuilder or scripts to formalize a procedure or workflow for solving complex problems.

To better support multiple-purpose and multiple-scale data and map production, standard tools have been made for generalization and cartography. For example, shared boundaries can be simplified consistently and marker symbols can be aligned to nearby line symbols. Although each of the existing tools involves limited contexts, the extendable and flexible geoprocessing framework has laid the foundation for advanced analysis and processes, such as optimization (see Section 5), to be included in the workflow.

Models using combinations of standard tools were created and used in the research, to prepare and enrich data for generalization. Implicit information, such as urban patterns, were analyzed, characterized, and made explicit as new data or attributes.

3.2 Urban Pattern Detection

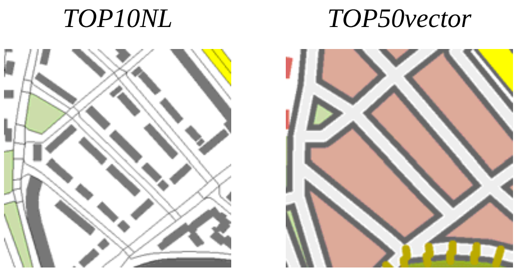
One goal of map generalization is to represent the geographic reality as faithfully as possible under map scale restriction. Geographic features are spatially and semantically related, and interfere with each other in many ways - some are topologically connected, others in relative positions. Geographic patterns - natural subdivisions, cultural areas, clusters, or alignments, are usually easy for human eyes to catch, but are not usually explicitly stored in a database. Both model and cartographic generalization share a common challenge – they must recognize, preserve, and often exaggerate these characteristics.

Patterns would be easy to recognize if features are positioned in perfect configuration, such as equally spaced, aligned to a straight line or other regular shape, symmetrically laid out, and so on. But, in the real world, features may form some patterns close to being regular, but never perfect, varying in many ways. Analytical methods and measures may be applied to help identify patterns with some success, but they are often sensitive to these variations or inconsistencies.

Geographic patterns are particularly important in urban areas, where the works of mankind are primary influences on the landscape. The following examples illustrate how geoprocessing tools were used creatively in urban area building pattern recognition.

3.3 Urban Building Pattern Detection

According to an analysis of the Netherlands TOP10NL and TOP50vector products and 50K cartographic map specifications done as part of this project, areas with buildings above a certain density that conceal the enclosed area from the road are aggregated into built-up area. Typical such high density building areas are shown, with their resultant built-up areas, in Figure 2.



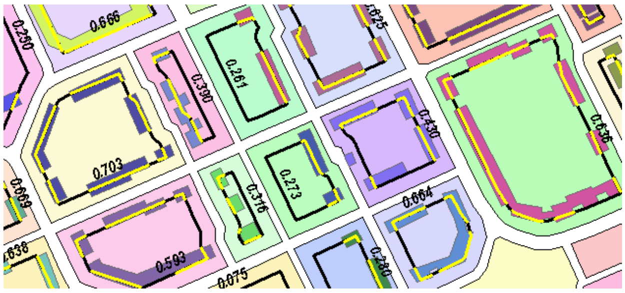


Fig. 2 - Aggregation into urban areas.

Fig. 3 – Detection of enclosing building patterns

3.3.1 Enclosing Pattern

The enclosing building patterns shown in Figure 2 seem to have the following characteristics:

- Buildings are within certain distance range from their associated street block borders.
- Many buildings have the longer side along the nearest street, although not necessarily parallel.
- Buildings have relatively smaller gaps between them – the smaller the total gap length, the stronger the enclosing pattern is.

A geoprocessing model was built to do the analysis and calculation using existing tools without any custom programming. The geoprocessing model and detailed steps were presented in a previous paper [Lee and Hardy, 2007]. The strategy was to identify buildings along an inner buffer from each street block, find where the buffer line overlaps these buildings and add them up to a total length of the overlapping segments, and then calculate the ratio of this total segment length over the buffer line length for that street block. A building with the longer side along the nearest street give a better chance of having its main length being passed through by the inner buffer, and therefore more truthfully contributing to the enclosing ratio calculation.

In Fig. 3, black lines are buffers constructed using a negative distance inside street block polygons; yellow line segments are where these buffer lines pass through buildings. The ratios of the line segments (total length per street block) over their corresponding buffers are as labeled. The segment ratio values can be used as a contributing factor in determining whether or not an enclosing building pattern exists inside a street block, therefore, these buildings should be aggregated into a built-up area. The higher this ratio is, the stronger the enclosing pattern is in a street block.

3.3.2 Well-chained Pattern

Some blocks may have buildings evenly around with similar gaps; while others may have the majority of buildings lined up around only part of the block. Further analysis can be done to distinguish seemingly "well-chained" buildings from others. Buildings appear nicely "chained" when the gaps between them are more even and short in length. Where a big gap occurs, the chain looks discontinuous. It makes sense to measure the gaps (black portions of the buffer lines) – shorter gaps indicate stronger chain pattern of buildings.

A geoprocessing model was built to extract closely chained buildings. The geoprocessing model and detailed steps were also presented in [Lee and Hardy, 2007]. In short, a buffer ring is created inside each street block at a distance such that it would pass majority of buildings. Then the gaps between buildings can be obtained using an overlay process that "erases" the ring by the buildings. Through a spatial query, the buildings touched by gaps shorter than a threshold can be identified.

The resulting closely chained buildings are shown in Figure 4a. If the majority of buildings are well chained in a street block, they are good candidates to be aggregated to built-up areas. They may or may not fill an entire block due to other factors. The result is displayed on top of existing TOP50vector urban area polygons (yellow background polygons) in Figure 4b. Where the chains discontinue (indicated by the green lines), no built-up areas are formed. Obviously a constant buffer distance may miss some buildings; additional analysis or adjustment to this approach can be made to improve the results. One possibility is to consider multiple buffers within a reasonable range and then choose the one that yields most chained buildings in each street block.

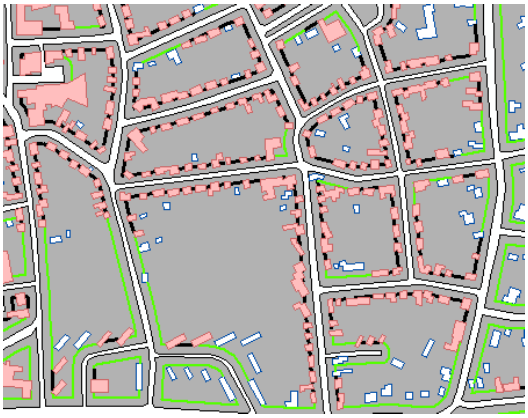


Fig. 4a - Closely chained buildings have short gaps (black lines in the image)

Fig. 4b - The identified closely chained buildings seem to fall in the existing TOP50vector urban areas

To summarize this example exercise: sufficient building density (e.g. above 0.20), high segment ratio (e.g. greater than 0.50), along with short gaps between buildings inside street blocks are quantitative measures that support decision making about urban areas. Possibilities for further study are to model the combination of the above three or more factors to determine the final candidate street blocks in which buildings should be aggregated into urban areas.

On the other hand, buildings or street blocks that don't meet the above criteria indicate that different generalization strategy should be applied based on additional measures and analysis. If time permits, further testing will be done to refine the building pattern detection, to use the pattern information to guide model generalization processes (selection, aggregation, simplification, etc.), and to feed the further generalized result along with enriched street blocks into the Optimizer.

3.4 Road Area to Centerline Collapse Use Case

Going from TOP10 to TOP50 roads, the most important generalization operation is the collapse of road polygons to road lines. In TOP10NL, the roads are represented as polygons as well as centerlines (Fig. 5a). In TOP50vector, roads are represented as centerlines only. As TOP10NL offers centerlines, it would appear logical to use these for the TOP50 roads. There is a drawback though, as in TOP10NL they are centerlines of individual lanes whereas in TOP50vector they are the centerlines of entire road segments, including refuges, shoulders and cycle lanes (Figure 5b).

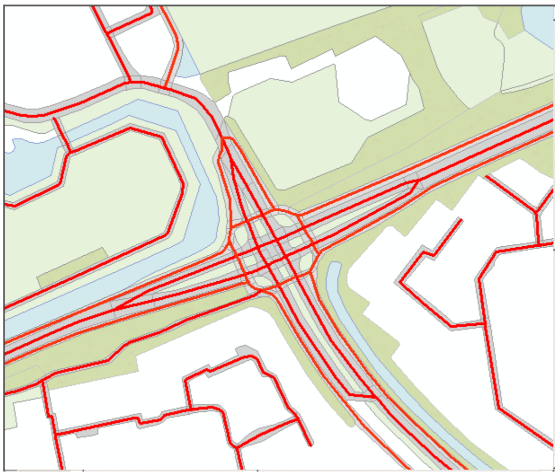


Fig. 5a - Centerlines in TOP10NL

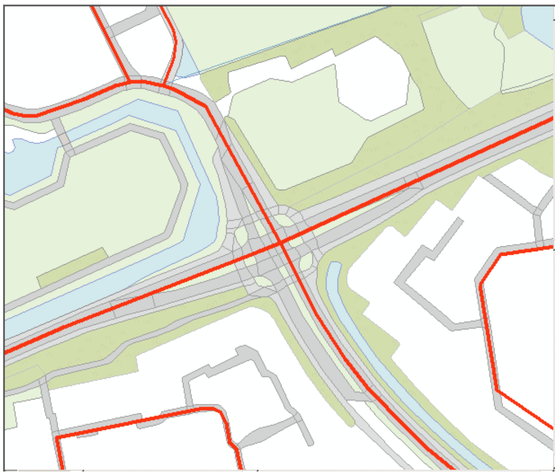


Fig. 5b - Road lines in TOP50vector

Choosing the most appropriate TOP10NL centerline to use as a TOP50 centerline is often not simple. Furthermore, in some cases, none of the TOP10NL centerlines are located at the center of the overall road.

So, in this project we choose to reuse the TOP10NL centerlines of single-lane roads, and to derive new centerlines for the multi-lane roads using the ArcGIS desktop centerline tool. The centerline tool however has limitations in that complex road intersections, common with multi-lane roads, are often not completely solved to a level adequate for 50K cartography. The project therefore fed examples of such use cases and current results back to the ESRI development team, to assist in evolution of the software over future releases.

Another point is that when deriving new centerlines for multi-lane roads it proves a disadvantage that TOP10NL does not differentiate between land cover and land use for grassland. Lacking this information makes it hard to distinguish grassy refuges from meadows and parks. By assigning the property ‘traffic divider’ to the grassy refuges and making it part of the road network instead of a land use parcel it would be much easier to derive centerlines running across them. Generalization of larger, complex crossings and roundabouts is rather troublesome without this additional information or reclassification. It is one example of the kind of information which is implicitly present in the source data (grass area topologically adjacent to two road areas), which is worth deriving and storing explicitly to aid in subsequent generalization.

4 TAILORED PROCESSES USING ARCOBJECTS

4.1 ArcObjects Toolkit

ArcObjects is a library of software components that make up the foundation of ArcGIS. ArcGIS Desktop, ArcGIS Engine, and ArcGIS Server are all built using the ArcObjects libraries. The same ArcObjects components are made available as a software development API (Application Programming Interface), for users to extend and tailor the capabilities of ArcGIS.

4.2 Use of ArcObjects to Create Additional Generalization Operators

Some parts of the generalization process are relatively straightforward, but standard tools have not been developed in ArcGIS. One example is the extension of terrain boundaries to road centerlines. In ArcGIS, it is possible to extend a linear feature to attach to another line, but this operation requires user interaction. Since in this project user interaction is not desirable, a custom batch tool (“Extend parcellines to road centerlines”) to extend all the lines in a feature class to lines in another feature class was developed in ArcObjects/.NET (Figure 6a). The tool is compatible with the ArcGIS (ArcToolbox) geoprocessing environment and can therefore also be used as a part of a ModelBuilder model (Figure 6b).

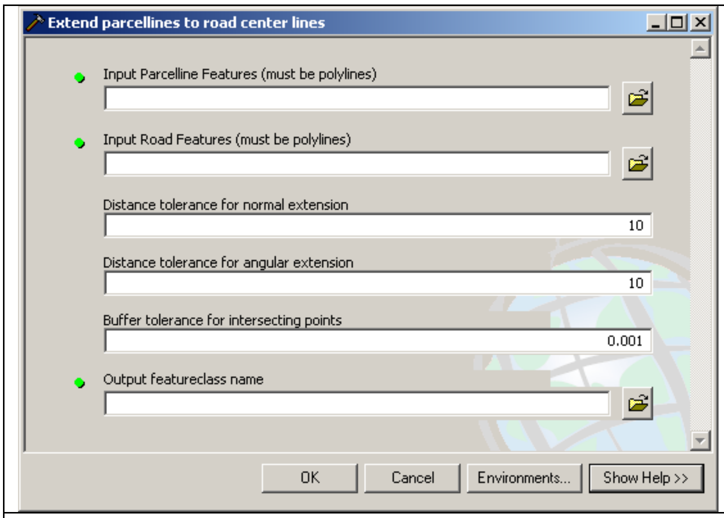


Fig. 6a – new geoprocessing tool written using ArcObjects

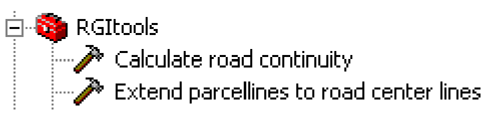
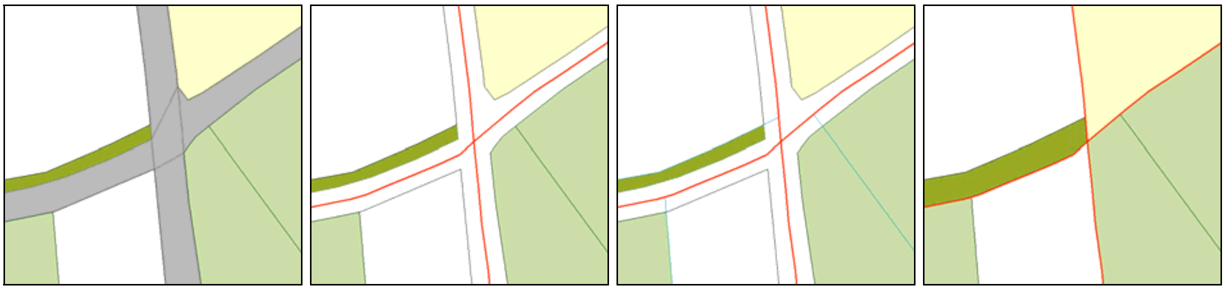


Fig. 6b – new tool appears in ArcToolbox

4.3 Extension of Land-Use Boundary to Road Centerline

By moving from roads as polygons (in TOP10NL) to roads as lines (in TOP50vector) there will be ‘gaps’ in the spatial tessellation (Figure 7b). This can be prevented by extending the terrain parcel to the road centerline. This involves the creation of new geometry (Figure 7c). The TOP50 terrain parcels become larger than their TOP10 counterparts (Figure 7d). One can argue whether this is desirable but for the moment we assume it is since it complies with the current TOP50vector data model.



(a) TOP10NL *(b) TOP10NL centerlines* *(c) Terrains extended* *(d) TOP50Vector*

Fig. 7 - Extending TOP10NL-terrains to the road centerlines

Extending the terrain instances to the road centerline is not a trivial task, for many reasons, notably:

- Firstly, ArcGIS does not currently offer a batch extend tool. Extending lines in ArcGIS is of course possible but requires user interaction, something to be avoided in this project. ESRI Netherlands therefore developed a custom batch extend tool in ArcObjects/.NET.
- Secondly, the length of the extended terrain boundary part is quite variable because of varying road widths and also the terrain boundary can meet the road centerline at a very small angle (almost tangential) or even not intersect with the road centerline at all.
- It is important that the centerlines are complete. Missing centerlines do not only lead to missing roads but also to unintentionally merged terrain parcels.

The procedure is as follows:

1. Merge the terrain_area feature class with the water_area feature class and with parking lots from the road_area feature class
2. convert the merged features to points (using option 'inside') for later use in reconstructing extended polygons in step 8
3. convert the boundaries of the merged features to lines
4. select parcel boundaries that are not road boundaries
5. add the study area boundary to the selection
6. store the selected lines as a FC (to-be-extended terrain boundaries)
7. extend the to-be-extended terrain boundaries to the road centerlines using the custom batch extend tool
8. take the extended terrain boundaries and the road centerlines and convert them to polygons using the points produced in step 2 as seeds for reconstructing polygons and as sources for retrieving polygon attribution

Analysis of the application of this process on a test dataset of the Wageningen area, showed encouraging results (Figure 8). This also shows that multi-lane roads cannot be properly collapsed to a single line when using TOP10NL centerlines and that is why additional centerlines will have to be derived for multi-lane roads.

Fig. 8 - TOP10NL parcels extended to TOP10NL road centerlines

5 OPTIMIZATION

5.1 Generalization as an Optimization Task

The concept of mathematical optimization of a system by convergent evolution has a long pedigree, with key points being the Metropolis algorithm [Metropolis 1953], and 'simulated annealing' [Kirkpatrick 1983]. There have been earlier academic applications of simulated annealing to generalization, notably for displacement [Ware and Jones 1998]. Statistical optimization (such as simulated annealing) is a useful technique for finding a 'good enough' solution to the class of problems where determining an exact solution would require exploring a combinatorial explosion of possibilities. The relevance of optimization to generalization is discussed further in [Monnot, Hardy, Lee 2007], which introduces the experimental design of an 'Optimizer' to run within the ArcGIS environment to tackle generalization processes. See Note 1 at the end of the paper regarding the status of this prototype development.

The Netherlands RGI project complemented the internal ESRI research, by providing realistic use-cases, test data, formalized constraints, example target results for comparison, and evaluation criteria.

5.2 Optimizer concepts

See [Monnot, Hardy, Lee 2007] for a detailed description of the concepts underlying the Optimizer, but it is important to understand that it is driven by a set of user-defined rules, with each rule made up of a constraint with one or more associated actions. Each constraint is defined as a feature satisfaction function, returning a value from zero (totally dissatisfied) to 1 (totally satisfied). Each action is defined as an algorithm which could improve the feature satisfaction against the constraint.

The task of the Optimizer is to maximize the overall system satisfaction, summed across all selected features and all active constraints. It is usual to apply the Optimizer in stages, applying it in turn to each spatial partition. The Optimizer implementation includes feature caching and partition iterating capabilities to facilitate this. The feature caching mechanism also provides topological structuring, and an 'undo' mechanism to allow the Optimizer to traverse the tree of possible actions, usually selecting the best available. This choice is however subject to a statistical chance (decreasing with a decaying notional 'temperature') of accepting a temporarily 'worse' state, in order to escape from a local minimum – e.g. temporarily overlap a feature in order to get to an open space beyond.

5.3 Preprocessing before Optimization

It is also important to note that the Optimizer is not a stand-alone tool. It fits into the ArcGIS geoprocessing framework, and for the Netherlands projects, we made extensive use of existing geoprocessing tools to do tasks such as to select, simplify and aggregate buildings to give a starting point for optimization (see Figure 9 below).

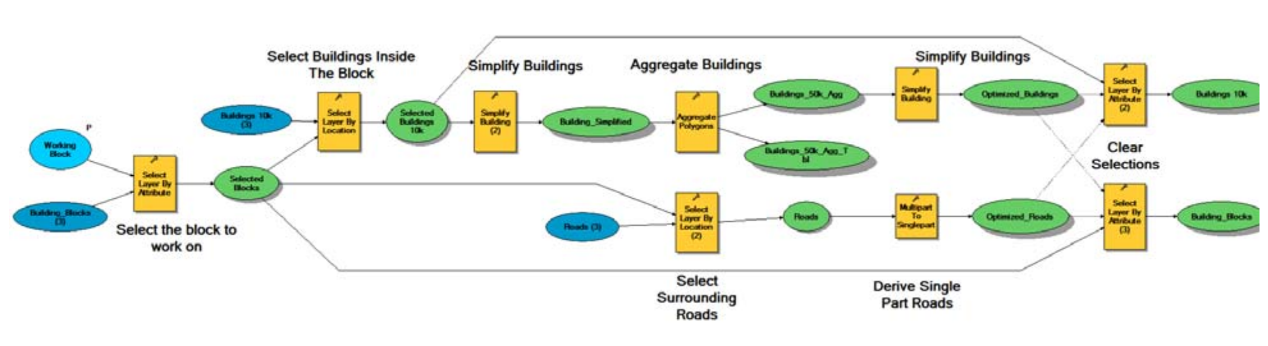


Fig. 9 - ModelBuilder script to prepare buildings and building blocks for subsequent optimization

5.4 Building Generalization Use Case

A primary use case for generalization in the 10K to 50K scale range is that of urban building generalization. In urban centers where building densities are very high, then the whole block is filled, but for less dense areas, then the buildings may need to be:

- Displaced away from roads
- Displaced away from buildings
- Not moved over significant space (like water)
- Enlarged to minimum size
- Compressed if influenced by opposing forces (between roads being exaggerated)
- Eliminated if too small, or no room

- Typified to a representative pattern
- Simplified in outline (or sometimes reduced to square or rectangle).
- Rotated to be parallel to road
- Aggregated (not done in Netherlands traditional rules)

Constraints for most of these were established in the Optimizer, and multiple data runs carried out, adjusting input parameters (e.g. minimum size thresholds). The results were encouraging, as shown in Figure 10, and enlarged in Figure 11.

Original 1:10K buildings reduced, with 1:50K roads

Resultant 1:50K buildings (at 200% for clarity)

Fig. 10 - Building generalization prototype by Optimizer

The Optimizer was run on all the primary urban partitions, as defined by the road network. This included many dense partitions which according to the Netherlands Kadaster rules would normally be converted to solid fill urban block terrain parcels. Applying the Optimizer process to these dense partitions allowed the identification and investigation of interesting situations of conflicting constraints, which otherwise would occur less often. Fig. 12 shows a useful diagnostic or metric, being the average satisfaction of all features in the partition.

Obviously just using the roads to define the edges of urban partitions is not adequate. As a bare minimum, lines of water features should be included, plus other interruptive features such as railways. This matches and extends the ‘trans-hydro’ partition concept of [Timpf 1988], which use the transport and hydrological network lines. However, at the scales concerned here, it became clear that it was also important to include certain land areas in the partitioning, notably woodland and water. This prevents buildings being displaced into woodland or over water.

Fig 11- Enlarged view of area where small buildings have been enlarged to rectangle, then displaced

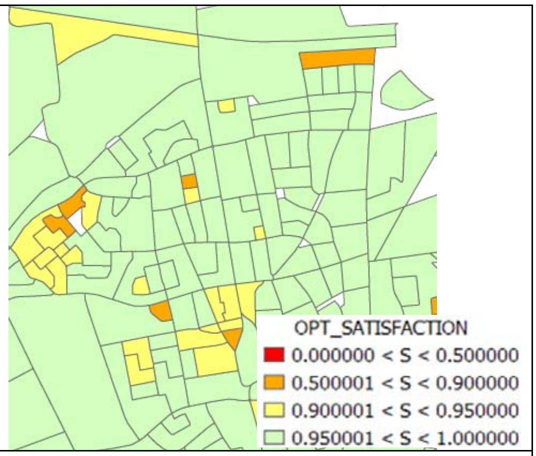


Fig 12 - Average satisfaction of all buildings in block can indicate areas where other techniques are needed

6 EUROSDR STATE-OF-THE ART STUDY

6.1 Testing by NMAs and Academics

According to the EuroSDR test plan, three testers, the Netherlands Kadaster (TDK), the Institut Cartogràfic de Catalunya (ICC) and University of Hannover were assigned to test ArcGIS 9.2 release. At the Barcelona meeting, a joint presentation given by Karl-Heinrich Anders from University of Hannover covered:

- some detailed testing examples and comments from TDK
- initial (pre-testing) comments from Univ. of Hannover
- no comments from ICC due to unexpected delay of their testing

As the presentation pointed out, ArcGIS is a complete and extendable GIS with rich documentation, good support, and many general tools. ModelBuilder is helpful for creating sequence of processes. In terms of generalization, some tools are available; more tools, such as displacement, enlargement, collapse polygons to lines, and so on, are needed. Areas to consider enhancements and functionality include support for additional constraints, global optimization, and ready-to-use generalization workflow; in other words more automation is desired to reduce manual editing and scripting work.

There were no real surprises in this test results, as previous analysis had indicated the gaps in the software capabilities, and many of these (lack of displacement etc) are very much in scope for the Optimizer development project.

6.2 Testing by Software supplier (ESRI)

The EuroSDR testing regime allows for software suppliers using development and experimental software. The RGI project work done using the Optimizer could not be directly submitted, as EuroSDR requires use of its own datasets (which are in a simplified and translated data model), constraints and testing forms. Therefore the ESRI team is repeating similar tests using the EuroSDR data for Kadaster. Results from this will be included in a later version of this paper.

7 CONCLUSIONS

- This kind of project, which combines representatives from NMAs, academia and industry to research practical progress in generalization can be of real benefit to all sides.

- Generalization, although still a hard problem, is moving from being an unsolved mystery, to being a well-studied process which is amenable to solution using tools in a commodity industry-standard GIS.

- Optimization is an appropriate tool to apply to the constraint-balancing aspects of generalization, such as displacement operations.

- Other operations can be carried out using standard GIS software, as combinations of existing tools in geoprocessing models, while more specific requirements can be solved using tailored tools constructed easily from available GIS components at the developer API level.

NOTES

1) This paper is a forward-looking research document, and many of the capabilities it describes are in the early stages of prototype development and review. Not all prototypes make the transition into product. As such, this paper must not be interpreted as a commitment by ESRI to provide specific capabilities in future software product releases.

2) The projects described above involved many people from NMAs, industry and academia. In particular, the contributions of Jantien Stoter of ITC Enschede, Maarten Storm and Harry Uitermark of Netherlands Kadaster, Jeroen van Winden and Ivo Visser of ESRI NL, and Jean-Luc Monnot and Weiping Yang of ESRI Inc are gratefully acknowledged.

3) Sample data extracts shown in illustrations is all copyright Netherlands Kadaster, with thanks.

REFERENCES

- ESRI Inc, “What is ArcGIS 9.2”, ESRI Press, Redlands, USA, 2006, http://www.esri.com/library/books/what-is-arcgis92.pdf [Downloaded 2008-02-21]

- Kadaster, “Handboek generalisatievoorschriften TOP50vector”, Apeldoorn NL, 2006

- Kirkpatrick, S.; Gelatt, C. D.; and Vecchi, M. P. "Optimization by Simulated Annealing." Science 220, 671-680, 1983.
- Lee, D "Geographic and Cartographic Contexts in Generalization", ICA Workshop on Generalisation, Leicester UK, 2004.
- Lee, D and Hardy P "Analyzing and Deriving Geographic Contexts for Generalization", International Cartographic Conference, July 2007, Moscow, Russia
- Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M.; Teller, A. H.; and Teller, E. "Equation of State Calculations by Fast Computing Machines." J. Chem. Phys. 21, 1953.
- Monnot, J; Hardy, P; and Lee, D "An Optimization Approach to Constraint-Based Generalization in A Commodity GIS Framework", International Cartographic Conference, July 2007, Moscow, Russia
- Timpf, S. 1998, "Hierarchical structures in map series", PhD Thesis, Technical University of Vienna
- Ware J.M. and C.B. Jones (1998) "Conflict Reduction in Map Generalisation Using Iterative Improvement'" GeoInformatica 2(4), 383-407.

[Issue Draft 0.9, of 2008-05-16]