Generalisation of webmaps of urban area

Module 11 – Elective: Generalisation in topographic mapping

September 1999

E. Mac Gillavry
GFM.2
Supervising staff: A. Brown & R. Hootsmans

[*Geographical*] *data sharing makes sense for the simple reason that there is only one Earth, and we share it.*

Kurt Buehler and Lance McKee, Open GIS Consortium Technical Committee

Foreword

This report is submitted in fulfilment of the requirements for the assessment of Module 11 (Elective: Generalisation in Topographic mapping) of the curriculum of the Master of Science course in Geoinformatics at the International Institute for Aerospace Survey and Earth Sciences (ITC) in Enschede, the Netherlands.

Contents

Chapter 1

Introduction

One of the main assets of a society is its environment, both the natural environment and the physical environment. In order to make decisions about the environment, good geographical information is necessary. This geo-information is derived from geographic data, data describing phenomena directly or indirectly associated with a location (and time, and orientation) relative to the surface of the Earth.

Geographic data have been collected in digital form for more than 30 years. The overall rate of collection increases rapidly with advances in technologies such as high-resolution satellite-borne imaging systems and global positioning systems (GPS), and with the growing number of people and organisations who are collecting and using geographic data. That number will continue to grow with the growing awareness among information technologists that indexing data by location is a fundamental way to organise and use digital data.

The modern day information-based society requires the development of a Geo-Information Infrastructure (GII), for it enables the integration of geographical data from the various sources. Integration is increasingly important because of growing environmental concerns, pressures on governments and businesses to perform more efficiently, and simply because of the existence of a rapidly growing body of useful geographical data and geo-processing tools.

Global sharing of digital geographical data via the Internet has become commonplace. The Internet has become a valuable tool for the dissemination of geographic information, providing new opportunities for the interactive visualisation of such information. By enabling the user to alter the scale at which data is displayed, the need for cartographic generalisation arises.

Exercise

The aim of the exercise is, given a large scale map of an urban area, to produce a generalised version which can be used for maps at 1:10 000 scale and, with minimum modification, also for 1: 25 000. The second aim is to produce a simple attribute database for the map features, in the first place to allow selection and symbolisation of features and secondly to give more information about each feature. The database should be structured in such a way, that generalisation from 1: 1250 to 1:25 000 scale maps is possible as well.

Methodology

The planning phase of this exercise is extremely important. After downloading the data, some considerations have to be taken into account before proceeding to the next step:

- What is the content of the 1:10 000 map and how does it compare with the content of the Land-Line data set?
- What generalisation operations have been done on the Land-Line map to produce the 1:10 000 map?
- What further adjustments are needed to produce a map at 1: 25 000 scale?
- How should the digital spatial data be structured to allow easy production of maps at the above mentioned scales?

Chapter

2

Generalisation for the Web

Introduction

In every map the objects and the forms of the Earth's surface are represented reduced. Only at larger scales, i.e. at lesser reduction, can the representation be approximately true to reality. As the map contents do not decrease proportionally in direct ratio to the reduction of the map size, an increasing density of the map contents arises at smaller scale. Generalisation is the process of creating a legible map at a given scale from a more detailed geographical dataset. It is the process of reducing the amount of detail so that the character or essence of the original features is retained at successively smaller scales. The purposes and benefits of generalisation are manifold (Spiess, 1995, p.31).

- Applying generalisation, we are able to describe reality with different degrees of abstraction, concentrating on the essential information for each group of users.
- Generalisation allows for modelling spatial databases.
- In order to visualise spatial databases some degree of generalisation is unavoidable: generalisation allows us to enhance visualisation
- Generalisation is the tool to render relevant information legible at a given scale
- Generalisation allows us to retain the optimal amount of the original information of a spatial database in a given map scale or image format.
- Generalisation allows us to remove noise in an image and enhance essential parts, not only graphical but also conceptual noise or redundant information.

Especially in built-up areas, many features compete for mapping space when represented at smaller scale than at which these have been surveyed. Trying to represent them all would turn the map illegible. Therefore, built-up areas always need some degree of generalisation. The next paragraph gives an overview of the different traditions towards this cartographic challenge.

Generalisation of urban area

The approach taken towards the generalisation of urban areas differs from county to country. On the one hand there is the continental tradition as in Switzerland and Germany. They are characteristic for the way most national mapping organisations (NMOs) on the continent perform the generalisation of urban areas. Topographic mapping focuses very much on the physical structure of settlements. There are two reasons for this. First, during the last century, geographers were very much concerned with the study of patterns of settlements. From this pattern they could derive the ethnographic background of its inhabitants. Second, because of the unstable political situation in Europe all through the ages, topographical maps were mainly the interest of the military. They are especially concerned with the protection of its inhabitants who live in cities. Good maps (in this case, detailed maps) of urban areas were a valuable asset.

On the other hand, there is the United States with very coarse generalisation. Maps of the 1: 24 000 and 1: 62 500 scale of the United States Geological Survey (USGS) have the same contents and same symbolisation (Thompson, 1979, p.42-48). Settlements are shown as area features, only with specific buildings, landmarks, distinguished. Even on the 1: 24 000 no individual buildings, except for landmark buildings, are displayed, though there is enough mapping space (Thompson, 1979, p.39). This results in easy generalisation, taking not too much time. There are two historical reasons for this. In the first place, the USGS is mainly concerned with natural resources, as is denoted by its "G", whereas there mandate to provide topographic maps stems from a later date. In the second place, the level of generalisation as maintained in Switzerland and Germany exists due to the fact, that they can boost on a long tradition of topographic mapping. This tradition does not exist within the USGS and therefore they can not rely on the experience of the staff.

Somewhere in between these two “extremes” there is the British Ordnance Survey. The generalisation of settlements on the OS maps is simpler when compared to the topographic maps of the continental NMOs. Settlements are represented as area features already at larger scales then one would expect in the continental tradition. The amount of detail relating to buildings and similar features which can be shown in built-up areas is limited (Harley, 1975, p.7). This applies particularly to individual buildings. This results in a lot of block-like structures on the map.

But not only the contents of the map may require generalisation. Another incentive for generalisation is the technical reproduction capabilities (Swiss Society of Cartography, 1987, p.15). In this case, the characteristics of the Web as a medium to communicate maps gives a raison d’être for generalisation. This aspect is looked into in the next section.

Webmaps: specific character

One of the limiting factors for web maps is the interplay between the size of the computer screen and the addressibility, often termed screen resolution. This defines the size of the map. Standard computer screen sizes are 14, 15, 17, 19, 21 and 23 inches. Standard configurations for addressibility, i.e. the number of pixels in the bitmap, are e.g. 640 x 480 or 1024 x 800. Trying to put a large array of pixels on a small screen leads to illegible fine details and small text. For example, a 14-inch monitor gives good resolution at 640 x 800 or 800 x 600, but not higher (Brown, 1993, p.130). Therefore, most webdesigners still take an adressibility of 800 X 600 as a standard for developing websites (<http://webreview.com/wr/pub/1999/06/18/poll//results.html>, <http://www.lynda.com/poll/vote.cgi?poll=19990708&option=results&imageField2.x=30&imageField2.y=2>).

The absolute size of maps as they appear on the Web browser depends on both pixel resolution and monitor size. As both the resolution and monitor size differ amongst users, this variation could result in image area varying by a factor of 9 (Stynes et al., 199., <http://www.geog.le.ac.uk/>). Absolute size is important when considering angular separation as well as intra-map comparison.

When designing maps for the Web, they need to fit on the screen, inside the Web browser. On a standard 14-inch monitor this leaves room for a document of 10.5 by 7.5 inches. Subtracting the space necessary for other page elements like a navigation toolbar, leaves a map size of just 9 by 5 inch (Plewe, 1998, p.181). This limited size of the map display is a constraining factor (Elzakker & Koussoulakou, 1997, p.9). Designers can get over the problem of legibility by generalisation. Different categories of data are put into different layers of the layer structure built into the software (Harrower, <http://maps.unomaha.edu/NACIS/> p.13). As scale is reduced, the less important layers can be switched off or new layers intentionally designed for small-scale representation may be shown.

To deal with limitation of map size and still facilitate interaction with an extensive database, the user will often have to zoom out (which will render small details illegible), or scroll and pan around the map, which means that he does not see the whole map at a glance (Brown, 1993, p.130).

Two drawbacks can be identified with the current approaches towards generalisation as outlined above. Firstly, they require the maintenance of a number of duplicate datasets at differing resolutions, thus introducing data redundancy. Secondly, the production of cartographically poor maps is possible - each representation operates over fairly broad scale bands, and is often displayed at a considerably different resolution to that at which it was captured.

These issues can be addressed using dynamic generalisation. Firstly, as every display is automatically derived on-the-fly from a single detailed dataset, data redundancy is avoided. Automated generalisation offers other advantages. Not only is the costly and time-consuming bottleneck of manual generalisation removed, but currency and consistency across the range of representations can be vastly improved. Secondly, the production of cartographically poor maps can be prevented by dynamically generalising the map according to the requested display scale, thus producing smoother transitions over scale. Dynamic generalisation also enables a range of map themes to be produced through selection of the features displayed. In the next two sections the subjects of generalising digital databases and automated generalisation are discussed in detail to give a background for the dynamic generalisation approach.

Generalisation of digital databases

The zooming- and panning abilities of current GISs and the layered data structure explains perhaps the historic lack of interest of the GIS community in cartographic generalisation (Müller *et al.*, 1995, p.4). However, there still is a need for the modelling of geographic information at different levels of abstraction. The automation of the generalisation process therefore remains a research subject with much interest. As an explicit programme or set of commands must be written to initiate each operation, it is easier to distinguish among the different generalisation processes (Robinson, 199, p.451). Generalisation can be separated into conceptual and graphic generalisation. These two types of generalisation are discussed in the next paragraph.

Conceptual generalisation

Conceptual or semantic generalisation consists of the decisions that have to be made before the actual representation takes place. Selection of relevant information from a geographic database implies powers of abstraction that depend upon an understanding of geographical concepts (Jones, 1997, p.271). This includes the following processes:

- Selection/omission of feature classes
- Reclassification of features into different or new feature classes
- Symbolisation of features and feature classes

Graphic generalisation

Graphic or geometric generalisation takes place as soon as features and feature classes are represented on a medium for communication. Graphic generalisation may be regarded as a process of increasing the level of graphic abstraction relative to the original surveyed form of geographical features (Jones, 1997, p.271). This type of generalisation encompasses the following processes:

- Simplification of shapes of features
- Exaggeration of size of features
- Displacement of features
- Graphic aggregation of features
- Graphic selection of features to preserve typifying characteristics

When automating generalisation the various processes as distinguished in this paragraph have to be implemented into a computer programme. The next paragraph addresses the various approaches for the automation of generalisation.

Approaches to automated generalisation

Artificially intelligent generalisation

Map generalisation is a tedious task, requiring skilled cartographers working for long periods of time. It is the common wisdom today that such labour-intensive tasks should be consigned to computers. However, this has been difficult to achieve. The expert system approach provides a structure for solving this problem. The collective lore used by humans in performing the generalisation process is systematically captured, classified and organised into an explicit list of rules (the “knowledge base”) and then applied by means of a logical inference programme to the generalisation task at hand. The concept of an artificially intelligent digital generalisation system suggests methodologies that reflect the theoretical aspects of human intelligence and, as such, could losely mimic the human generalisation process (Shea, 1991,p.6).

Object-Orientation

People have tried for years to build centralised “knowledge bases” of generalisation rules. In such systems, the map features themselves have just been passive items containing coordinates and attributes, acted upon by the centralised rules. In the object-oriented world, this is turned upside down (Hardy, 1999,<http://www.laser-scan.com>). The map features themselves become objects that have generalisation behaviours defined in the database schema. The application itself becomes much thinner, and contains no knowledge about what, how, or when. It merely provides a framework for invoking and sequencing the generalisation processes by sending messages to selected objects.

Each such object inherits behaviours from its object class definition and from superclasses in its class inheritance hierarchy. These behaviours allow the object to decide for itself what to do when receiving a message to generalise itself, e.g. it can inspect its relationships with its neighbours to decide whether to move itself. However, as the object modifies itself, any objects that are directly linked to it or spatially adjacent can also be told to reassess themselves, so that effects propagate.

Delauney triangulation

Generalising geographic features can be performed based on the determination of the structure of such features. This structure is determined by examining the space surrounding the feature. This can be achieved using Delaunay triangulation. In triangulating a structure it is desirable to create triangles that are as equiangular as possible. In doing so, individual triangles tend to be locally representative of the value of the surface. Another desirable characteristic of a triangulation procedure is that it should produce a unique triangulation, independently of the starting point or orientation of the dataset. The Delauney triangulation meets these objectives (Jones, 1997, p.200).

This approach has several applications. In the first place, triangulation can be used in the generalisation of contour lines. Digital Terrain Models (DTMs) can be stored in a Triangulated Irregular Network (TIN) structure. Triangulation provides a model for the controlled reduction of surfaces in that it is possible to apply vertical error tolerances to the selection of important points used in triangulation and to apply line generalisation procedures to linear features and form lines that are integrated with and constrain the triangulation (Jones, 1997, p.280). In the second place, triangulated data structures can be used to aggregate separate features by identifying the space between them and re-attributing the space between them with the classification of the aggregating features (Jones, 1997, p.284). Triangulation can also assist in finding appropriate vectors to be used in moving nearby features towards each other in order to close the gap between them. In the third place, triangulation can be used for generating a correct skeleton from vector data (Jones, 1997, p.284). Joining the circumcentres of the triangles forms the skeleton.

Raster approach

Another option for structuring space is the raster approach. Techniques to “generalise” raster data sets are essentially equivalent to simple pixel-based image processing operations, not respecting the object nature of “raster polygons.”

This approach has several applications. In the first place, this method can be used for line and area boundary simplification. Rasterised representations of linear features can be reduced in detail by using the mathematical morphology image processing techniques of dilation and erosion (Jones, 1997, p.279). In the second place, a raster procedure can be used for the aggregation of separate features. The area boundaries are dilated first such that nearby features merge into each other and then the image is eroded, reducing the size while retaining contiguity between the aggregated features (Jones, 1997, p.283).

Dynamic generalisation

Though the approaches described in this paragraph differ, they are not mutually exclusive. On the contrary, these approaches can be applied complementary. In order to support feature-based (geographical) generalisation, an artificially intelligent approach, using a “knowledge base” can be used (Mark, 1991, p.103). Adopting an object-oriented approach to representation enhances the implementation of such concepts. A key concept of the object-oriented approach is that of methods defined on objects (Hardy, 1999, <http://www.laser-scan.com>). These methods are stored in the “knowledge base”. When a method on an object is invoked by sending a message to the object, a behaviour bound to it is executed, possibly using values and references held by the object. Such behaviours can be for example the execution of pixel based image processing algorithms (raster approach).

For the implementation of dynamic generalisation for webmaps, the object-oriented approach seems a natural fit for the client side in of a DGI application, where generalisation methods, or operations, are only executed when called upon, resulting in the required "just-in time" character.

When dealing with DGI-technology, not only the map has to be transmitted across the Web, but also requests for new map views, searches and GIS-analyses (Plewe, 1998, p.63). Zooming, scrolling and panning therefore makes the map less attractive and much less interactive: it will take users more time to see parts of the map (Plewe, 1998, p.181). The limiting factor, is that the Web data base is stored on a distant server; the generated map must be formatted and compressed to allow transmission on the Web and its display on the browser in an acceptable time frame (Roche, 1998, p.5). The availability of adequate bandwidth is critical to enabling geographic applications to operate, by collecting data "just in time" over the Internet (Barr, 1998, p.15).

The amount of data-transmission taking place over the Web not only depends on the actual file size of the map view, but also on the client-server configuration, it is also important to take notice of the thickness of the client-side. The thickness depends on the computing capabilities of the client-side software: the browser with plug-ins. The thicker the client, the less data is transmitted, the faster the new map view is being created

However, it is common to expect the client side to be thin, both in its footprint and in its bandwidth requirements. To keep the client side thin means that higher level abstractions are required for communication over the Internet. The most common abstraction with spatial databases is the feature. However, in many cases it is necessary to communicate a "map" to the client, and this is ideal work for object-oriented selection, generalisation and representation mechanisms. An object-oriented database therefore is not only capable of answering Web queries about the properties of individual features, but can also serve up fully symbolised maps.

Chapter
3

Methodology

Introduction

There are always to aspects in developing a spatial information system. On the one hand there is the development of a spatial database filled with geographical data. On the other hand there is the development of a set of functionalities to handle thematic and spatial operations on the geographical data stored in the database.

To be able to perform dynamic generalisation on geographic features, the implementation should be carried out using a powerful object-oriented spatial information system and using a rule-based approach (Mackaness & Glover, 1999, <http://www.geo.ed.ac.uk>). In this case, the geographical data (see Introduction and Appendix) is structured in an object-oriented database. The functionalities for this object-oriented spatial information system are modelled as methods. The methods described here only refer to the functionality of the system to generalise the data stored in the database for portrayal at scales smaller than surveyed. These methods are stored as rules.

This has been implemented before using Laser-Scan's LAMPS2 mapping software. This is an automated mapping system built on Laser-Scan's object-oriented database Gothic, and as such is particular suited to the task of automated generalisation. Relevant functionality includes dynamic cartographic representation, some existing generalisation capabilities and a WYSIWYG screen display.

However, the Laser-Scan software is not available for the exercise. In principle any software or software combination can be used which allows editing of lines and additions of new lines, points, polygons and text, if necessary in new layers. Coupling with a database should also be possible, and it should be possible to produce webmaps. For this exercise, the software package ArcView GIS 3.0 of Environmental Systems Research Institute Inc. is used together with the Internet Map Server (IMS) 1.0 extension. In the next two sections the two aspects in the development of a spatial information system are described.

Data

Conceptual model

The conceptual design is the first step in database design where the contents of the intended database are identified and described. It leads to an understanding of the natural structure of data at a high level, independent of how the data is processed. It is developed based on knowledge of the organisation through the external model, not from any other information requirements. It is independent of any specific database technology. The conceptual model is then mapped onto an implementation model, or logical model. The conceptual model is produced with hindsight to the implementation system, creating a hybrid conceptual/logical model, because of time and resource constraints.

The complete set of Land-Line data contains 43 feature layers. The sample set of the area of Port Talbot used for the exercise contains only 17 layers. Out of this set of layers an initial selection is made (see Table 1. Initial selection of layers), based on the assumption that roads, parcels and buildings are the main objects in the urban area that need generalisation when represented at smaller scales than surveyed (Mackaness & Glover, 1999, <http://www.geo.ed.ac.uk/>; Robinson, 1995, p.241; Thompson, 1979, p.35). This selection should not be interpreted as a first step of conceptual generalisation, but a way to limit the scope of the exercise.

ArcView Shape file	Description	Land-Line layer name
buildout.shp	building outline	G8010001
buildpk.shp	building pecks	G8010004
roadpk.shp	pecked road metalling	G8010021
linegen.shp	general line detail	G8010030
align.shp	general alignments	G8010032
landlim.shp	veg.landform limit	G8010036
linemin.shp	minor line detail	G8010052
water.shp	water detail	G8010059
roadcl.shp	centre line public road	G8010098

Table 1. Initial selection of layers

From this initial selection of data layers three object classes have to be derived:

- Road (line)
- Building outline (area)
- Parcel (area)

These object classes are further described using object cards (Appendix B). Attributes are descriptors that describe solely that object. Unique identifiers are one or more attributes which uniquely identify the object. Relationships are associations between objects that need to be built.

Logical model

Though the Land-Line dataset is said to be apt for analysis using spatial information systems, there still seems to be a legacy from the analogue era. The structuring of the dataset is still very much organised towards the production of analogue. To represent these features in a logical model, the data in the layer files have to be restructured into new features. For example, the parcels are represented in the layer with general line detail as being just lines, not areas. Furthermore, the lines are not topologically digitised.

Buildings are represented not as individual buildings, but only the outlines are stored in the layer “building outline”. A semi-detached house for example is not individually represented. Its character can only be derived from the parcel border that splits the building outline. However, the building outline does not always constitute in itself the building. There are also the layers “building pecks” and “minor line detail” that form together with the layer “building outline” a whole building. There are also constructions as for example sheds, that are not represented in the layer “building outline”, but in the layer “general line detail”, not as area features, but as line features.

Another difficulty is the roads. There are two sources of difficulties. First, there are both centre lines, stored in the layer “centre line public road” and road casings, stored in the layer “pecked road metalling.” The centre lines are used for the generalisation, where the road casings provide input for the symbolisation of these lines. Second, apart from the public road, there are also private roads that are important to the topography of the area. These are stored in yet another layer “general alignments” and have to be put into the layer “pecked road metalling.

The inappropriateness of the dataset for spatial analysis is most apparent when dealing with parcels. Their cadastral and geographic nature would make one think that these features would be represented as areal objects. However, they are part of the general line detail, not even distinguished from other geographic, real world features that make up the layer general line detail.

The structuring of this data set therefore has to be re-arranged to reflect an object-oriented spatial information system. This editing is performed using Macromedia Freehand 8.0.1 together with Avenza MAPublisher 3.0.

Functionality

The behaviour of an object is described by its methods. These methods are rules that the object invokes when applicable or required, based on certain assumptions. The rules are small programmes that have to be executed. In this exercise the rules only dictate the generalising behaviours of the objects. Since there is no object-oriented software available, the programmes are only written in pseudo-language. The application of the methods to the objects is executed manually by inferring the rules and applying them on an object in a layer. In the next paragraphs the derivation of the methods for the different objects classes is discussed, based on a review of the literature and on the comparison of the maps of 1: 1250 scale and 1: 10 000 scale.

Road centre lines

Roads play an important role in the process of generalisation. In the first place, as a geographical feature there is a high priority for representation on various scales. Especially maps at 1: 50 000 scale are suitable for emphasising the communications network (Harley, 1975, p.9). Almost all roads in the 1: 10 000 maps are also present in the 1: 50 000 map (Højolt, 1999, <http://www.geo.unizh.ch/>). Since all maps ranging from scale 1: 1250 upto 1: 50 000 are generated from the same dataset, the accurate representation of roads on all scale levels is important. In the second place, roads are important for they are starting points for the generalisation by means of map partitioning. Map partitioning is a means for organising and decomposing map space into smaller, though still geographically meaningful divisions: meso level situations (Brazile & Edwardes, 1999, <http://www.geo.unizh.ch/>; Lamy & Ruas, 1999, CD-ROM). These situations are composed of buildings surrounded by a street cycle (Ruas, 1999, <http://www.geo.unizh.ch/>).

When generalising roads, the best way is to use the centre lines ((Brown, 1998, p.2). For representing the road by means of road casing, several buffers can be shown on top of each other. As roads are considered to be important geographical features, they should be represented at true scale, whenever possible (Harley, 1975, p.9; Thompson, 1979, p.26). In this exercise the buffer threshold is derived from the length of the vector perpendicular to the centre line public road from its origin until the intersection with the pecked road metalling.

Also great care has to be taken to preserve topology with a high degree of accuracy compatible with scale (Harley, 1975, p.9; Thompson, 1979, p.27). Too small irregularities in the lineament of the centre lines that will be covered by the envisioned line width, as set by the buffering threshold, have to removed (Swiss Society of Cartography, 1987, p.21). This requires some **simplification** of line detail (see Figure 1). There is a whole corpus of cartographic literature on only this subject itself. For simplicity's sake, the Douglas-algorithm is used, which is very good at deleting small irregularities while retaining important characteristics (Jones, 1997, p.277).

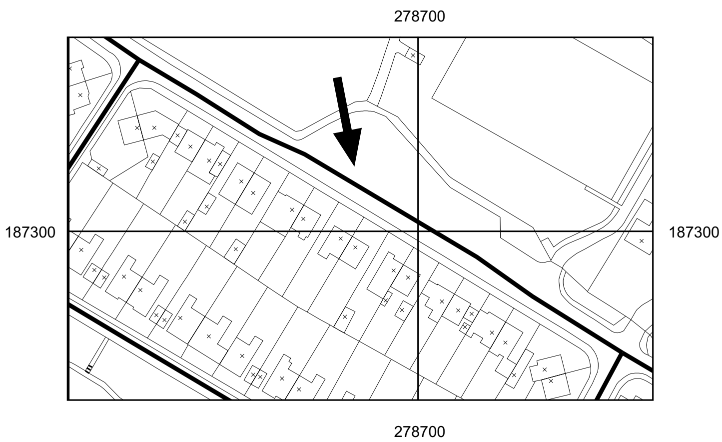

Figure 1. Here, the road centre line shows small irregularities that have to be removed

When representing roads at true scale, most roads are too narrow for clear delineation to scale, so **exaggeration** is needed, by means of using minimum symbol widths. This minimum width is defined using different thresholds. For example, the Swiss Federal Office of Topographic defines this width at 15 meter in reality when using double line symbology for maps of 1: 25 000 scale (Swiss Society of Cartography, 1987, p.19). Here, the arbitrary threshold of the US Geological Survey, defined as 40 feet in reality for mapping at scale 1: 24 000, is observed (Thompson, 1979, p.26).

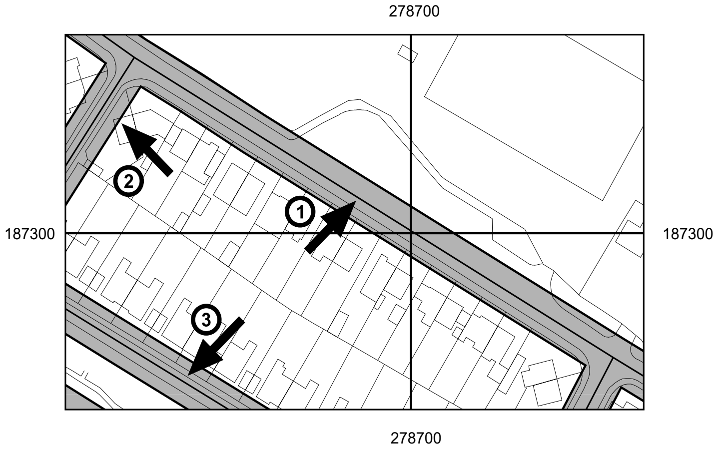

Figure 2. After simplication (1), several overlaps occur (2) and (3) due to buffering. Note that at some points buffers exceed the width of the origial road: exaggeration.

The method of exaggeration is implemented for generalisation to 1: 10 000 by buffering the centre line with a threshold of 6 meter, because the representation of the features in the standard single database at both scales will be the same.

However, there can be several parallel roads resulting in overlapping buffers. In order to maintain the topology of the roads, the roads have to be **displace**d. When there is an even number of parallel roads, all roads are displaced in opposite direction with equal distance until the buffers do not overlap anymore. When there is an odd number of roads, the position of the centre road is maintained. The other roads are displaced in opposite direction with equal distance until the buffers do not overlap anymore.

Figure 3. Overlapping buffers

Also for map partitioning centre lines are very adequate (Robinson, 1997, p.241). The algorithms used to derive street cycles are easier and faster to derive from linear centre line features (using some network-analysis: **connection**) than from area features. Based on this analysis, the buildings are assigned a **relationship** with the street cycle.

Building outlines

Aside from roads, buildings and groups of buildings are by far the most common manmade features shown on large-scale topographic maps (Thompson, 1979, p.35). In the previous chapter some words already have been dedicated to the generalisation of buildings themselves. Their generalisation is intertwined with the generalisation of road, because in a dense urban area, their graphic representations may overlap often. This overlapping may stem from different generalisation procedures. In the first place, due to the aggregation of single buildings, the space in between the buildings that is reclassified, may have been a road. In the second place the exaggeration of either the road or the building may be a source of overlap.

The first source is already taken care of by the **relationship** method of the object class “road”. This ensures that buildings can only belong to one meso level situation at a time: the street cycle, thus maintaining topology. The second source is dealt with later in this paragraph. First, the attention is drawn towards the generalisation issues that apply only to the object class “building” itself.

An important aspect when generalising buildings is their character attribute. The character of a building defines the application of methods to the object. Buildings are classified as being either general buildings or landmark buildings. A landmark building is any building that is distinctive in relation to its surroundings. Its distinctive character can be derived from its public function, historical association, conspicuous size, height, appearance or location. For this exercise, the presence of landmark buildings is derived from the textual layers indicating amongst others school buildings.

For example, its character attribute ensures the retention of small landmark buildings, whereas general buildings of the same size are deleted, thus resulting in a method for **graphical selection**. For landmark buildings, regardless of the actual size of the building, they are always retained, its symbol is never drawn smaller than a specified minimum size. This is an **exaggeration** method to ensure legibility. For both methods the same value can be applied. In the former method this value serves a threshold, in the latter as a standard. The US Geological Survey applies a minimum symbol size of 40 x 40 feet in reality for representation at 1: 24 000 (Thompson, 1979, p.36). In this exercise however, the graphical selection does not have to take place according to a threshold. All building outlines are stored in the layers G8010001 and G8010004. Smaller buildings (e.g. sheds) are stored in the layer G8010030 (see for explanation of layers Table 1. Initial selection of layers). Thus, the generalisation is not graphical but conceptual.

For any map there is a minimum separation below which neighbouring features become hard to distinguish (Lonergan & Jones, 1999, <http://www.comp.galm.ac.uk/>. Here, the character attribute of the building comes again into play to decide whether to apply the method of **aggregation** or **displacement**. If two general buildings are close to each other according to a certain neighbourhood threshold or share a common boundary, the buildings are aggregated. The aggregation should involve only general buildings that belong to the same meso level situation (Jones, 1997, p.283).

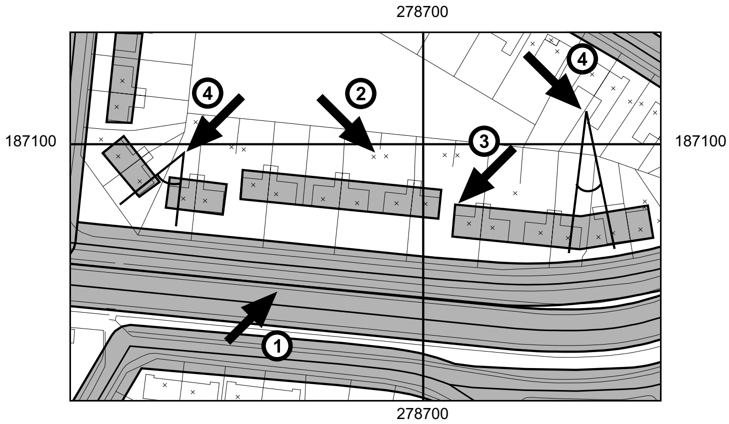

Figure 4. The buffers have been displaced (1) and small building outlines have been removed (2). Distance to road (3) and orientation (4) are important aspects when applying the aggregation method

Furthermore, the distance from the building outline to the road edge should be taken into account. If this distance differs too much for the features applying for aggregation, they should not be aggregated. Comparing the Land-Line dataset with the 1: 10 000 scale map results in another requirement for aggregation. The orientation of the general buildings involved in the aggregation process should not be too different. In most cases this excludes general buildings on a parcel that is situated on a corner, or, putting it topologically, on an intersection of two streets from the aggregation. This requirement ensures the preservation of the overall character of the pattern of buildings. It could also be derived from this comparison, that there should be a rule that states that houses that are on parcels situated at street corners should not be considered for aggregation. This derivation should not be made. First, this derivation requires a method to be invoked that applies to the object class “parcel” to detect if the parcel is on a corner, whereas parcels are not yet taken into account in the generalisation process. Second this would prevent the aggregation of building outlines which are both residing on parcels situated on street corners.

When taking only these requirements for the aggregation of general buildings into account, there still may be problems preserving the character. Consider for example a situation of semi-detached houses along a road with equal spacing, for this is more or less the case in this exercise. When applying the rules as set until now, the generalisation process will result in the aggregation of all the building outlines along a road in the dataset: resulting in the impression of a row of terraced houses! Therefore, a third rule has to be applied for the aggregation of general buildings to prevent this situation from occurring, e.g. the introduction of a rule stating a minimum threshold for the number of general building to be aggregated. For this exercise the threshold of 4 general buildings is observed.

If one of the buildings meeting the requirements for aggregation based on the neighbourhood threshold is a landmark building, the general building has to be displaced. In this exercise there are no cases where general buildings and landmark buildings have mapping conflicts for they are situated in different street cycles. If the buildings under consideration are both landmark buildings and do not share common edges, both have to be displaced. When it is necessary to **displace** one or more buildings of a group of landmark buildings, the positions of the other buildings is proportionally adjusted so that the pattern of the buildings in the group approximates the true pattern on the ground (Thompson, 1979, p.37). This minimises the loss of positional accuracy and maintains the gestalt and topology of the features (Mackaness & Purves, 1999, CD-ROM).

A simple solution to the implementation of a method for displacement involving buildings with different character attributes is to find the shortest vector connecting the two features and use this vector to displace the general building until they are separated by the minimum acceptable distance (Lonergan & Jones, 1999, <http://www.comp.glam.ac.uk/>). This is most likely the neighbourhood threshold as set for the method for aggregation. This method can be refined for the displacement when both features involved are landmark buildings by sharing the displacement between the features in some proportion based on their individual characteristics, e.g. the area attribute. Displacement never exceeds the meso level situation bounded by the street cycle (Mackaness & Purves, 1999, CD-ROM).

However, this approach may cause problems in the subsequent procedure of simplification. The simplification algorithm may result in a generalised representation that reduces again the minimum distance between two features below the neighbourhood threshold. This would require some degree of re-iteration of the generalisation process. This would have negative implications for the speed of the overall process.

The approach applied in this exercise is the “happiness-algorithm” (Mackaness & Purves, 1999, CD-ROM). For each feature meeting the requirements for displacement the algorithm considers a number of alternative locations close to its current location. For each location, the algorithm calculates the total “happiness” of the feature and moves the feature to the location at which the “happiness is at a maximum. “Happiness” is defined as the location among a set of alternate locations that produces the greatest reduction in overlap between itself and its local neighbours.

The rationale for observing this approach is that the features that serve as input for the method are not the building outlines themselves, but the bounding circle that encompasses the building footprint. This bounding circle is fitted with at least two points on the building's footprint touching the bounding circle. The circle acts as a required buffer of separation. Most likely a simplification method will not result in a generalised representation of the building footprint that exceeds this bounding circle. Therefore, this approach bypasses the need for re-iteration of the displacement and simplification methods, resulting in faster processing.

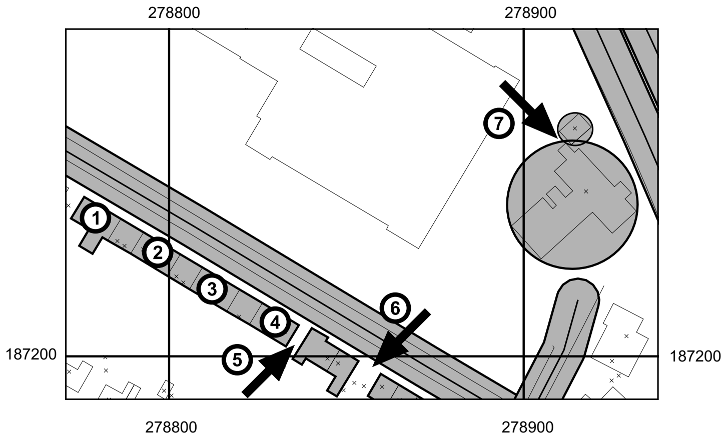

Figure 5. At most 4 general buildings (1-4) can be aggregated, although the space (5) to next building is small enough. If the space (6) is too large the buildings are not aggregated. In case of landmark buildings, buildings are displaced using bounding circles (7).

The original geometry of the building outlines in the Land-Line dataset is far too detailed to be represented at scales smaller than 1: 1250. Therefore, the building outlines should be **simplified**. Simplifying building outlines means finding a simpler representation of the original buildings by reducing details in their boundaries, while maintaining the essential shape and size of the buildings (Lee, 1999, < http://www.geo.unizh.ch/>). For aggregated buildings a simple algorithm could be applied that creates the minimum bounding rectangle, than scales it to obtain the area of the input feature again and rotates it to reflect the principle orientation of the original building (Mackaness & Glover, 1999, <http://www.geo.ed.ak.uk>). For simplifying individual buildings (landmark buildings and buildings situated on the corner of the street) the simplification algorithms of ESRI's latest ARC/INFO release can be applied (Lee, 1999, <http://www.geo.unizh.ch/>). This can also be applied to aggregated building outlines, though their internal structure should not be too complex.

The original geometry of the building outlines in the Land-Line dataset is far too detailed to be
Now that all the generalisation issues that apply only to the object class "building" itself have been looked into, the focus should again be shifted to those generalisation issues that apply to the object class "building" taking into account other object classes.

The first source of overlapping of buildings and roads has already been treated. The second source is a result of the generalisation procedures applied to either of the object classes. When interference occurs between a building and a linear feature of regular alignment, in this case the road, the linear feature is held in position (though it may have been **displace**d to prevent overlapping with other roads) and the building feature is **displace**d (Højolt, 1999, <http://www.geo.unizh.ch >; Thompson, 1979, p.36).

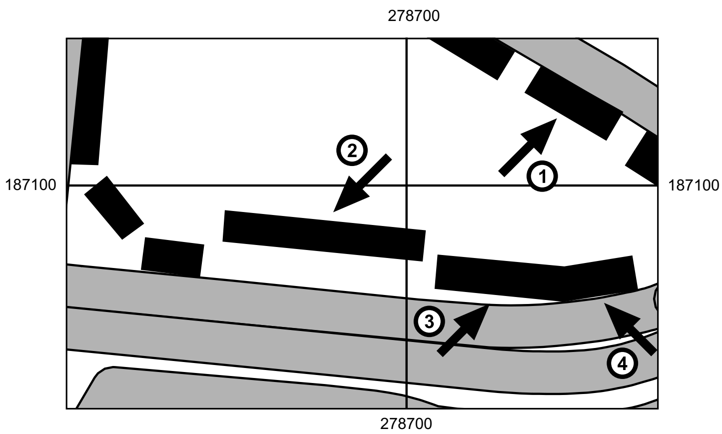

Figure 6. Buildings (1) close to the edge of the road (buffer) are moved towards the edge. Others (2) maintain their position. After displacement rotating (3) and smoothening (4) may be necessary

The displacement is in the direction normal (perpendicular) to the alignment of the linear feature towards the centre of the street cycle and in the smallest amount necessary for legibility. This requirement applies straightforward to building outlines that, in reality, share edges with the road sidelines. These should be displaced simultaneously, so that the common edges are conserved as common edges (Højolt, 1999, <http://www.geo.unizh.ch >).However, this applies also to other buildings that are within a certain distance to the edge of the road, resulting from the **buffer** method as applied to the road centre lines. Comparing the Land-Line dataset against the map 1: 10 000 scale one sees that some generalised building outlines are displaced towards the road edges, whereas others remain located separated from the road edge.

After displacement there may still be map space between the building outline and the edge of the road. For example, the building may still have an irregular boundary after simplification, or the orientation is different from the alignment of the road. Applying another simplification method can solve the first example. This method **smoothen**s the irregular building outline to a line, parallel to the road edge (either straight or curved). If there still is map space in between, the building outline has to be displaced once more towards the edge of the road. This requires a re-iteration of the generalisation process. A **rotation** method solves the second problem. If the angle between the road edge and the building outline is smaller than a certain threshold, the building outline is rotated so that the edges of the building outline and the road coincide (Swiss Society of Cartography, 1987, p.30). If this angle is exceeded the situated remains in the status quo to preserve the character of the situation. These two methods have to be applied in this order, because reversion introduces further problems in situations where straight building outlines form an angle with a curved road edge. In this regard, the sequence of road simplification method and building outline smoothening method is also important.

Parcels

In urban areas the pattern of parcels is important. Therefore, their appropriate generalised representation should be treated with care matching its importance. Comparing the Land-Line dataset with the map 1: 10 000 scale it can be derived that within residential blocks, defined as street cycles, the longest line is retained. This separates the parcels that border different streets. However, this still assumes a linear representation of the features. Transferring this rule to an areal approach to the parcel object, this rule should be rephrased. Every parcel should be assigned an attribute indicating the road it borders. Only parcels having the same road attribute can be **aggregated**. This rule also maintains the centre line separating the parcels that border different streets. Generalised road centre lines may have caused overlapping if the width of the road was too small for representation true to scale. Also displacement of the centre line may have caused overlapping. Therefore, the object class “parcel” has to have a **reclassification** method. That part of the parcel that is overlapped by the road has to be reclassified to road. This can be implemented using a clipping overlay.

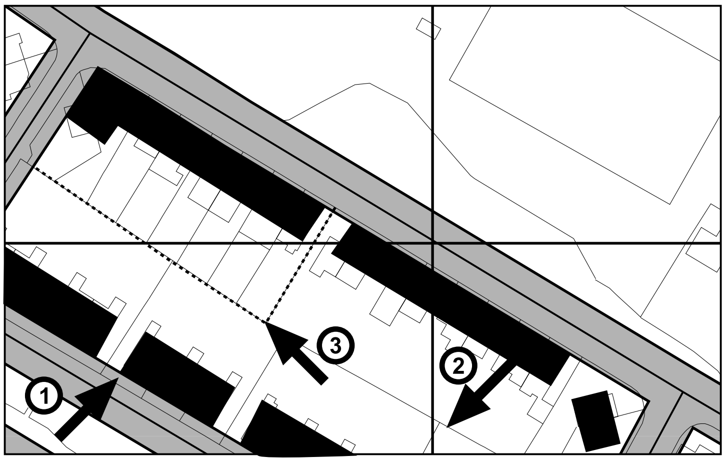

Figure 7. Overlap of road buffer (1) is reclassified. Centre line (2) of parcels is maintained. Parcels with common edges intersecting with buildings are aggregated (3).

These methods still have not changed the number of represented parcels. Comparing the Land-Line dataset with the 1: 10 000 scale map however, shows a reduction of the number of parcels. The aggregation method has to be refined to mimic this reduction. Only those parcels can be aggregated that satisfy the requirement that their common border intersects with a building outline. This rule can be applied in the case of semi-detached houses, though a caveat should be mentioned. In case of a row of terraced houses, this rule results in a generalised situation, where the number of parcels will then be reduced to one. Every common border coincides with a building outline, thus applies for aggregation. The other extreme is a situation where only detached houses or villas are present. The number of parcels is not reduced at all for none of the parcel boundaries coincides with a building outline. Taking a closer look to both maps, it shows that the parcels situated on a street corner are not aggregated with other parcels in order to maintain the character of the situation. Introducing yet another requirement for aggregation can simulate this. Parcels can only be aggregated if the building outlines on the parcels have similar orientation. This has already been checked for the aggregation of building outlines themselves, so this does not introduce further problems.

Methods: formulation, formalisation and sequence

The functionality of the spatial information system in the realm of generalisation is supplied by the formulation of methods that can be applied to objects stored in an object-oriented database. The formulation of these methods in the previous section has been derived from the literature review and from a comparison of the contents of the Land-Line dataset and the 1: 10 000 scale map.

As proposed in the second chapter, these methods are based on rules stored in a rule base. The methods now have to be translated into several rules. In order to implement a system capable to perform dynamic generalisation, these rules have to be formalised. However, these description of these rules is sometimes under-specified, sometimes over-specified (Robinson, 1997, p.240). Furthermore there often is considerable latitude in the interpretation of the rules. Nevertheless, Appendix C gives the formulation of these rules for the three object classes in a semi-natural language. The methods referring to (spatial) relationships between objects are not described.

Not only the rules themselves pose problems, for they do not operate in a vacuum. Their context, in this case the presence of other rules, have to be taken into account. Already in the previous section the sequence of rules has been referred to many times. Also this sequence has to be formalised. If this were not the case, the outcome of the generalisation process is different a lot from the intended sequence as pursued in this exercise. This formalised sequence is called the procedural knowledge.

Chapter 4

Conclusion and recommendations

Due to the increasing complexity of modern society there is a growing need for information about the natural and manmade environment: geographical information. Sharing this information is very important. The presence of the networked environments, as the Internet, facilitates this sharing. To support the various levels of decision making, different levels of detail of geographical information is needed. Sometimes integration of these levels is necessary. Generalisation procedures facilitate this integration. Furthermore, the medium of the Internet requires generalisation procedures to cope with the limited amount of map space.

The approach of dynamic generalisation is taken up to facilitate the generalisation of webmaps of urban area. The approach requires an object-oriented approach. Therefore, objects and their methods have been designed based on a review of literature and a comparison of the original Land-Line dataset and the 1: 10 000 scale map of Ordnance Survey. This structure has been applied to different parts of the Land-Line dataset to visualise the need and the operation of the methods.

During the implementation phase of the structure several problems were encountered. First, the Land-Line dataset is structured in layers, not as objects. Furthermore, the structuring of the layers was not fit for use in a spatial information system. There still remains the legacy of the analogue era. The way the features look is important, not their structuring. This is especially the case with the layer of general line detail. This has no other application than the representation of various lines. These lines may represent parts of buildings, parcels, smaller manmade constructions etc. It is therefore recommended to re-structure this layer by separating the different information contents into different layers, though still doing this pursues no object-orientation. Especially with the growing need for the dissemination of maps across the Internet, this approach has to be implemented into the Land-Line datasets. For example, utility companies can use these large-scale maps on mobile laptops during inspection in the field.

Still, the rule-based approach combined with object-orientation is no panacea for generalisation. This procedure always will require some input from cartographers. For most applications however, the results of the current approach are satisfactory.

Literature

Barr, R. (1998c) Dangling Segment – Stop the world, I want to get off. GIS Europe 7 (9), p.14-15

Brazile, F & A. Edwardes (1999) Organising Map Space for Generalization through Object-primary Partitioning. <http://www.geo.unizh.ch/~fbrazile/tmp/partitions.pdf>

Brown, A (1987) Generalisation for topographic maps (II) – a reader. ITC, internal document.

Brown, A. (1993) Map design for screen displays. The Cartographic Journal 30 (2), p.129-135

Brown, A. (1998) GFM2 module 15 – topographic cartography, exercise: generalisation for topographic maps of an urban areas. ITC, internal document

Buttenfield, B.P. & R.B. McMaster (1991) Map Generalization. Making rules for knowledge representation. Harlow: Longman Group.

Elzakker, C.P.J.M. & A. Koussoulakou (1997) Kaarten en het gebruik ervan op het Internet. Kartografisch Tijdschrift XXIII (4), p.5-12.

Hardy, P.G. (1999) Active Object Techniques for Production of Multiple Map and Geodata Products from a Spatial Database. <http://www.laser-scan.compapers/ica99gwg/pghica99gwg.htm>

Harley, J.B. (1975), Extracts from: Ordnance Survey Maps – a descriptive manual. Southampton: Ordnance Survey. In: Brown, A. (ed.) Generalisation for topographic maps (II) – a reader. ITC, internal document.

Harrower, M., C. P. Keller & D. Hocking (1996) Cartography on the Internet: thoughts and a preliminary user survey. <http://maps.unomaha.edu/NACIS/CP26/Thoughts.pdf>

Højolt, .P (1999) From 1: 10 000 to 1: 50 000 at 300 km^2 pr <http://www.geo.unizh.ch/~weibel/ottawa99_WS/hojholt/300km2u.pdf>

Jones, C.B. (1995) Geographical Information Systems and Computer Cartography. Harlow: Addison Wesley Longman Ltd.

Lamy, S. & A. Ruas (1999) The Application of Agents in Automated Map Generalisation. Proceedings 19th International Cartographic Conference, Ottawa, August 14 to 21, 1999. Ottawa: Canadian Institute of Geomatics. CD-ROM.

Lee, D. (1999) Practical Solutions for Specific Generalization Tasks. <http://www.geo.unizh.ch/~weibel/ottawa99_WS/dlee/DanLee-ICA-workshop99.pdf>

Lonergan, M.E. & C.B. Jones (1999) The deformation and displacement of areal objects during automated map generalisation: an approach to conflict resolution through maximising nearest neighbour distances. <htp://www.comp.glam.ac.uk/pages/staff/mlonerga/workshoppaper.htm>

Mackaness, W.A. & R. Purves (1999) Issues and Solutions to Displacement in Map Generalisation. Proceedings 19th International Cartographic Conference, Ottawa, August 14 to 21, 1999. Ottawa: Canadian Institute of Geomatics. CD-ROM.

Müller, J.C., J.P. Lagrange & R. Weibel (1995), GIS and Generalization. London: Taylor & Francis Ltd.

Müller, J.C., R. Weibel, J.P. Lagrange & F. Salgé (1995) Generalization: State of the art and issues. In: Müller, J.C., J.P. Lagrange & R. Weibel (eds.), GIS and Generalization, p.1-17. London: Taylor & Francis Ltd.

Plewe, B. (1998) GIS Online: information retrieval, mapping, and the Internet. Santa Fe: Onword Press

Robinson, G.J. (1995) A hierarchical top-down bottom-up approach to topographic map generalization. In: Müller, J.C., J.P. Lagrange & R. Weibel (eds.), GIS and Generalization, p.235-245. London: Taylor & Francis Ltd.

Roche, C. (1998) Cartography and the Internet. In: Kuunders, R.E. (ed.) Kartografie en Internet. NVK publicatiereeks, 24. Enschede: ITC, p.4-6

Ruas, A. (1999) The role of meso level for urban generalisation. <http://www.geo.unizh.ch/~weibel/ottawa99_WS/ruas/aci-ws-99-ruas.htm>

Shea, K.S. (1991) Design considerations for an artificially intelligent system. In: Buttenfield, B.P. & R.B. McMaster (eds.) Map Generalization. Making rules for knowledge representation, p.3-20. Harlow: Longman Group.

Spiess, E. (1995) The need for generalization in a GIS environment. In: Müller, J.C., J.P. Lagrange & R. Weibel (eds.), GIS and Generalization, p.31-46. London: Taylor & Francis Ltd.

Stynes, K., J.Wood, J.Dykes, P. Fisher & D. Uwin (1998) Publishing Cartography on the Web. <http://www.geog.le.ac.uk/argus/ICA/K.Stynes/>

Swiss Cartographic Society (1987) Cartographic Generalisation – Topographic Maps. Cartographic Publication Series 2.2nd Edition

Thompson, M. M. (1979) Extracts from: Maps for America. U.S. Geological Survey. 1st edition. In: Brown, A. (ed.) Generalisation for topographic maps (II) – a reader. ITC, internal document.

Appendices

Appendix A: The Land-Line data set

Description

Land-Line offers a comprehensive dataset depicting man-made and natural features ranging from houses and factories, roads and rivers, to marshland and administrative boundaries. There are almost 229,000 tiles covering the whole of Great Britain, surveyed and digitised at three different scales according to location:

- Urban Land-Line: 1:1 250 scale mapping Larger towns, cities.
- Rural Land-Line: 1:2 500 scale mapping Cultivated country, small towns.
- Moorland Land-Line: 1:10 000 scale mapping Moors, mountains, estuaries, etc.

Land-Line has many applications. With the use of appropriate software, Land-Line can be viewed on screen, manipulated, merged with other data and then plotted onto paper or film. Typical applications include Geographical Information Systems (GIS); estate, land and asset management; environmental maintenance; reference and research; CAD; and project planning.

The data for this exercise are available as a sample set on the Internet website of the Ordnance Survey®, the British national mapping organisation. Ordnance Survey® sample data is available to customers who wish to evaluate data before committing themselves to an order. It is also available for system suppliers to help them with software development and testing. The dataset contains vector data of an urban area (Port Talbot/Aberafan) in the south of Wales, 500m x 500m in extent, in DXF format, designed for output at scale 1:1 250.

Visual impression

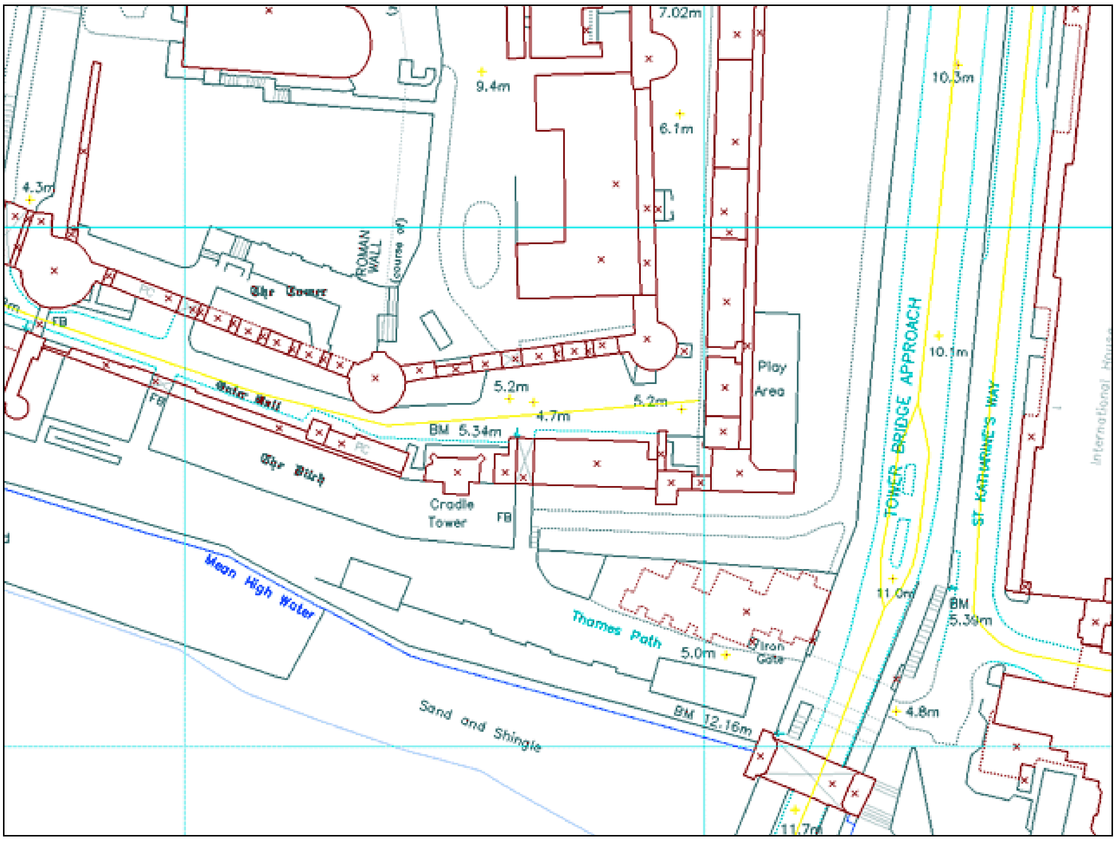

Figure 8 A plot of a Land-Line® dataset

Content *

Land-Line represents the built environment in clear and concise detail using 37 feature codes in NTF and 43 layers in DXF. Land-Line layer names conform to the British Standard BS 1192 Part 5 for Layer Names in DXF. The layer names are 8 characters long, in the range G8010001 to G8011212.

description	layer	type	line	block	colour
building outline	**G8010001**	**POLYLINE**	**continuous**		**red**
building pecks	**G8010004**	**POLYLINE**	**dashed**		**red**
parish boundary	G8010007	POLYLINE	continuous		magenta
district/London b bdy	G8010008	POLYLINE	continuous		magenta
county boundary	G8010009	POLYLINE	continuous		magenta
electoral boundary	G8010010	POLYLINE	continuous		magenta
boundary post/stone	G8010011	INSERT	continuous	BDYPOST	magenta
bdy mereing symbol	G8010013	INSERT	continuous	BDYMEREI	magenta
narrow gauge railway	G8010014	POLYLINE	continuous		blue
railway track	G8010015	POLYLINE	continuous		blue
pecked road metalling	**G8010021**	**POLYLINE**	**dashed**		**blue**
trig point	G8010025	INSERT	continuous	TRIGP	blue
bench mark	**G8010026**	**INSERT**	**continuous**	**BENCHM**	**blue**
spot height	**G8010027**	**INSERT**	**continuous**	**SPOTH**	**blue**
general line detail	**G8010030**	**POLYLINE**	**continuous**		**white**
general alignments	**G8010032**	**POLYLINE**	**dashed**		**white**
underground alignments	G8010033	POLYLINE	dashed(x2)		blue
veg/landform limit	**G8010036**	**POLYLINE**	**dashed**		**green**
overhead alignments	G8010043	POLYLINE	dashed(x2)		blue
pylon	G8010049	INSERT	continuous	PYLON	blue
minor line detail	**G8010052**	**POLYLINE**	**continuous**		**white**
point feature	**G8010057**	**INSERT**	**continuous**	**POINT**	**white**
water detail	**G8010059**	**POLYLINE**	**continuous**		**cyan**
flow arrow	**G8010069**	**INSERT**	**continuous**	**FLARROW**	**cyan**
mean high water	G8010071	POLYLINE	continuous		cyan
mean low water	G8010072	POLYLINE	continuous		cyan
parliamentary boundary	G8010079	POLYLINE	continuous		magenta
centre line public road	**G8010098**	**POLYLINE**	**continuous**		**yellow**
general building seed	**G8010321**	**INSERT**	**continuous**	**BUILDSED**	**red**
glasshouse seed	G8010323	INSERT	continuous	GLASSEED	white
copyright symbol	**G8010570**	**INSERT**	**continuous**	**COPY**	**white**
header (map footnotes)	G8010571	TEXT	standard		white
grid lines and values	**G8010572**	**LINE/TEXT**	**dot**		**white**
neatline and corner values	**G8010573**	**LINE/TEXT**	**standard/ continuous**		**white**

description	layer	type	line	block	colour
not used	G8010574	TEXT	continuous		white
not used	G8010575	TEXT	continuous		white
road names and numbers	**G8011000**	**TEXT**	**continuous**		**blue**
boundary text	**G8011005**	**TEXT**	**standard**		**magenta**
building numbers and names	**G8011006**	**TEXT**	**standard**		**red**
miscellaneous text	**G8011009**	**TEXT**	**standard**		**white**
water feature text	**G8011010**	**TEXT**	**standard**		**cyan**
land parcel numbers	G8011013	TEXT	standard		green
vegetation/landform limit (suppressed)	G8010035	POLYLINE	dashed		green
positioned coniferous tree-symbol	G8010372	INSERT	continuous	TREEC	green
positioned non-coniferous tree-symbol	G8010373	INSERT	continuous	TREENC	green
top of slope	G8010374	POLYLINE	continuous		red
top of cliff	G8010375	POLYLINE	continuous		red
bottom of slope/cliff	G8010376	POLYLINE	dashed		brown(34)
boulders-symbol	G8010377	INSERT	continuous	BOULDERS	brown(34)
boulders(scattered)-symbol	G8010378	INSERT	continuous	BOULDSC	brown
coniferous trees-symbol	G8010379	INSERT	continuous	TREESC	green
coniferous trees (scattered)-symbol	G8010380	INSERT	continuous	TREECSC	green
coppice/osier-symbol	G8010381	INSERT	continuous	COPPICE	green
marsh/saltmarsh/reeds-symbol	G8010382	INSERT	continuous	MARSH	green
non-coniferous trees-symbol	G8010384	INSERT	continuous	TREESNC	green
scattered non-coniferous trees –symbol	G8010385	INSERT	continuous	TREENCSC	green
orchard –symbol	G8010386	INSERT	continuous	ORCHARD	green
heath –symbol	G8010387	INSERT	continuous	HEATH	green
rock –symbol	G8010388	INSERT	continuous	ROCKS	brown(34)
rock(scattered)-symbol	G8010389	INSERT	continuous	ROCKSCAT	brown(34)
rough grassland-symbol	G8010390	INSERT	continuous	ROUGHGR	green
scrub –symbol	G8010392	INSERT	continuous	SCRUB	green
upper level of through communication –symbol	G8010395	INSERT	continuous	ULC	magenta
cliff –symbol	G8010396	INSERT	continuous	CLIFF	brown(34)
slope –symbol	G8010397	INSERT	continuous	SLOPE	brown(34)
water –symbol	G8010400	INSERT	continuous	WATER	cyan
scree	G8011210	INSERT	continuous	SCREE	brown(34)
positioned boulder	G8011211	INSERT	continuous	BOULDER	brown(34)
ridge/rockline	G8011212	POLYLINE	continuous		red

* The dataset of the sample area contains only the layers in bold face

Appendix B: Object cards

ORGANISATION	…	
OBJECT	NAME	**ROAD CENTRE LINE**
	DESCRIPTION	centre line indicating the course of a road
	SYNONYMS	?
	RULES	length > 0 m
	DATA TYPE	line
	ID_ATTRIBUTE(S)	ROAD_ID
ATTRIBUTES	NAME	ROAD_ID
	DATA TYPE	number
	CONSTRAINTS	compulsory ,unique
	NAME	WIDTH
	DATA TYPE	meters
	CONSTRAINTS	positive
RELATED OBJECTS + relation builders	ROAD CENTE LINE– connects PARCEL – borders	
DYNAMICS	remove, buffer, exaggerate, displace, simplify	

ORGANISATION	…	
OBJECT	NAME	**BUILDOUT**
	DESCRIPTION	perimeter of a man-made construction
	SYNONYMS	line demarcating the areal extent of dwelling
	RULES	AREA > 0 m^2
	DATA TYPE	area
	ID_ATTRIBUTE(S)	BUILD_ID
ATTRIBUTES	NAME	BUILD_ID
	DATA TYPE	number
	CONSTRAINTS	compulsory ,unique
	NAME	AREA
	DATA TYPE	square meters
	CONSTRAINTS	positive
	NAME	class
	DATA TYPE	binary classification
	CONSTRAINTS	0 = general, 1 = special
RELATED OBJECTS + relation builders	PARCEL – contain BUILDOUT – neighbourhood	
DYNAMICS/METHODS	graphical selection, aggregate, displace, exaggerate, simplify, smoothen, rotate	

ORGANISATION	…	
OBJECT	NAME	**PARCEL**
	DESCRIPTION	a piece of land on which a subject has a right of legal ownership
	SYNONYMS	plot
	RULES	AREA > 0 m^2
	DATA TYPE	area
	ID_ATTRIBUTE(S)	PARCEL_ID
ATTRIBUTES	NAME	PARCEL_ID
	DATA TYPE	number
	CONSTRAINTS	compulsory ,unique
	NAME	AREA
	DATA TYPE	square meters
	CONSTRAINTS	positive
RELATED OBJECTS + relation builders	BUILDOUT – contain ROAD CENTRE LINE– borders PARCEL – borders	
DYNAMICS	aggregate, reclassify	

Appendix C: Rule base for methods

Road

- SIMPLIFY

IF there are minor irregularities THEN
simplify the road using a Douglas-algorithm
END IF

- BUFFER/EXAGGERATE

IF the road edges are 2x meters apart, measured perpendicular to the road centre line THEN
buffer the centre line with a threshold of x meter AND x ≥ 6 meter.
ELSEIF the road edges are 2x meters apart, measured perpendicular to the road centre line AND x < 6 meters THEN
exaggerate the road by buffering the centre line using a threshold of 6 meter
END IF

- DISPLACE

IF the buffers overlap AND the number of parallel roads is even THEN
displace both roads until there is no more overlap AND the sum of the displacement vectors is zero AND the multiplication of the displacement vector and the vector indicating the course of the centre line at the point both vectors meet is zero.
ELSEIF the buffers overlap AND the number of parallel road is odd THEN
displace all road centre lines except for the centre line, observing the same rules for the displacement vectors as set above.
END IF

Building

AGGREGATE/DISPLACE GENERAL BUILDING

IF a general building is within a certain distance of another general building AND these buildings are within the same block of buildings AND their main orientation is not too different AND the difference in distance to the road for each of them is smaller than a certain length THEN
    aggregate these buildings
ELSEIF a general building is within a certain distance of a landmark building AND these buildings are within the same block of buildings THEN
    displace the general building
END IF

AGGREGATE/DISPLACE LANDMARK BUILDING

IF a landmark building is within a certain distance of another special building AND they share a common edge THEN
    aggregate these buildings
ELSEIF a landmark building is within a certain distance of another special building AND they do not share a common boundary AND these buildings are within the same block of buildings THEN
    displace the buildings using the “happiness-algorithm”
END IF

SIMPLIFY

IF the scale of representation is smaller than 1: 1250 AND is a complex aggregated feature THEN
    simplify the building outline by taking its minimum bounding box, then scale it to obtain the area of the input feature and then rotate it to reflect the principle orientation of the input feature
ELSEIF the scale of representation is smaller than 1: 1250 AND is a simple aggregated feature or a single building outline THEN
    simplify the building outline using the simplification algorithms of ESRI’s ARC/INFO
END IF

DISPLACE

IF the general or landmark building outline overlaps the road centre line buffer AND share a common boundary with the road edge or are within a certain distance to the edge of the road THEN
    displace the building outline to the edge of the road centre line buffer using a displacement vector perpendicular to the course of the road centre line in the direction of the centre of the street cycle it has been assigned to.
ELSEIF IF the general or landmark building outline overlaps the road centre line buffer AND do not share a common boundary with the road edge or are not within a certain distance to the edge of the road THEN
    displace the building outline to the centre of the parcel it is situated on using a displacement vector perpendicular to the course of the road centre line
END IF

SMOOTHEN/DISPLACE

IF after displacement there still is some map space between building outline and road centre line buffer THEN
    smooth the edge of the building outline parallel to the edge of the road centre line buffer and displace the building outline towards the edge of the road centre line buffer until they share a common edge
END IF

ROTATION

IF the angle between the building outline and the road centre line buffer < 22.5° THEN
    rotate the building outline until the angle is 0°
ELSEIF the angle between the building outline and the road centre line buffer 22.5° < 45° THEN
    rotate the building outline until the angle is 45°
END IF

Parcel

AGGREGATE
IF two parcels assigned to the same road centre line share a common edge AND this edge intersects a building outline AND the building outlines on the parcels have a not too different orientation THEN
 aggregate the parcels
END IF

RECLASSIFY
IF the road centre line buffer and the parcel overlap THEN
 reclassify the overlap as being road
END IF