# ALGORITHMS FOR THE AMALGAMATION OF TOPOGRAPHIC DATA

Nicolas Regnauld
Ordnance Survey
Research and Innovation
Romsey Road
Southampton SO16 4GU
United Kingdom
Email: nregnauld@ordsvy.gov.uk

*Abstract*
*This document proposes a set of algorithms that need to be developed to cover the amalgamation operation for generalising MasterMap data. The aim is to be able to generalise MasterMap at any scale and complying with any user requirements. Of course the set of algorithms probably won't achieve that completely, but we believe that it should provide a reasonable alternative for most of the requirements.*

*Keywords*
*Generalisation algorithms, Amalgamation, Delaunay Triangulation, Topographic maps*

## Introduction

Amalgamation is one of the main operations used in map generalisation. Its generic behaviour can be defined as the replacement of several features by a single one. Depending on the scale reduction targeted and the specifications of the product, the features that require amalgamation varies. It affects the types of features that can be amalgamated, the distance threshold under which features should be amalgamated, the types of features that would prevent the amalgamation of others if they lie in the gap, etc.

Previous researches on amalgamation algorithms mainly followed two lines. One is the use of morphologic operators (expand then shrink) (see [shylberg 93], [ Su et al 97], [Cámara and Francisco 2000]). They are particularly adapted to raster data. The other trend is to use a Delaunay triangulation to provide explicit spatial relationships between features and guide the amalgamation process [Jones et al 95], [Peter and Weibel 99]. These techniques are used in the context of vector data.

In this paper, we study the requirements for algorithms to support amalgamation of Ordnance Survey (OS) MasterMap features in order to generate topographic maps from the OS product range. MasterMap is the new Ordnance Survey's detailed database. It contains vector data, which gives an initial preference for the triangulation based techniques. In addition, the triangulation will be necessary for other generalisation operations (especially displacement), so a common structure for spatial analysis would be convenient. Finally, the morphologic operators are very hard to fully control. They generate global solution with no easy local tuning. For all

these reasons, the morphologic operators are not addressed further in this paper. We think they are not flexible enough for our purpose.
The first section of the paper describe the role of amalgamation in conflict resolution, compared with other operations. The second section presents the main types of features that require amalgamation for topographic maps. The thirds section is a brief discussion about when the grouping of features that can be amalgamated should be performed in the generalisation process. Section 4 a presents a general classification of amalgamation techniques, based on observations of OS data and existing maps. Section 5 contains the specifications for the amalgamation algorithms required for generalising OS MasterMap data.

# 1 Amalgamation among other generalisation operators

Amalgamation can refer to two main actions. It can correspond to the fusion of two adjacent polygons usually due to their reclassification to a single theme. Or it can correspond to the aggregation of several polygons which are initially not touching. In this report, we focus more on the second problem, which takes part of the more general problem of conflict resolution.

- Why do we amalgamate?

Amalgamation is mainly used to reduce proximity conflicts. A proximity conflict occurs when gaps between features are to small to be perceived. Amalgamation solves this by eliminating these gaps, therefore merging the features together. Amalgamation is then used in competition with displacement, which takes the reverse action: increase the gap between features until they can be visually distinguished. Shrinking is also a possibility to increase the gap between features. To answer the question, we use amalgamation when features are too close, and we can't or don't want to increase the gap between them.

- Why is this gap suddenly two small?

The first reason why the gaps between features appear too small, is the scale reduction. Scale reduction by a simple zoom out reduces everything: the features and the gaps between them. So some of the gaps are going to get smaller than what the human eye can perceive.
In addition, scale reduction can be an indirect cause for the gaps getting too small. Scale reduction may necessitate to apply other algorithms (like enlargement and displacement), that would fill the gaps between objects.

Figure 1shows the main generalisation operations involved to solve the size and proximity conflicts. Other operations, like shape simplification, may have marginal effects on them, but they are not used to solve these conflicts. The plain green arrows show that an operation can be used to reduce or solve a particular type of conflict. The dotted green arrow means the same, except that the operation may not be powerful enough to solve the conflict, i.e. the amalgamation of two very small objects can generate a polygon which is still too small or too narrow. Red arrow show frequent (expected) bad effects on an operation. For example, moving a feature away from feature A can make it closer to Feature B, creating a new conflict.

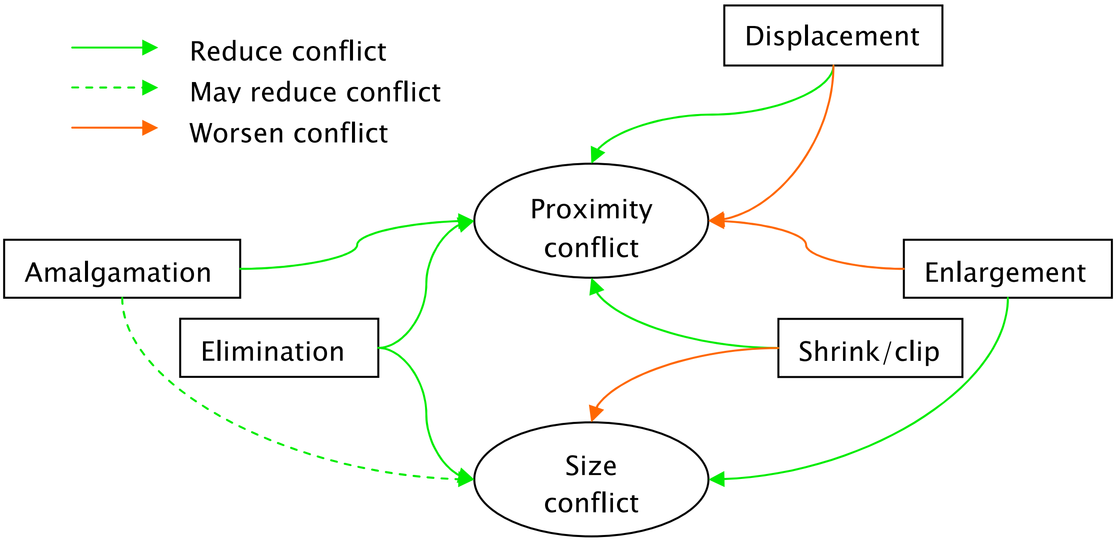
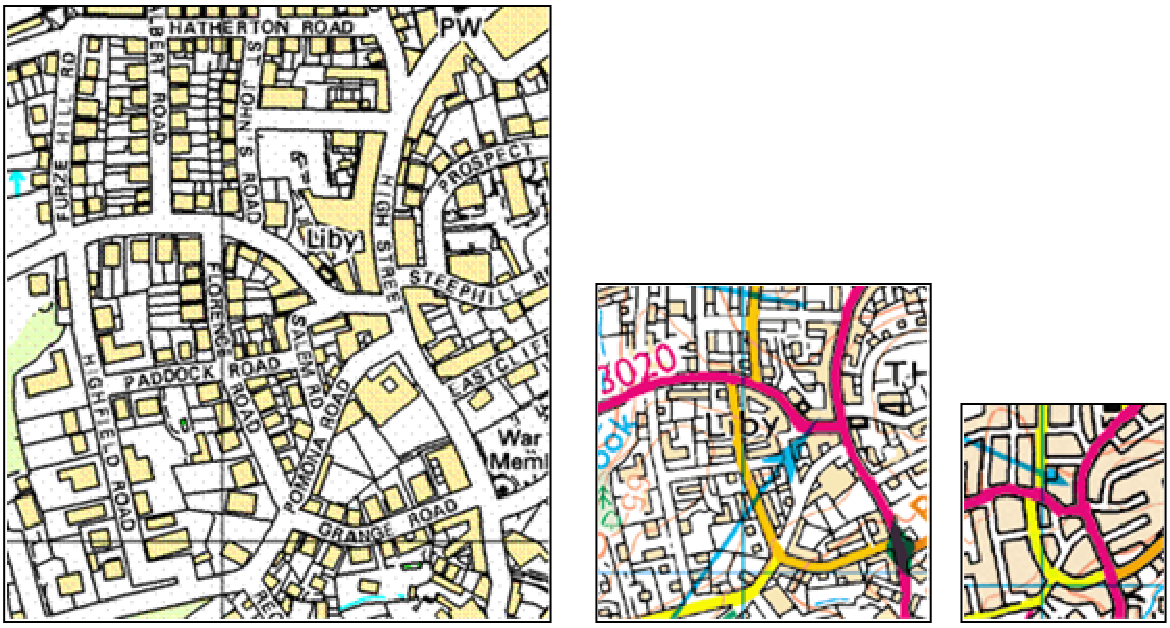



**Figure 1: Operations and conflicts**

The sad thing is that we often prefer the operations that have bad side effects than the other ones. Amalgamation and elimination are modifications which are immediately noticeable on a map. Enlargement, shrink or displacement, when used wisely are much more difficult to spot, and therefore preferred.
This means that we use amalgamation only when the shrink and displacement are not applicable, or can not solve the proximity conflict. Amalgamation is particularly useful when it solves two conflicts at once: group two close small features together to solve the size and proximity conflict.

## 2 Amalgamating what?

Amalgamation is used to unify several area features. This operation is rarely needed at large scale, where distinct features tend to be kept individually. It starts to be very useful at medium scales (1:25 000 to 1:100 000). At these scales, the main two dimensional features shown in topographic maps are buildings, water bodies and vegetation (mainly forest). We will restrict our study to these three types of features.
Buildings are very different to the two other ones. They are mainly small features, with distinctive shapes (orthogonal), and often grouped in dense clusters. On the other hand, forests or lakes/ponds, are often large, and distant from each other. This means that amalgamation is required for buildings at larger scales than it is needed for forests and water.
The 3 figures (2, 3 and 4) below show examples of maps extracts, taken from Ordnance Survey maps series (10k, 25k, and 50k). They are not shown to scale, but in a way where we can compare easily the evolution of the features through the scale reduction. For buildings (Figure 2), the amalgamation is quite dramatic. At 1:25k, few buildings have kept their individual character. At 1:50k, buildings are only shown by large amalgams.

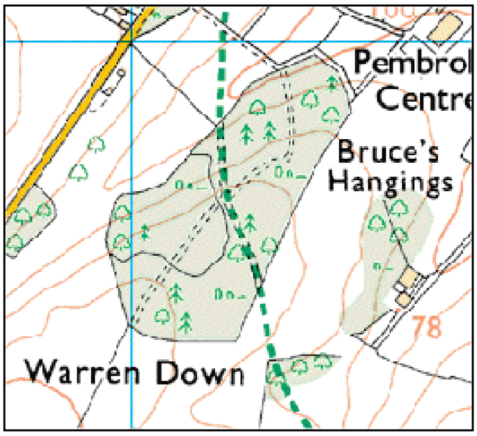



a) 10k b) 25k c) 50k

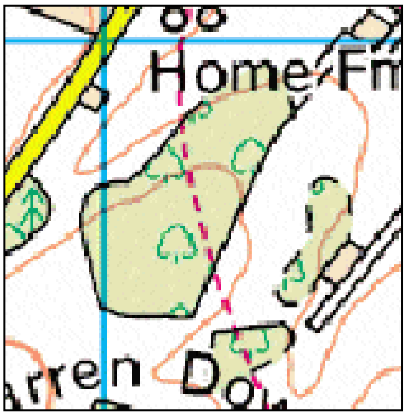
Figure 2: Building amalgamation

For the forests, Figure 3 shows that the extent of the features varies very little between MasterMap and the 25k Map series (MasterMap is not a map, but the main large scale database of Ordnance Survey). The main difference seems to be in the classification of the polygons. Comparing MasterMap and the 50k series shows that the smallest features have been omitted, some small ones have been enlarged and/or displaced, but no amalgamation took place (on the feature at the bottom of the picture, displacement was preferred). The main difference is that the adjacent features with similar types of vegetation have been merged.

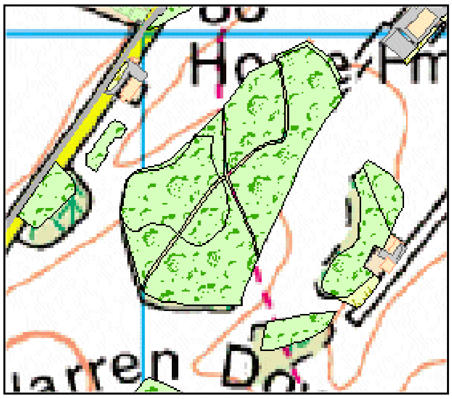
a) MasterMap




b) 25k series




c) 50k series



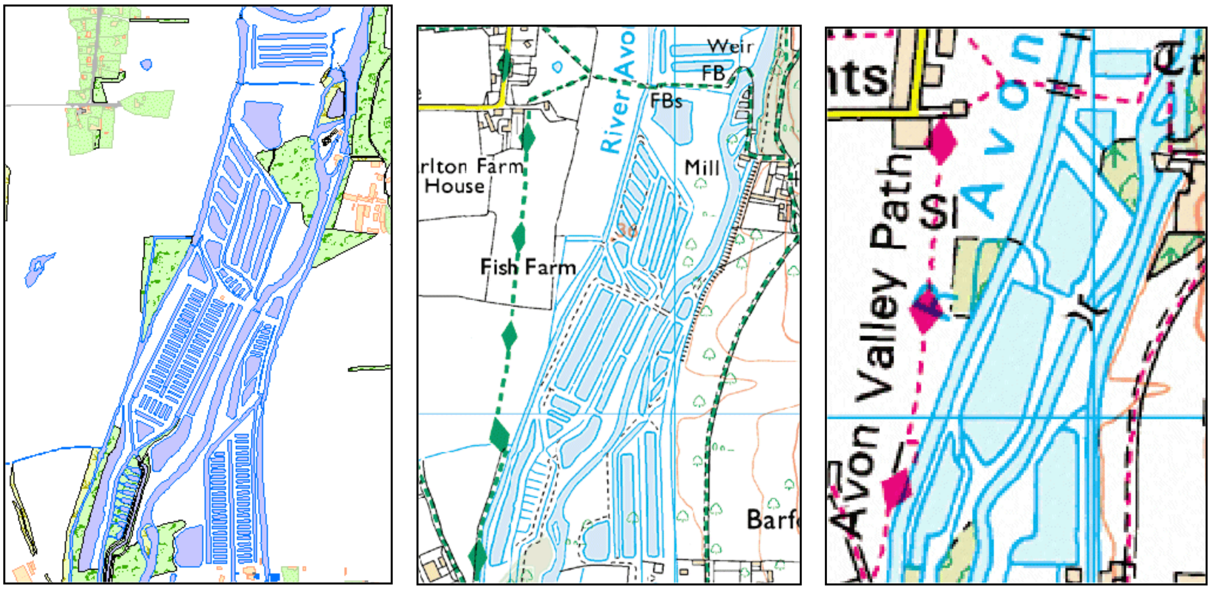
d) MasterMap + 50kseries

**Figure 3: Forest amalgamation**

We have distinguished water bodies from other natural land coverage because they can have particular functions which influence the way they should be generalised, and especially the way they should be amalgamated. In MasterMap, most of the water bodies are represented by polygons. Some of them represent flowing water or drainage bodies. In this case, from a functional point of view, they should be seen as linear components of an hydrologic network. Amalgamation in this case should be limited to features that are contiguous in the network sense. Then it creates a polygon that is still functionally linear. If you amalgamate features that are not contiguous in the network, then you produce a polygon that has no linear function, and can be seen as an inconsistency in the network. Other water bodies that have a static function (water storage), are truly two dimensional objects, that can be amalgamated together.

The first step before amalgamating water bodies is to classify them as functionally linear or area, and build an hydrologic network. Research has been done already on that issue [Regnauld et al].

Figure 4 shows the representation of a fish farm in MasterMap, and its depiction in the 25k and 50k OS Map series. The basins are amalgamated together but never with the river itself, whatever the proximity might be. The polygons part of the river play the role of obstructing features in the constitution of clusters of basins.




a) MasterMap b) 25k c) 50k

Figure 4: Water bodies amalgamation

## 3 Main types of amalgamation

The classification below is an attempt to identify different families of amalgamation algorithms. They have been deduced from observation of OS large and medium scales maps and data. Therefore, this classification needs to be extended to cover small scales and other mapping styles (from other map producers).

### 3.1 Amalgamation with displacement

This type of amalgamation involves the displacement of the features toward each other and then amalgamate them. The displacement can be shared by the features, or limited to one of them. Usually, the bigger the feature the more resistant to displacement, so we tend to displace the smallest feature. The displacement can be full or partial. We call a full displacement a displacement that leads to the features touching or overlapping. In this case a merging is enough to complete the amalgamation. If the displacement only makes them closer, then it must be followed by an amalgamation of the other type (by bridging the features, see next section). The amalgamation by displacement does not increase the area occupied by the theme of the amalgamated features, or at least less than other methods. It also frees some space on the side of the displaced object which is opposite to the displacement direction. The main drawback of this method is that it changes the position of the features, which affect the spatial relationships with the neighbouring features. Note that moving large features is not wise, it can create conflicts in regions outside the amalgamation area.

## 3.2 Amalgamation by bridging features

It can also be necessary to amalgamate two (or more) features without having to displace them. This is particularly useful when the features are large, or resistant to displacement for any other reason. In such a case, the amalgamation consists in bridging the features. The way the bridging is done depends on the nature of the features, it is important to preserve the shape specificity of the feature type. For example, bridging buildings could be done by adding a corridor, while two lakes should be joined using a patch with curved boundaries.
The main advantage of this method is that it preserves the position of the features. The drawback is that it increases the density of the theme, which increase the overcrowding effect.

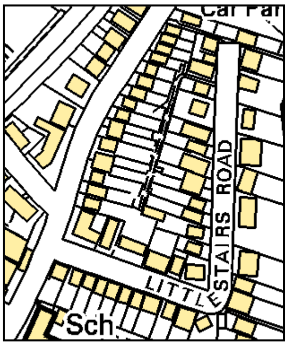
## 3.3 Amalgamation by flooding the features

This consists in replacing a group of features by a single polygon. The boundary of the resulting polygon are related to the extent of the initial group rather than the geometry of the initial features. This is useful to cope with dense areas of small objects, like buildings in urban areas.

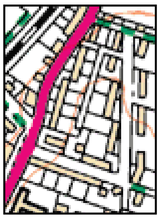
## 3.4 Amalgamation by  sampling

This consists in reducing the number of elements in a group of features. This is used when the objects are too small, and have a particular shape (nearly a symbol), that would be lost if another amalgamation technique was used. This is useful for small groups of houses for example. This is often referred in generalisation as typification [Ormsby and Mackaness 1999], [Regnauld 2001], [Sester 2003].

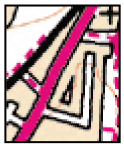
Figure 5 illustrates the different types of amalgamation. The pictures are not displayed at their true scale.




a) Initial data (10k)

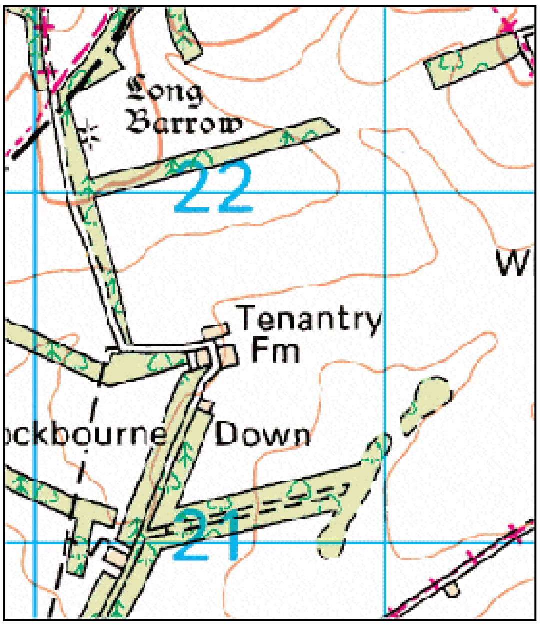


b) displacement amalgamation (25k)

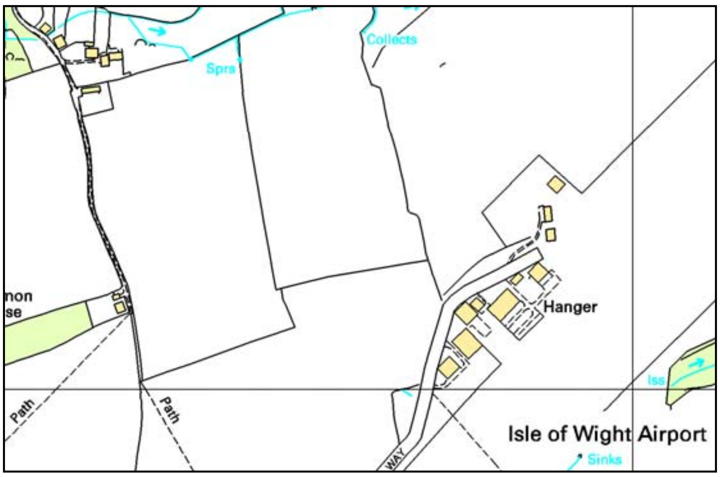


c) Flooding amalgamation (50k)

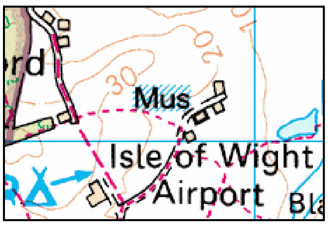
d) Initial data (MasterMap)




e) Bridging amalgamation (50k)




f) Initial data (10k)




g) Sampling amalgamation (50k)

**Figure 5: Examples of the different types of amalgamation**

## 4 Choosing the features to amalgamate

We usually amalgamate features when they are too close to each other and can't be displaced or shrank. This looks straightforward, but in fact the difficulty is to decide when to search for the objects that should be amalgamated. Looking for objects too close together in the initial state, or after enlargement, or after displacement, or after a combination of different algorithm would come up with different solutions. The practical solution consists in amalgamating whatever is too close after having applied the other algorithms. But this would probably lead to amalgamating objects that were not that close initially, which would look like a strange choice when evaluating the result.

Another practical solution would be to start with amalgamating object that are too close, with no hope of conflict resolution by other means. But this can lead to big objects which would be very hard to move later if required due to the proximity of features of other types.

The sensible solution would be to detect in the initial state the candidates for amalgamation, and amalgamate these ones later if still required.

The ideal way to perform amalgamation depends on the technique used and will be discussed more in depth when prototypes will be available.

## 5 Algorithms proposed

In this section, we specify a set of tools that would be needed to achieve amalgamation in the context of the automatic generalisation of topographic data, with the OS map styles. This set is of course not exhaustive, but should provide a good range to allow to tackle the most common amalgamation requirements.

1. Triangulation
   The triangulation is the key component for the amalgamation methods that we propose below. It helps the detection of the groups of features that need to be amalgamated, and it can also help the amalgamation itself.
   Triangulation is widely used in the generalisation context to represent the proximity relationships between features [DeLucia and Black 87], [Jones et al 95], [Ruas 98]. We need to model the triangulation in robust and flexible way, to be able to use it at different stages of the generalisation process, and for different purposes.

2. Flooding amalgamation
   We propose to try two different techniques to amalgamate features of a given type. The first one (growing tide) gradually increases the amalgam until covering all the footprints of the original features, while the second (decreasing tide) starts by filling a whole region and inserts holes where significant areas don't contain any feature. Growing tide is essential for amalgamating buildings for small scales outside the city centre. Decreasing tide will be more efficient to

amalgamate buildings in city centres, and can also be useful for amalgamating other types of features.

*Growing Tide:*

The aim of this algorithm would be to amalgamate buildings, without completely filling up the space. It would be especially useful where buildings are organised in a row alongside a road, with no direct special constraint on their opposite side (back garden). In such a case, the extent of the amalgam can be generated from the extent of the road symbol.

*Grouping criteria:*

- Distance between buildings (< threshold)

  Using a triangulation
- Prevent linking across obstructing features

*Algorithm principle:*

- Build an envelop around the buildings

  The envelop can be deduced from a triangulation
- Find the portions of roads (or other features) which are too close to the envelop, and create a buffer around them (the width of the buffer being the minimum width for a building amalgam at the target scale)
- Add to the amalgam any part of the envelop which are not covered by the buffers or other features.

*Parameters*

- The group of buildings to amalgamate
- The other features in a close neighbourhood
- The minimum width for the amalgamation symbol
- The minimum distance between features

*Decreasing Tide*

*Grouping criteria:*

- Features inside a block formed by the structuring network (roads+ rivers+railways).

*Algorithm Principle:*

- Build an envelop around the buildings of a well delimited area (city block), deduce the free spaces, see [Boffet 2001] chap. C, section 3.2.
- Fill the area with a single amalgam polygon, and insert holes where significant free spaces stand. This requires clipping to take into account the increased width of the road features.
- Other features that we want to keep should be treated like holes.

- Refinements will be needed to avoid narrow corridors in the resulting amalgam (between its external boundary and the boundary of its holes). This can be done by moving and/or shrinking (even deleting) the holes.

*Parameters*

- Features constituting the block
- The features inside the block
- List of obstructing types
- The minimum width for the amalgamation symbol
- The minimum distance between features

3. Sampling amalgamation
   This is the alternative amalgamation for the buildings at scales like the 50k. This is necessary to cover the amalgamation of buildings in low density area. We propose a light version of the typification method proposed in [Regnauld 2001]. This is a light version because it would be operated in low density areas, where clusters are obvious.

   *Grouping criteria*

   - Distance between buildings (< threshold)
   - Only buildings of similar types are grouped together (regarding the target classification)

   *Amalgamation technique*

   - Build a skeleton for the group
   - Place buildings (min size) along the skeleton, keeping the orientation consistent with the closest initial buildings.

4. Displacement amalgamation
   This algorithm is necessary to achieve the amalgamation of buildings at scales around the 25k, with the OS style. However, it is not clear yet whether the current style will be maintain or if alternative solutions will be studied, like Typification [Regnauld 2001]. Therefore we have no immediate plan to implement this type of amalgamation. For very local building amalgamation with displacement, we can use the method presented in [Lichner 79] and [Jones et al 95], which rotates and displace a small building to fit it against another nearby building.

5. Merging
   Merging adjacent features. Very simple, it just has to be generic enough to be used for different types of features. It must allow the user to specify which types can be amalgamated, and what types of obstructing features would prevent the amalgamation.

*Grouping criteria*

- Adjacency
- Compatibility between the feature types
- No obstructing feature along the frontier

*Parameters for the grouping procedure*

- Clusters of amalgamable themes
- List of obstructing themes by cluster

*Amalgamation technique*

Straightforward, most GIS will have the functionality to do a logical "OR" between several polygons.

6. Cluster amalgamation

Cluster amalgamation will be the most often used for amalgamating natural features. It has to be generic enough to cope with the different types of features that we may want to amalgamate, and the different scales. It is potentially useful at all scales, specially at small scales to create regions of woodland.

*Grouping criteria*

- Proximity
- Area of initial features (all too small)
- cumulated area of features (> threshold)
- Compatibility between the feature types

*Parameters for the grouping procedure*

- Maximum distance for amalgamating
- Minimum size for showing a feature
- Clusters of amalgamable types

*Grouping technique*

- Build a triangulation between all the amalgamable and obstructing features
- Select the edges of the triangulation that link to amalgamable objects, shorter than the minimum distance parameter. By amalgamable we mean features of amalgamable types, smaller than the minimum size parameter.
- Group the features connected by the selected edges.

*Amalgamation technique*

Here we can reuse the principles from [Jones et al 95] based on the triangulation. Every triangle linking two features of the cluster should be filled if one of the bridging edges is associated with a distance inferior to the threshold (maximum distance for amalgamation). Once this is done, the refinement using Bezier curve could be used to enhance the visual rendering. Finally, this technique may lead to holes, and possibly too narrow bridges. Holes too small should be filled, and bridges too narrow should be discarded if they don't disconnect the cluster, widened otherwise.

*Parameters for the amalgamation algorithm*

- Maximum distance for amalgamating
- Minimum width (for bridges)
- Minimum size for holes

7. Linear amalgamation

   Linear amalgamation would aim at amalgamating features which form a linear pattern. This would be useful for hedges, which are often fragmented into patches of trees, while having a strongly linear structure. It will also be very useful when we will deal with the river network, but this one has to be available first, in order to fully specify this algorithm.

## Conclusion

Having studied the requirements for amalgamation algorithms for OS topographic maps, we are now in the process of implementing them. We hope to provide some results at the workshop, at least for a number of them. The priority is given to the flooding algorithms, and the sampling one.

The results should provide valuable information to validate the range of algorithms chosen, and probably to enhance or refine it. The workshop should be a good occasion to discuss the extension or refinement of this set of algorithms to address other styles of map than the OS ones. It will also be an occasion to discuss the integration of amalgamation in the whole process of generalisation, including the way the user could specify the level of amalgamation that he wants.

## References

Boffet, A., 2001, Méthode de création d'information multi-niveaux pour la généralisation cartographique de l'urbain, PhD Thesis from the Université de Marne la Vallée.

Cámara, M.A.U., Francisco, F.J.A., 2000, "Mathematical morphology applied to raster generalisation of urban city block maps", Cartographica, 37:1, 33-48.

DeLucia, A. and Black, R.T., 1987, "A comprehensive approach to automatic feature generalisation", Proceedings of 13th International Cartographic Conference, 173-191.

Jones, C.B., Bundy, G.Ll. and Ware, J.M., 1995, "Map generalisation with a triangulated data structure", Cartography and GIS, 22:4, 317-331.

Lichtner, W., 1979, "Computer-assisted processes of cartographic generalisation in topographic maps", Geo-Processing, 1, 183-199.

Ormsby, D. and Mackaness, W., 1999, "The development of phenomenological generalisation within an object-oriented paradigm", Cartography and GIS, 26:1, 70-80.

Peter, B. and Weibel, R., 1999, "Using vector and raster-based techniques in categorical map generalisation", presented at 3rd ICA Workshop on Progress in Automated Map Generalisation.

Regnauld, N., 2001, "Contextual Building Typification in Automated Map Generalization", Algorithmica, 30:2, pp 312-333.

Regnauld, N., Mackaness W.A., Greenwood, J., "Creating a hydrologic network from its cartographic representation", to be published.

Ruas, A., 1998, "A method for building displacement in automated map generalisation", International Journal of GIS, 12:8, 789-803.

Sester, M., 2003, "Optimising approaches for generalisation and data abstraction", to appear in International Journal of GIS: Special Issue on Map Generalisation.

Su, B., Li, Z., Lodwick, G. and Müller, J-C., 1997, "Algebraic models for the aggregation of area features based upon morphological operators", International Journal of GIS, 11:3, 233-246.