Problems Arising From A Simple GIS Generalisation Algorithm

Mark R. Johnston, Christine D. Scott and Robert G. Gibb

Landcare Research, Private Bag 11-052, Palmerston North, New Zealand
Phone: +64 6 356-7154   Fax: +64 6 355-9230
{JohnstonM,ScottC,GibbR}@landcare.cri.nz

Presented at SIRC 99
The 11th Annual Colloquium of the Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
December 13–15th 1999

ABSTRACT

Practical GIS generalisation using existing Arc/Info operations invariably follows a simple, generic generalisation algorithm. Typically the algorithm consists of dissolving adjacent polygons with the same attribute value, line generalisation, selection of all polygons of small area, and elimination of selected polygons by merging with the neighbouring polygon of largest area. This paper addresses two fundamental problems with this simple algorithm: performing topologically consistent line generalisation that preserves polygon adjacencies; and establishing criteria and operators for selection of polygons as candidates for elimination. In particular, we consider the utility of the polygon *skeleton* in each of these contexts. These problems arose in attempts to reproduce a generalised Soil Map of New Zealand from higher resolution soils polygons of the New Zealand Land Resource Inventory. We have developed additional tools that operate on existing Arc/Info coverages and contribute further generalisation logic.

Keywords and Phrases. GIS Generalisation, Topologically Consistent Line Generalisation, Douglas-Peucker-Ramer Algorithm, Virtual Adjacency.

1 Introduction

Geographic information needs to be made available at multiple levels of abstraction (Müller, Weibel, Lagrange, and Salgé 1995). A map is one means of representing geographic information. The scale of a map may be defined as “the ratio between the size of an object on the map and its real size on the ground” (Müller et al. 1995). Although geographic data may be represented on maps at a range of different scales, the validity of observations or inferences drawn from such maps depends upon the appropriateness of the **thematic resolution** (the fineness of category definitions for a particular attribute) and **spatial resolution** (the smallest object represented) of the representation. Questions regarding geographic phenomena often have some inherent resolution. Models applied to these questions may be valid only over a specific range of resolutions. Moreover, the available data may not be at the required resolution, hence the need for generalisation.

Weibel and Dutton (1999) define **generalisation** as “a process which realises transitions between different models representing a portion of the real world at decreasing detail, while maximising information content with respect to a given application.” More simply, this means that generalisation is a process which transforms a spatial dataset to a coarser representation of the real world with the objective of preserving as much *valid* information at the final resolution as possible. Müller et al. (1995) clearly distinguish between *cartographic generalisation* and *GIS generalisation* (also called *model generalisation*). The former is concerned with improving legibility of graphical visualisation of spatial data, while the latter is oriented toward reduction of spatial data to a coarser thematic or spatial resolution. In this paper we are concerned with GIS generalisation in the context of a polygonal dataset representing natural resources. This is considered by Weibel (1995) as one of the three "essential building blocks" for automated generalisation together with knowledge acquisition and the post-generalisation assessment of the results.

Rijkse and Hewitt (1995) produced a **Soil Map of New Zealand** by applying the New Zealand Soil Classification (Hewitt 1998) to the spatial soils coverage held in the New Zealand Land Resource Inventory (NZLRI) (NWASCO 1979). The *soil order* classification of New Zealand soils divides soils into 15 major categories: allophanic, anthorpic, brown, gley, granular, melanic, organic, oxidic, pallic, podzols, pumice, raw, recent, semiarid, and ultic. Arc/Info line generalisation and polygon elimination GIS operations were applied to approximately 90 000 (nominally 1:50 000 scale) NZLRI polygons. This was followed by considerable editing to produce a map at a scale of 1:1 million, a twenty-fold change in scale. Two salient issues have been identified with this style of map production. First, we wish to produce a map at the 1:1 million scale that contains much less detail. Much of the detailed line work and very small areal polygon features are not valid at the final scale. Second, it is often the case that the spatial structure of a natural resource follows the topography of the land, e.g., recent soils often follow river structures but traditional area-based elimination of small polygons tends to under-represent these soils in the generalised map.

2 A Simple GIS Generalisation Algorithm

Traditional vector GIS terminology may be more rigorously defined in terms of graph theory (for definitions of basic graph theory see, e.g., Bondy and Murty (1976)). A **coverage** is a planar graph, G, embedded in the Euclidean plane $\mathbb{R}^2$, with vertex set V and edge set E. Each edge is a line segment in the planar embedding. Although G need not necessarily be connected, each component of G must be strongly connected, i.e. if any edge is removed from G there is no change in the number of components. A **polygon** is a face of a coverage and is a connected subset of $\mathbb{R}^2$ but may not necessarily have a connected boundary. The **world polygon** corresponds to the infinite face. A non-world polygon is **external** if it is adjacent to the world polygon. A non-world polygon is **isolated** if it is adjacent to exactly one other polygon (which could be the world polygon). The set of **nodes** consists of all vertices of degree ≥ 3 and one nominated vertex on the boundary of each isolated polygon. A **polyline** is a path in G and an **arc** is a polyline between nodes that contains no other nodes. Associated with a coverage is at least one **attribute** that is a function defined on the set of polygons, P.

Example. Figure 1 shows the planar embedding of a graph whose nodes are denoted by '$\bullet$' and labelled $N = \{n_1, n_2, \ldots, n_{12}\}$. Edges are denoted '—', and those vertices that are not nodes are denoted '$\circ$'. The polyline $(n_3, v_1, v_2, v_3, n_7)$ is an arc. The polygons are the set $P = \{p_{\text{world}}, p_1, p_2, \ldots, p_9\}$. The boundary of polygon p_3 is the path $(n_3, v_1, v_2, v_3, n_7, \ldots, n_8, \ldots, n_2, \ldots, n_3)$. The boundary of polygon p_7 is the set of two paths $\{(n_7, \ldots, n_{12}, \ldots, n_9, n_8, \ldots, n_7), (n_{11}, \ldots, n_{11})\}$. Note that there are three arcs with nodes n_5 and n_6. Polygons p_8, p_9 are isolated and polygons p_1, p_2, p_4, p_7, p_9 are external.

GIS generalisation of a coverage may be decomposed into two generic operations:

Line Generalisation is the appropriate representation of the *one-dimensional polyline* resolution that reduces the amount of detail within coverage arcs. Formally, line generalisation is a transformation from a source coverage S to a generalised coverage G that results in a one-to-one correspondence between the nodes, arcs and polygons, preserves the planar embedding of the nodes, and each arc in the generalised coverage approximates the corresponding arc in the source coverage, often by fewer vertices.

Polygon Generalisation is the appropriate representation of the *two-dimensional polygon* resolution. Formally, polygon generalisation is a transformation from a source coverage S to a generalised coverage G such that the polygons of G correspond to a subset of the polygons of S. New nodes and arcs may be created and old nodes and arcs removed.

A simple approach, typified by Arc/Info and the Kaleidos package of Arc/Info AML scripts by Daroussin (1991), assumes that line generalisation is independent of polygon generalisation. A difficulty with the independence of line and polygon generalisation is to identify the relationship between the line resolution (the smallest significant line detail) and the areal resolution (the smallest significant areal detail). In cartographic generalisation, significance refers to what is able to be plotted at the map scale. However, in GIS generalisation, significance refers to what degree of detail has a valid interpretation given the required resolution. This ultimately comes down to what valid inferences may be drawn from the detail of data concomitant with the questions posed. If a polygon is large, then we must rely on line generalisation, but if a polygon is small, then we must determine significance

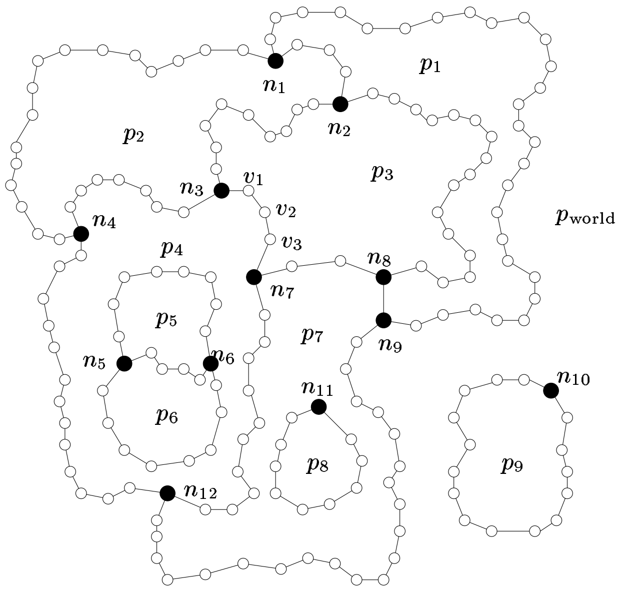


Figure 1: Graph Theoretic Definitions of GIS Terminology

via polygon generalisation. Polygon generalisation is commonly restricted to the merger of each *insignificant* polygon with its *most significant* neighbour. Hence, GIS generalisation has become synonymous with Algorithm 1 SIMPLE GIS GENERALISATION (below), which assumes that adjacent polygons of the same class are merged both before and after generalisation.

Algorithm 1 SIMPLE GIS GENERALISATION
  Perform line generalisation.
  Select insignificant polygons.
  Merge each insignificant polygon with its most significant neighbour.
end

This paper attempts to address these two component problems in the context of the SIMPLE GIS GENERALISATION heuristic: GIS line generalisation (Section 3) and GIS polygon generalisation (Section 4). Finally, Section 5 draws some conclusions and makes recommendations for future research.

3 GIS Line Generalisation

Line generalisation is one of the most studied problems in the cartographic literature (McMaster 1987a; McMaster 1987b; McMaster 1989). White (1985) suggests that the most visually effective polyline generalisation heuristic is the Douglas-Peucker-Ramer algorithm (DPR) proposed independently by Ramer (1972) and Douglas and Peucker (1973). DPR builds a new generalised polyline that lies within a predefined distance ϵ of the original polyline. Algorithm 2 DOUGLAS-PEUCKER-RAMER (see over) gives a recursive definition of DPR where for $v_i, v_j, v_k \in V$, $d(v_i, v_j, v_k)$ is the perpendicular distance between v_i and the line segment $\overline{v_j v_k}$. It is known (Bader and Weibel 1997; Saalfeld 1999) that DPR does not necessarily preserve the non-intersection property of a polyline.

We are interested in two particular aspects of line generalisation in the context of GIS generalisation: the preservation of topology during line generalisation, i.e., preventing intersection of generalised arcs; and the evaluation of the "goodness of fit" of a generalised polyline to the original.

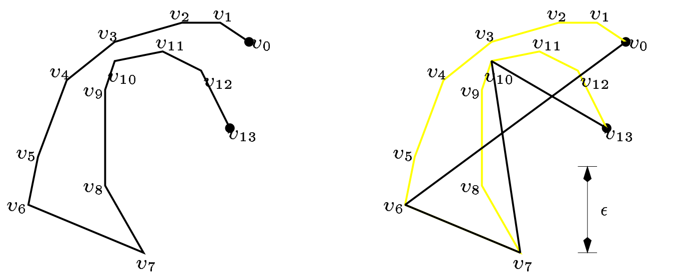
3.1 Detection and Correction of Topological Conflicts

Suppose DPR has been applied to an individual polyline and results in a generalised polyline which intersects itself. Figure 2 shows the polyline $(v_0, \ldots, v_{13})$. With an appropriate choice of ϵ, as illustrated, DPR may

Algorithm 2 DOUGLAS-PEUCKER-RAMER

Input: $L = (v_0, \ldots, v_n)$ *// a polyline*
Input: $\epsilon > 0$ *// smallest significant distance*

if ($n \geq 2$ **and** $\max_i\{d(v_i, v_0, v_n)\} > \epsilon$) **then**
  $i^* \leftarrow \text{argmax}_i\{d(v_i, v_0, v_n)\}$
  $L_\alpha \leftarrow \text{DOUGLAS-PEUCKER-RAMER}((v_0, \ldots, v_{i^*}), \epsilon)$
  $L_\beta \leftarrow \text{DOUGLAS-PEUCKER-RAMER}((v_{i^*}, \ldots, v_n), \epsilon)$
  return $\text{append}(L_\alpha, L_\beta)$
else
  return (v_0, v_n)
end
end



(a) Original Polyline

(b) Self-Intersecting Polyline resulting from DRP

Figure 2: Example of Self-Intersecting Polyline resulting from DRP

select the subset polyline $(v_0, v_6, v_7, v_{10}, v_{13})$. There are several possible methods of resolving this conflict. The obvious method is to continue to apply DRP with a smaller value of ϵ. However, this may introduce additional detail to parts of the polyline which do not currently conflict. An alternative that resolves this problem is to continue to apply DRP only to those edges of the generalised polyline which currently conflict and continue until there are no conflicts. In the worst case, it is possible that the edges of the generalised polyline that originally conflicted are refined all the way to the corresponding sections of the original polyline. The problem of deciding which offending edge to attempt to resolve first is complicated. One decision rule is to consider first those edges with the maximum number of intersections. However, Figure 3 illustrates a problem with this correction method since applying DRP to the polylines on the left and right may both result in the topologically incorrect polyline in the centre. If DRP results in the centre polyline, then we cannot easily determine whether to first consider the edge with two intersections or one of the edges with one intersection.

Now suppose that an arc within a coverage is to be generalised so that it does not intersect any other coverage arcs. A generalised polyline resulting from applying DPR may be displaced by a distance as large as ϵ from the original polyline. Hence, the generalised polyline may lose its proper topological relationship to other point or linear features that lie at a distance less then ϵ from the original polyline. Saalfeld (1999) observes that since DPR only selects a subset of the original polyline, potential conflicts can only occur with vertices of other polylines that lie within the closed convex hull of the original polyline. Hence, the search for potential topological conflicts of a given arc may be reduced to only those vertices on other arcs which satisfy these distance and

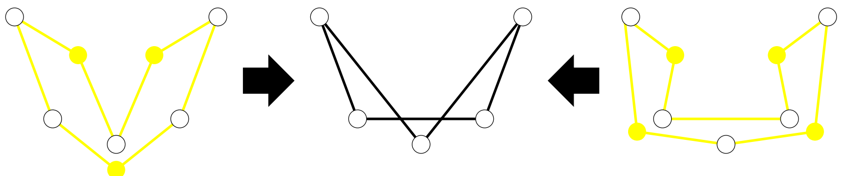
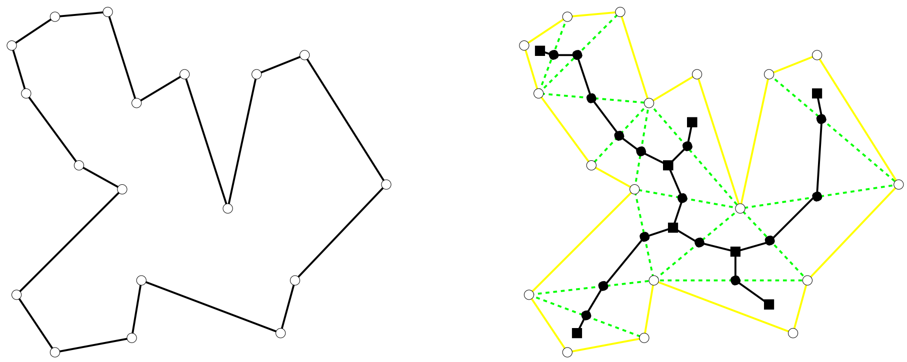


Figure 3: Problem with Correction of Self-Intersection



(a) Polygon (b) Constrained Delaunay Triangulation and Skeleton

Figure 4: Constrained Delaunay Triangulation and Skeleton of a Polygon

convex hull conditions. To correct such conflicts, Saalfeld suggests that if we choose an ϵ slightly less than the minimum distance of an offending vertex from the polyline, then we remove all possibility of any other feature conflicting with the feature undergoing generalisation. The choice of which offending edge to tackle first is not important since each offending edge may be considered independently. Although this method of Saalfeld guarantees topological consistency of a coverage, it applies only when arcs are generalised sequentially. Ideally, we wish to generalise all arcs in parallel.

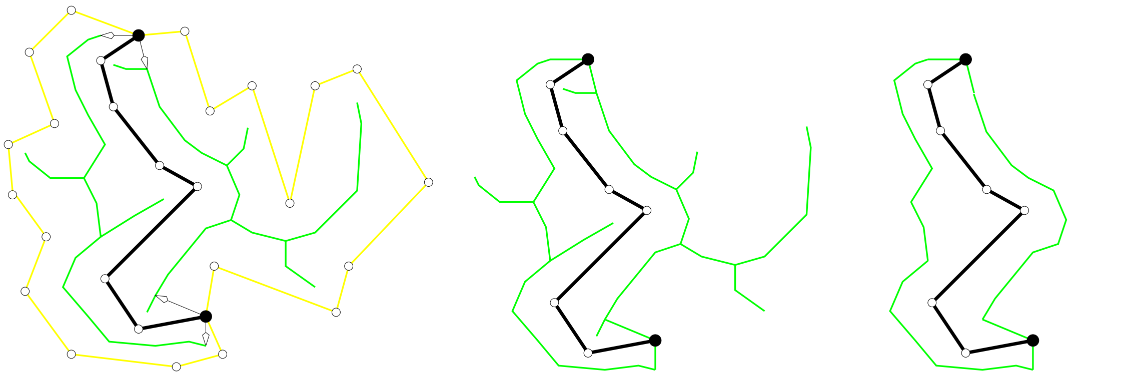
3.2 Polygon Skeletons and Constrained Line Generalisation

Within the context of GIS generalisation, the major inconvenience of DPR is that it operates exclusively in ignorance of other arcs in the same GIS coverage. Hence, it may produce generalised arcs that do not retain the topological integrity of the coverage. Although Saalfeld's scheme tells us which vertices or edges of other arcs may potentially conflict, it tells us nothing about preserving the characteristics of the *polygons* on either side. We wish to balance the degree of generalisation with preservation of the caricature of the polygons while maintaining topological consistency. An alternative approach is to determine a constraining arc on either side of a given arc that the generalised version of that arc may not intersect. The problem then becomes how to determine the constraining arcs.

This begs the question of what the characteristic structure of a polygon is. A polygon may be very simple, e.g., a triangle, or may have a very complicated structure. In natural resource datasets, there are often polygons with "spidery" structures. We need to be able to determine which parts of the polygon structure are significant (and hence need preserving) and which are insignificant (and hence may be generalised or pruned).

The **constrained Delaunay triangulation** of a polygon is a triangulation of the polygon vertices that includes the polygon edges and has the property that the smallest internal triangle angle is maximised when all triangles are removed from concavities and holes (Shewchuk 1996). The **skeleton** of a polygon is a "1-d analogue for an areal object and is also useful during merge and displacement operators" (Bundy, Jones, and Furse 1995). Figure 4(a) shows a generic polygon. Figure 4(b) shows the corresponding constrained Delaunay triangulation (dashed lines) and skeleton (solid lines). The skeleton vertices of degree two (solid circles) lie on the midpoint of each triangle edge not on the polygon boundary, while the skeleton vertices of odd degree (solid squares) lie at the centre of triangles of odd degree. A polygon whose boundary is disconnected, i.e. the polygon contains at least one "hole," has a corresponding skeleton that is a tree plus one additional edge for each hole.

Figure 5(a) shows a particular arc (in bold) to be generalised, and its two adjacent polygons with their corresponding skeletons. The arrows indicate the connection between the two arc nodes and the two skeletons; each arc node is connected to the closest skeleton vertex on each skeleton. Figure 5(b) shows the original arc to be generalised and the two skeletons which act as constraining trees. The skeletons are augmented by grafting the nodes to the closest skeleton vertex. To incorporate this method into DPR, we have modified Algorithm 2 by replacing the condition "$n \geq 2$ **and** $\max_i\{d(v_i, v_0, v_n)\} > \epsilon$" by the condition "$n \geq 2$ **and** $(\max_i\{d(v_i, v_0, v_n)\} > \epsilon$ **or** (v_0, v_n) intersects either constraint tree)." Figure 5(c) shows the original arc to be generalised and two constraining arcs obtained by pruning the skeletons of all extraneous branches which do not lie on the path from node to node through the tree. Using the full skeleton as a constraint maximizes the preservation of the character of the corresponding polygon, whereas the other extreme, using a completely pruned skeleton, corresponds to preserving only the topological consistency of the coverage when arcs are generalised in parallel. Intelligent use



(a) Selected arc and adjacent polygons and skeletons

(b) Polygon Skeletons as Constraints

(c) Pruned Polygon Skeletons as Constraints

Figure 5: Specification of Constraints to Line Generalisation

of pruning may produce more useful constraints since insignificant parts of a polygon's structure may be ignored while significant parts of a polygon can be retained in the generalised coverage.

3.3 Assessment of Line Generalisation Algorithms

Line smoothing can be distinguished from line generalisation, e.g., a line derived from a raster image may be smoothed to remove the saw-tooth or steps-and-stairs look. Here, we focus on polylines as a representation of part of a polygon boundary. In particular, there are two types of requirements which lead to the representation of a polyline by another polyline.

Line simplification is the representation of a polyline so that the caricature of the line is represented by fewer points. This often occurs from a desire to compress the storage associated with a line, from limitations of the plotting device, or from a wish to represent line resolution concomitant with polygon representation, i.e. representing no more detail in the line than would be in a polygon.

Line smoothing is the representation of a polyline so that there are fewer sharp angles and the aesthetics of the line are improved. This often occurs where a line is to be extracted from a raster representation.

McMaster (1987a) provides a suite of numerical evaluators for line simplification. However, often these are defined in terms of line simplification algorithms that simply select a subset of the existing vertices, such as DRP. For such old-new polyline pairs it is computationally efficient to calculate the area enclosed between the two polylines or the perpendicular distance of any point on one polyline from the other polyline. However, when the new polyline is not a subset of the old polyline, there can be ambiguities as to the exact area between the two polylines. This is often the case in line smoothing, since few of the original vertices are retained, and many additional vertices are added either to *densify* the polyline or through interpolating a set of vertex locations. We have developed a simple algorithm that detects the intersections of the two polylines in strict order from one end through to the other. Hence, each *component* of the pair of polylines may be evaluated independently and the results accumulated.

4 GIS Polygon Generalisation

A criterion for elimination of insignificant polygons is to select those polygons of small area. In *soil mapping*, size is not necessarily a good criterion. For example, there are often significant clusters of non-adjacent polygons that should be preserved in some generalised form. In addition, there are often isolated polygons that should somehow be represented in the final output. Hence, the relationship between polygons beyond simple adjacency (an artifact of the vector model) should be considered. The *select* and *eliminate* paradigm for GIS generalisation can be improved by considering alternative selection criteria and different elimination operations.

4.1 Selection of Polygons for Elimination

Two common principles for selection of polygons for elimination are to apply a DISSOLVE operator that merges adjacent polygons that have the same attribute value, and to select those polygons whose area lies below a given

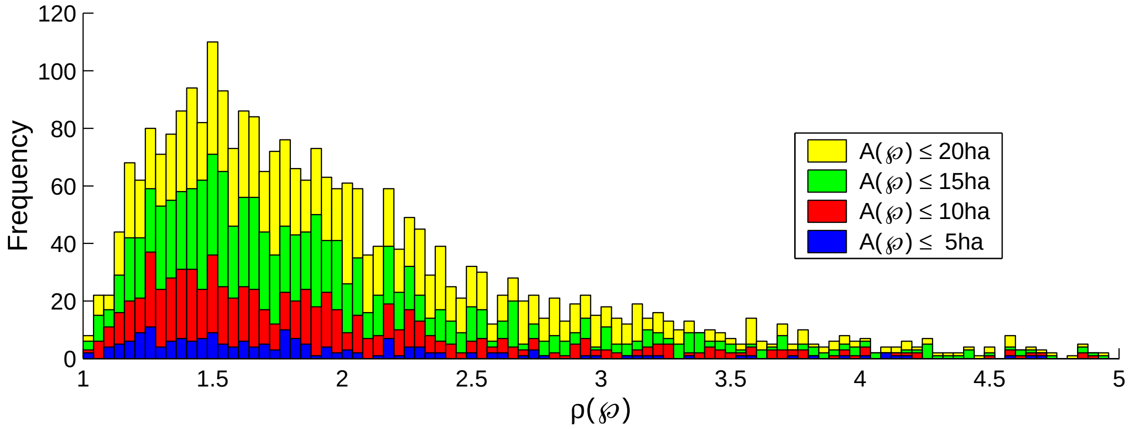


Figure 6: Histogram of $\rho(\wp)$ for North Island NZLRI Polygons

threshold for elimination. We have attempted to relax these two requirements by considering both the area and perimeter of an individual polygon, and by considering sets of polygons with the same attribute value that are nearly adjacent.

The standard approach ignores the significance of items that have small area but span a large spatial extent, such as a river network and the soils associated with rivers. These types of polygons have been identified as very important features to preserve in the generalisation process. We have used a perimeter squared to area ratio. A circle makes the most efficient use of space. For a polygon $\wp$, with area $A(\wp)$ and perimeter $P(\wp)$, let $\rho(\wp)$ be the ratio

$$\rho(\wp) = \frac{P(\wp)^2}{4\pi A(\wp)}.$$

This ratio is chosen because it is independent of scale and the most efficient ratio, that of the circle, has $\rho(\text{circle}) = 1$. Note that a square has $\rho(\text{square}) = \frac{4}{\pi}$, and any rectangle has $\rho(\text{rectangle}) \geq \frac{4}{\pi}$. Figure 6 illustrates the distribution of ρ over the North Island NZLRI polygons for four different levels of area. From this we may conclude that rather than consider area alone, we may consider both an area and a ρ constraint.

Simple generalisation logic ignores the nature of natural resource datasets: clusters of polygons that are not adjacent often form a natural structure such as a riparian strip. In this case, evaluation of the significance of individual polygons is not sufficient; we need to consider natural clusters or structures of nonadjacent polygons. Two polygons are **physically adjacent** if they share a common boundary arc. However, we may define two polygons as **virtually adjacent** if they are both part of some set (*cluster*) of polygons that should be considered as one for the purpose of elimination. In an operational sense, virtual adjacency may be defined in terms of distance between polygons or (since we want to model linear structures of sets of polygons) the distance between polygon skeletons. In this respect, dissolving adjacent polygons with the same attribute value is a cosmetic operation only.

4.2 Elimination Operators

Once a set (or cluster) of polygons has been selected as insignificant at the target resolution, the polygons need to be eliminated from the coverage. One method for "filling the holes", that of the Arc/Info ELIMINATE operator, is to merge a selected insignificant polygon with the polygon that either has the largest area of those polygons adjacent to the selected polygon, or has the longest common boundary with the selected polygon. Neither area, nor common boundary length are necessarily good criteria for assigning the area previously occupied by the selected polygon.

Traditionally, a polygon is eliminated by merging it with one adjacent polygon and removing the common boundary. This may alternatively be viewed as the *collapse* of a polygon onto part of its boundary and the reassignment of its area to the adjacent polygon. Two apparently separate activities may initiate a partial polygon collapse. In the context of polyline generalisation, a tree structure may provide appropriate generalisation of a polyline, e.g., an isthmus adjacent to a coastline. Also a "skinny" part of a polygon that has small area and corresponds to a branch of the polygon skeleton deemed insignificant to the overall structure of the polygon may be replaced by a polyline attached to the reduced polygon. These two possibilities are essentially perspectives on the same operation, demonstrating that line generalisation and polygon generalisation are not necessarily independent problems. The most general elimination operator is the *partial collapse* whereby one or more parts

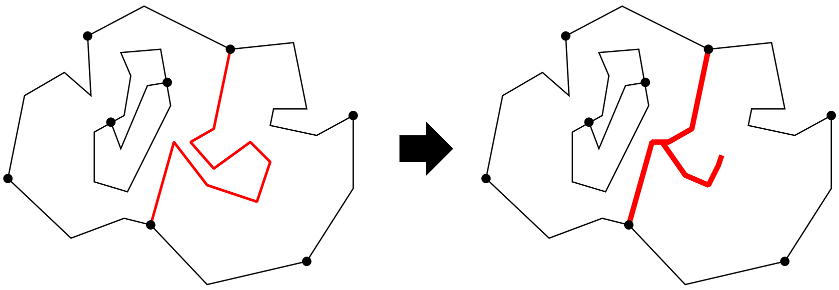
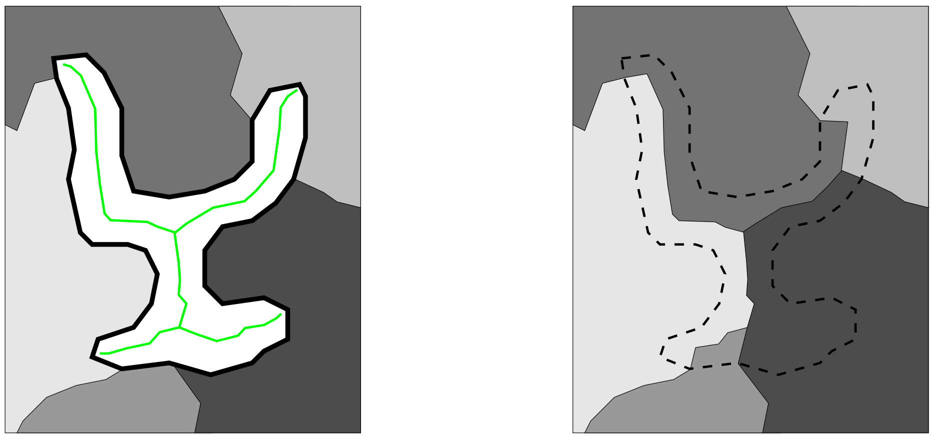


Figure 7: Partial-Collapse Operator



(a) Polygon (plus skeleton) selected for elimination (b) Area redistributed amongst adjacent polygons

Figure 8: Collapse-and-Eliminate Operator (adapted from Bader and Weibel (1997))

of a polygon are replaced by a linear feature (polyline or tree) connected to the remaining part of the polygon. Figure 7 illustrates one particular operator of this type. A special case is developed by Bader and Weibel (1997). Figure 8 illustrates a polygon skeleton collapsing for the entire polygon where the area of the eliminated polygon is redistributed amongst all adjacent polygons, and the new polygon boundaries are constructed from the skeleton of the eliminated polygon.

5 Conclusions

Although the distinction of GIS generalisation from cartographic generalisation is well defined in the literature, the required tools, logic and algorithms for GIS generalisation are less well developed. This paper has described the development of topologically consistent line generalisation within a polygonal dataset and discussed alternative criteria and operations for elimination of polygons or clusters of polygons. The result is encapsulated in Algorithm 3 ENHANCED SIMPLE GIS GENERALISATION which includes the idea of iteratively changing the spatial and thematic resolution in small steps, converging to the desired resolution. The *skeleton* of a polygon has proved useful in defining constraints on individual polygons that preserve the topological consistency of the coverage, chaining nearly adjacent polygons that represent some natural structure, partially collapsing polygons to linear features, and redistributing the associated area. Although the skeleton of a polygon is not unique, it appears to capture the significant structure of a polygon and lends itself well to collapsing an areal feature to a linear feature, i.e. collapsing part of a polygon onto the corresponding branch of the polygon skeleton.

Promising future research directions include further development of the use of skeletons and clustering of polygons. There are many stages of the generalisation process at which the branches of a polygon skeleton may be pruned, either in constrained line generalisation or for collapsing polygons. Also, skeleton-based operators may be useful for exaggeration of polygons, merging of polygons that are not adjacent, and for clustering of polygons. When a polygon is eliminated, its area is currently redistributed amongst its neighbouring polygons, regardless of the attribute value of these neighbours. An alternative is to redistribute the polygon area to the most similar neighbour, involving some kind of clustering operation. Rather than assigning the dominant attribute value to the cluster, we may record the proportional area contributed by each original polygon that has been

Algorithm 3 ENHANCED SIMPLE GIS GENERALISATION
loop over increments in resolution **do**
Perform topologically constrained line generalisation.
Select candidate insignificant polygon clusters.
Eliminate each insignificant polygon amongst its most appropriate neighbouring polygons.
end
end

merged into the cluster.

References

Bader, M. and R. Weibel (1997). Detecting and resolving size and proximity conflicts in the generalization of polygonal maps. In *Proceedings of the 18th International Conference of the International Cartographic Association*, Stockholm, Sweden.

Bondy, J. A. and U. S. R. Murty (1976). *Graph Theory with Applications*. North Holland.

Bundy, G. L., C. B. Jones, and E. Furse (1995). *Holistic generalization of large-scale cartographic data*, pp. 106–119. GISDATA 1. Taylor and Francis.

Daroussin, J. (1991). Kaleidos: une boite à outils complémentaire pour arc/info — actes du séminaire inra “gestion de l’espace rural et système d’information géographique”. Technical report.

Douglas, D. H. and T. K. Peucker (1973). Algorithms for the reduction of the number of points required to represent a digitzed line or its caricature. *Canadian Cartographer 10*(2), 112–122.

Hewitt, A. E. (1998). *New Zealand Soil Classification* (2nd ed.). Number 1 in Landcare Research Science Series. Lincoln, Canterbury, New Zealand: Manaaki Whenua Press.

McMaster, R. B. (1987a). Automated line generalization. *Cartographica 24*, 74–111.

McMaster, R. B. (1987b). The geometric properties of numerical generalization. *Geographical Analysis 19*(4), 330–346.

McMaster, R. B. (1989). The integration of symplification and smoothing algorithms in line generalization. *Cartographica 26*, 101–121.

Müller, J.-C., R. Weibel, J.-P. Lagrange, and F. Salgé (1995). Generalization: state of the art and issues. In J.-C. Müller, J.-P. Lagrange, and R. Weibel (Eds.), *GIS and Generalization: Methodology and Practice*, pp. 3–17.

NWASCO (1979). *Our Land Resources*. National Water and Soil Conservation Organisation.

Ramer, U. (1972). An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing 1*, 244–256.

Rijkse, W. C. and A. E. Hewitt (1995). *Soil map of the North Island and Soil map of the South Island, New Zealand Soil Classification 1:1000000 Scale*. Lincoln, New Zealand: Manaaki Whenua, Landcare Research.

Saalfeld, A. (1999). Topologically consistent line simplification with the Douglas-Peucker algorithm. `http://gscgas.eng.ohio-state.edu/dgpkfinl.pdf`.

Shewchuk, J. R. (1996). Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator. In M. C. Lin and D. Manocha (Eds.), *Applied Computational Geometry: Towards Geometric Engineering*, Volume 1148 of *Lecture Notes in Computer Science*, pp. 203–222. Springer-Verlag. From the First ACM Workshop on Applied Computational Geometry.

Weibel, R. (1995). Three essential building blocks for automated generalization. In J.-C. Müller, J.-P. Lagrange, and R. Weibel (Eds.), *GIS and Generalization: Methodology and Practice*, pp. 56–69. Taylor & Francis Ltd.

Weibel, R. and G. Dutton (1999). Generalising spatial data and dealing with multiple representations. In P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind (Eds.), *Geographical Information Systems: Volume 1 Principles and Technical Issues* (2nd ed.)., pp. 125–155.

White, E. R. (1985). Assessment of line-generalization algorithms using characteristic points. *The American Cartographer 12*(1), 17–27.