Barkowsky, T., Latecki, L. J., & Richter, K.-F. (2000). Schematizing maps: Simplification of geographic shape by discrete curve evolution. In C. Freksa, W. Brauer, C. Habel, & K. F. Wender (Eds.), *Spatial Cognition II - Integrating abstract theories, empirical studies, formal models, and practical applications* (pp. 41-53). Berlin: Springer.

Schematizing Maps: Simplification of Geographic Shape by Discrete Curve Evolution[1]

Thomas Barkowsky[*], Longin Jan Latecki[†], Kai-Florian Richter[*]

[*]University of Hamburg, Department for Informatics
and Cognitive Science Program
Vogt-Kölln-Str. 30, 22527 Hamburg, Germany
{barkowsky, 4krichte}@informatik.uni-hamburg.de

[†]TU Munich, Zentrum Mathematik and
University of Hamburg, Department of Applied Mathematics
Bundesstr. 55, 20146 Hamburg, Germany
latecki@math.uni-hamburg.de

Abstract. Shape simplification in map-like representations is used for two reasons: either to abstract from irrelevant detail to reduce a map user's cognitive load, or to simplify information when a map of a smaller scale is derived from a detailed reference map. We present a method for abstracting simplified cartographic representations from more accurate spatial data. First, the employed method of *discrete curve evolution* developed for simplifying perceptual shape characteristics is explained. Specific problems of applying the method to cartographic data are elaborated. An algorithm is presented, which on the one hand simplifies spatial data up to a degree of abstraction intended by the user; and which on the other hand does not violate local spatial ordering between (elements of) cartographic entities, since local arrangement of entities is assumed to be an important spatial knowledge characteristic. The operation of the implemented method is demonstrated using two different examples of cartographic data.

1 Map Schematization

Maps and map-like representations are a common means for conveying knowledge about spatial environments that usually cannot be surveyed as a whole, like, for example, local built areas, cities, or entire states or continents. Besides the large spatial extent of these *geographic spaces* (Montello, 1993) an important characteristic is their complexity what regards possible aspects that can be depicted in a map. Dependent on its scale, a general topographic map is intended to depict as much spatial information as possible, since the specific purpose the map will be used for is not known in advance.

[1] This work is supported by the Deutsche Forschungsgemeinschaft (DFG) under grants Fr 806-8 ('Spatial Structures in Aspect Maps', Spatial Cognition Priority Program) and Kr 1186-1 ('Shape in Discrete Structures').

Performing a given task, however, usually only requires a rather small subset of spatial knowledge aspects extractable from a general purpose geographic map. Therefore special purpose *schematic maps* are generated which are only suitable for restricted purposes, but which, on the other hand, ease their interpretation by concentrating on relevant aspects of information by abstracting from others.

Schematic public transportation network maps are a common example of schematic maps (e.g. Morrison, 1996). In this type of map-like representations most entities not directly relevant for using busses, underground trains, etc. are omitted (see Fig. 1 as an example). So, these kinds of maps concentrate on stations, the lines connecting them, and some typical features helpful for the overall orientation within the city at hand. Especially, they usually abstract from detailed shape information concerning the course of the lines and other spatial features, like waters. As a consequence, schematic maps often convey qualitative spatial concepts thus adapting to common characteristics of mental knowledge representation (Freksa et al., 1999).

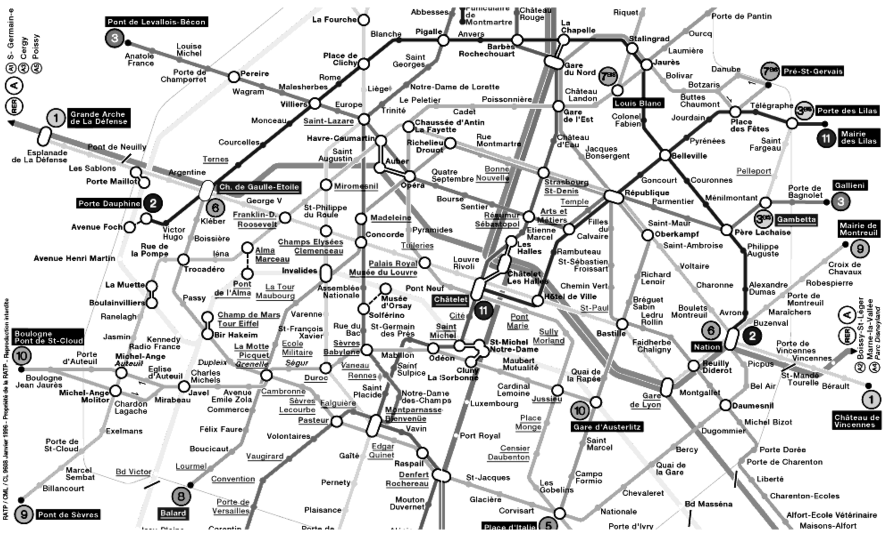

Fig. 1. Example of a schematic public transportation network map (Paris)

Intending to derive a schematic map with reduced shape information from detailed spatial data, we need techniques for reducing spatial accuracy of shape aspects to abstract from the exact course of linear entities or the detailed shape of areal geographic objects.

Shape simplification processes are needed not only for the generation of schematic map-like representations but also in general cartographic contexts. When a map of a smaller scale (e.g. an overview map of a larger area) is intended to be constructed using reference map data of larger scale, it is usually necessary to perform a reduction in spatial detail. Therefore, apart from the use of symbolic replacement, displacement, and resizing of cartographic entities, an important step in cartographic generalization is the simplification of details of spatial features (Hake & Grünreich, 1994). Automated cartographic generalization is a major research issue in the area of *geographic information systems* (GISs) (for an overview, see Müller et al., 1995; Jones, 1997).

We employ discrete curve evolution (Latecki & Lakämper, 1999a, 1999b, 1999c) as a technique for simplifying shape characteristics in cartographic information, which will be presented in the following section.

2 Discrete Curve Evolution

The main accomplishment of the discrete curve evolution process described in this section is automatic simplification of polygonal curves that allows to neglect minor distortions while preserving the perceptual appearance. The main idea of discrete curve evolution is a stepwise elimination of kinks that are least relevant to the shape of the polygonal curve. The relevance of kinks is intended to reflect their contribution to the overall shape of the polygonal curve. This can be intuitively motivated by the example objects in Fig. 2. While the bold kink in (a) can be interpreted as an irrelevant shape distortion, the bold kinks in (b) and (c) are more likely to represent relevant shape properties of the whole object. Clearly, the kink in (d) has the most significant contribution to the overall shape of the depicted object.

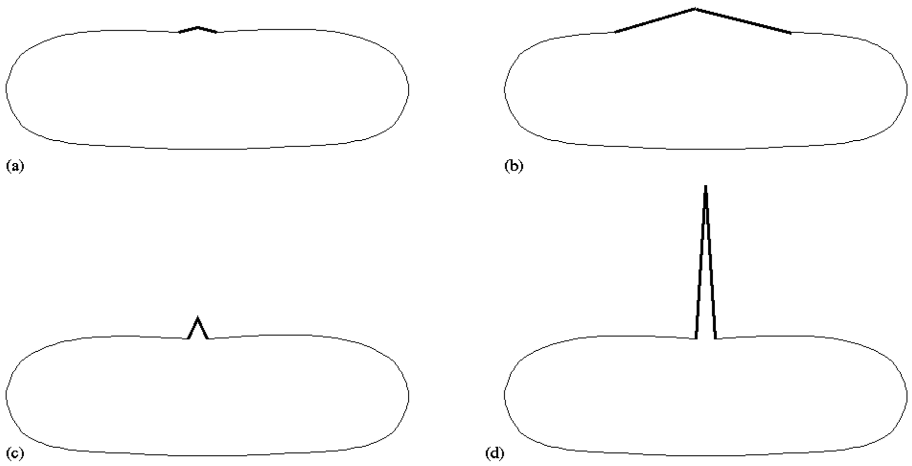

Fig. 2. The relevance measure *K* of the bold arcs is in accord with our visual perception

There exist simple geometric concepts that can explain these differences in the shape contribution. The bold kink in Fig. 2 (b) has the same turn angle as the bold kink in (a) but is longer. The bold kink in (c) has the same length as the one in (a) but its turn angle is greater. The contribution of the bold kink in Fig. 2 (d) to the shape of the displayed object is the most significant, since it has the largest turn angle and its line segments are the longest.

It follows from this example that the shape relevance of every kink can be defined by the turn angle and the lengths of the neighboring line segments. We have seen that the larger both the relative lengths and the turn angle of a kink, the greater is its contribution to the shape of a curve. Thus, a cost function *K* that measures the shape relevance should be monotone increasing with respect to the turn angle and the lengths of the neighboring line segments. This assumption can also be justified by the rules on salience of a limb in (Siddiqi & Kimia, 1995).

Based on this motivation we give a more formal description now. Let s_1, s_2 be two consecutive line segments of a given polygonal curve. More precisely, it seems that an adequate measure of the relevance of kink $s_1 \cup s_2$ for the shape of the polygonal curve can be based on turn angle $\beta(s_1, s_2)$ at the common vertex of segments s_1, s_2 and on the lengths of the segments s_1, s_2. Following (Latecki & Lakämper, 1999a), we use the relevance measure *K* given by

$$K(s_1, s_2) = \frac{\beta(s_1, s_2) l(s_1) l(s_2)}{l(s_1) + l(s_2)} \tag{1}$$

where *l* is the length function. We use this relevance measure, since its performance has been verified by numerous experiments (e.g. Latecki & Lakämper, 2000a, 2000b). The main property of this relevance measure is the following:

- The higher the value of $K(s_1, s_2)$, the larger is the contribution of kink $s_1 \cup s_2$ to the shape of the polygonal curve.

Now we describe the process of *discrete curve evolution*. The minimum of the cost function *K* determines the pair of line segments that is substituted by a single line segment joining their endpoints. The substitution determines a single step of the discrete curve evolution. We repeat this process for the new curve, i.e., we determine again the pair of line segments that minimizes the cost function, and so on. The key property of this evolution is the order of the substitution determined by *K*. Thus, the basic idea is the following:

- In every step of the evolution, a pair of consecutive line segments s_1, s_2 with a smallest value of $K(s_1, s_2)$ is substituted by a single line segment joining the endpoints of $s_1 \cup s_2$.

Observe that although the relevance measure *K* is computed locally for every stage of the evolution, it is not a local property with respect to the original input polygonal curve, since some of the line segments have been deleted. As can be seen in Fig. 3, the discrete curve evolution allows to neglect minor distortions while preserving the perceptual appearance.

A detailed algorithmic definition of this process is given in (Latecki & Lakämper, 1999a). A recursive set theoretic definition can be found in (Latecki & Lakämper, 1999c). Online examples are given on the web page www.math.uni-hamburg.de/home/lakaemper/shape. This algorithm is guaranteed to terminate, since in every evolution step, the number of line segments in the curve decomposition decreases by one (one line segment replaces two adjacent segments). It is also obvious that this evolution converges to a convex polygon for closed polygonal curves, since the evolution will reach a state where there are exactly three line segments in the curve decomposition, which clearly form a triangle. Of course, for many closed polygonal curves, a convex polygon with more then three sides can be obtained in an earlier stage of the evolution.

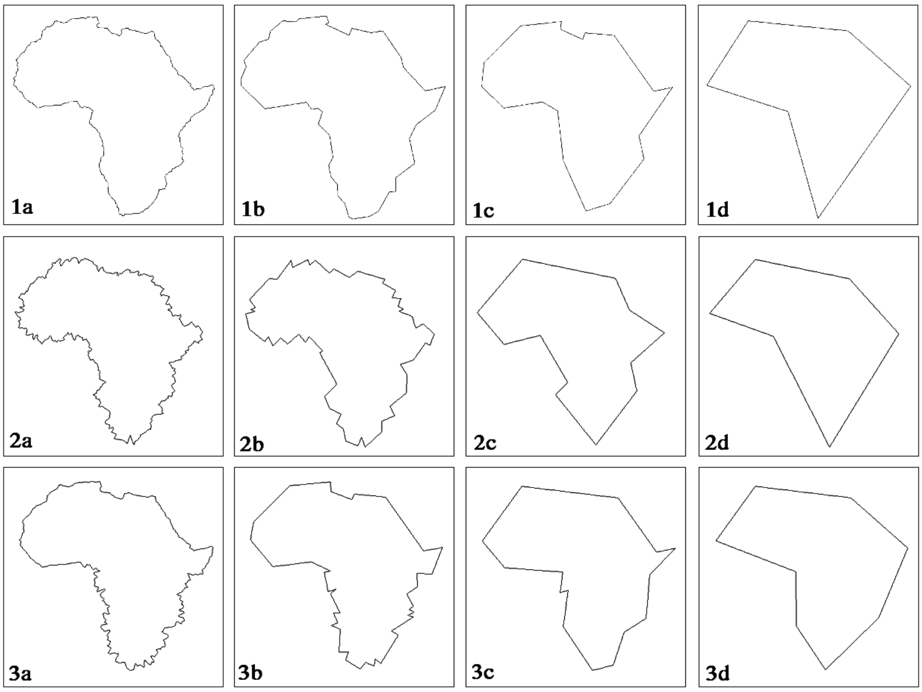

Fig. 3. Examples of discrete curve evolution. Each row shows only a few stages

3 Application to Map Schematization

In the previous section we have seen how discrete curve evolution is used to simplify the shape of closed curves. Entities depicted in geographic maps are point-like, linear, or areal. Shape simplification can be applied both to linear and areal entities. However, when dealing with geographic information, the discrete curve evolution method has to be extended in several respects.

We have seen that the relevance measure of a kink is computed from the two adjacent line segments. Therefore, for geographic objects represented by a single isolated point and for the endpoints of linear objects in spatial data sets no relevance measure can be computed as there is none or only one adjacent line segment (cf. Fig. 4a). As a consequence, these points cannot be included in the discrete curve evolution process (they should not be intended to be eliminated, either).

On the other hand, if points belong to more than one object, we may not want to eliminate them, as this might seriously change spatial information. Consider as an example a linear object (e.g. a state border) being in part connected to the boundary of an areal object (e.g. a lake) as illustrated in Fig. 4b. When such points are eliminated or displaced in either of the two objects by the discrete curve evolution the spatial feature of being connected to each other may be (at least in part) modified. We will call such points *fix points*.

These two cases make it plausible that not every point in an object may be eliminated or displaced by the discrete curve evolution. Therefore, *fix points* are introduced as points not to be considered by the simplification procedure, be it that it is not possible to assign a relevance measure to them, be it that they must not be eliminated.

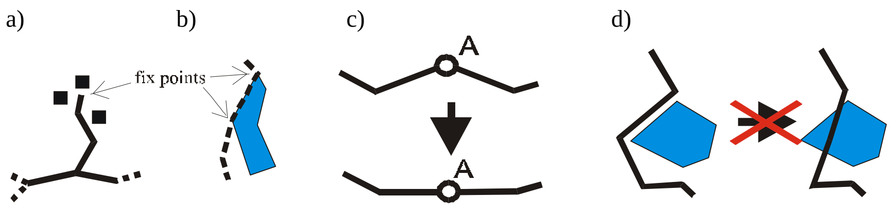

Fig. 4. Examples for cases in cartographic data which require the extension of the discrete curve evolution method

Thus, point-like cartographic entities are considered as fix points as the simplification process does not apply to them; consequently, their positions are not changed during shape simplification. However, in geographic data, point-like objects are often located on linear objects (for example cities lying on roads they are connected by, or train stations in the above public transportation network example). In this case the position of the point-like objects will have to be moved when the linear object they belong to is modified. So, when a line segment is straightened in the discrete curve evolution process a point-like object located on this line has to be projected back onto the line afterwards (Fig. 4c). We will call such point-like entities *movable points*.

According to the above discussion, we obtain the following classification of points in a given map:

1. *Fix points* – points that represent geographic features that cannot be removed and whose position cannot be changed.
2. *Movable points* – points that represent geographic entities that cannot be removed but whose position can be changed.
3. *Removable points* – points that can be deleted; they usually occur as parts of polygonal curves and do not represent individual geographic entities.

However, changing a point's position or deleting a point is only possible if the local arrangement of points is not changed. Before we state this restriction more precisely, we give an intuitive motivation for it. There are many cases in which we will not want to straighten kinks between line segments. Consider, for example, the case of a point-like or an areal entity (e.g. a city, a lake) being located on the concave side of a kink belonging to another object (e.g. a street). If the kink is to be straightened in a discrete curve evolution step it may be the case that the obtained straight line segment intersects the other object (Fig. 4d). In the case of a point-like object located next to the line segment the simplification step may result in the point being located on the other side of the line. Clearly, we usually should not modify the spatial relationships between geographic objects in such a severe way. The restriction that local arrangements of spatial entities must not be changed is a fundamental spatial knowledge aspect used in reasoning with geographic maps (cf. Berendt et al., 1998a, 1998b).

Therefore, we will check whether local ordering relations (cf. Schlieder, 1996) between objects are violated before a simplification is performed (i.e., the schematization process is performed in a context-sensitive way)[2]. Notice that this check only concerns (parts of) entities in immediate vicinity to the line segments in question. Violations of spatial ordering between objects in farther distances - although they may occur during the schematization process as a whole - can be assumed not to affect the extraction of appropriate information from the map by a map user.

In the next section we will show how these requirements are implemented by the simplification process for cartographic representations.

4 The Algorithm

As elaborated in the previous section, the algorithm for performing the simplification task based on the discrete curve evolution requires some additional considerations. However, it is still fairly easy.

Point-like entities are considered as *fix points* as stated in the previous section. What regards linear and areal objects we start by computing the relevance measures for all (non-end) points and sort them by their respective relevance measures. Notice that points that have a relevance measure of zero do not need to be dealt with, since they are already on a straight line.

Then we consider the first point in the sorted list (i.e. the point which has the lowest relevance measure). Two cases have to be distinguished: (A) The coordinates of the point at hand are unique or (B) there is more than one object at the given point's coordinates.

Case A: Unique coordinates. Since the point's coordinates are unique there is no interaction with any other object in this point. Endpoints of linear objects are treated as *fix points* in this case as no relevance measure can be computed for them. For non-endpoints two cases have to be distinguished:

1. If the point at hand is a *removable point* we may delete it unless this does not violate the local arrangement of spatial entities. Therefore we check whether there is any other point inside the triangle formed by the original and the new straightened line. If so, it is not possible to straighten the line since this would change the local

[2] Clearly, it would also be possible to modify both, the object to be simplified and the conflicting object (e.g. by moving it aside). However, due to the degrees of freedom (as well as further possible conflicts) this does not seem to be a feasible approach.

spatial arrangement of the two objects involved. The point at hand is treated as *fix point* in this case.

If there is no point in the triangle the line is straightened by removing the point at hand. Finally, the relevance measures for the adjacent points are recalculated. When relevance measures are not normalized with respect to the total length of entities (see previous section) it is sufficient to recalculate just the relevance measures for the immediately neighboring points since in this case the relevance measure is only based on local shape characteristics. Afterwards we resort the list of relevance measures and start all over again.

2. If the point at hand is a *movable point* (i.e. it has to be preserved) it is treated like described before. Instead of removing it, however, its new coordinates on the straightened line are calculated and the point is projected to this position. After moving the point its relevance measure is zero for it is now on a straight line; therefore, it will not receive further consideration in discrete curve evolution.

Case B: Multiple objects at the given coordinate. If the point's coordinates belong to more that one object, there are interactions between at least two objects in the point at hand. Three cases have to be distinguished:

1. If the coordinates of the immediately neighboring points of all objects involved are pairwise identical, then the objects the points belong to have common coordinates for at least three subsequent vertices. In this case the middle vertex (the point position at hand) is a *removable point* as long as simplification is performed simultaneously for all objects involved. Therefore we can proceed in analogy to removable points as described for case A (see above). *Movable points* (which shall not be eliminated from either object) can also be treated in analogy to movable points in case A. In this case all points at the middle vertex's coordinates are simultaneously transferred to their new common position on the straightened line.

2. The current point's position coincides with an endpoint of one object and for the other objects holds that the coordinates of the immediately neighboring points are pairwise identical. In this case the point at hand is treated as a *movable point* (which is - regarding the endpoints - in contrast to case A). The respective endpoint is projected onto the straightened line. So, although the endpoints still have no relevance measure, they are moved in accordance with the points they coincide with.

3. In any other case the point at hand has to be treated as a *fix point* and cannot be modified.

If the point at hand has been identified as a fix point by the test under cases A or B we continue with the next point in the sorted list of relevance measures (i.e. the point with the next lowest relevance measure). If there is no non-fix point left in the list or the abort criterion (see below) is reached we are done and the algorithm terminates.

Usually a certain value of the relevance measure is chosen as abort criterion but it is also possible to specify a value of turn angle as a stop condition of the simplification. These abort criteria specify the degree of schematization and have to be chosen in accordance with the schematic map's intended purpose of use or other design requirements. So, if the point at hand chosen for elimination has a higher relevance measure (or a larger turn angle), then the abort criterion is fulfilled and the algorithm stops. It is important to realize that this check is made against the point that is about to be changed and not against the point with the lowest relevance measure. Since points that do not fulfill the above preconditions may not be considered any more, it is possible that a point is tested whose relevance measure is substantially higher than the point with the lowest relevance measure.

The algorithm has been implemented and tested with different types of geographic data. In the following section we demonstrate the operation of the algorithm showing two examples.

5 Examples

The first example employs a simple schematic public transportation network map as basic data for the algorithm. Schematic transportation maps usually abstract both from the exact positions of stations and the detailed course of lines connecting the stations.

Figure 5a shows the original data set containing linear and point-like entities. The schematic map exhibits the exact positions of subway and city train stations of a part of the Hamburg public transportation network. Obviously, some stations belong to more than one subway line. The stations at the ends of the lines have to be considered as fix points as they do not allow for computing a relevance measure for them. The line courses connecting the stations have already been abstracted to straight lines. In the course of the simplification process minor kinks in the line courses are eliminated (Fig. 5b and c) while the overall appearance of the network map is preserved.

This simple example already exhibits the most important features of the simplification process developed in the previous sections:

- As the point-like stations (depicted as squares) must be preserved (i.e. must not be deleted during the simplification process) they are always projected back onto the lines during the evolution process. So they are treated as *movable points*. As only straight line connections are used in this example, there are no *removable points*.
- Endpoints of lines (usually to be treated as *fix points*) which are simultaneously part of another line are moved according to the simplification of the other line.
- Simplification is only performed as far as the local spatial arrangement between lines and stations is not violated. In other words, in the simplification process no station changes sides with respect to any line (seen from a given direction).

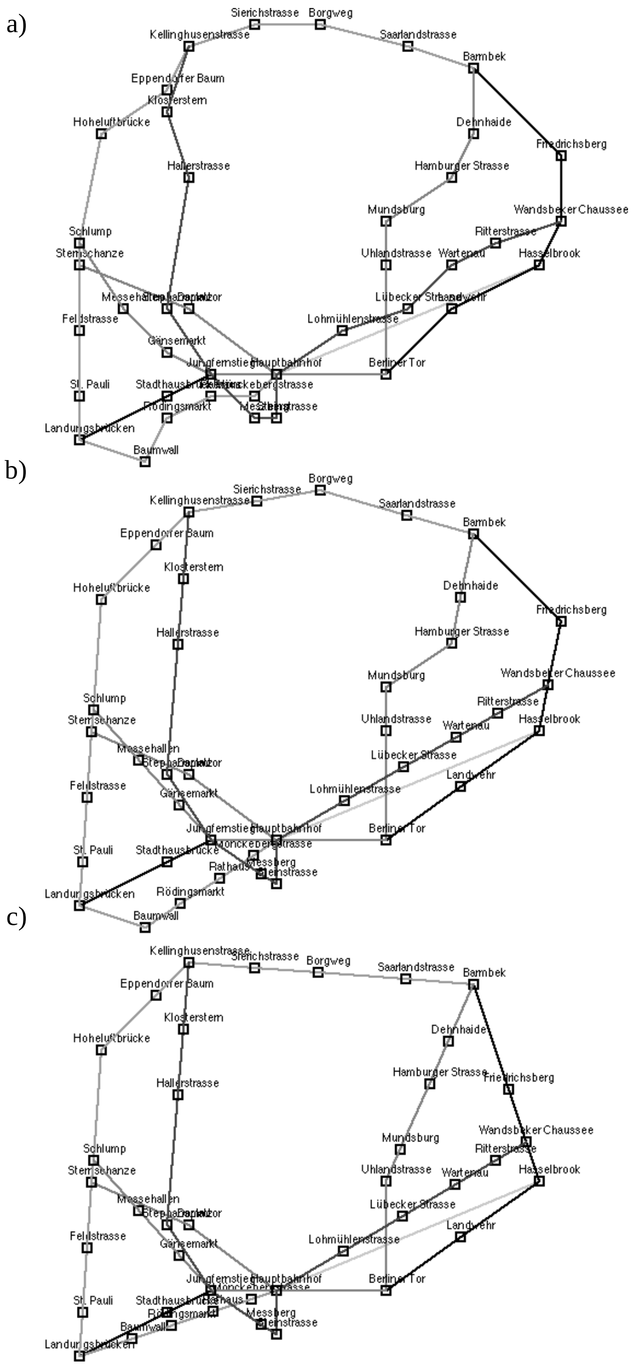

Fig. 5. Simplification of a public transportation network map (Hamburg). Original plan (a) and two schematizations up to different simplification levels

The employed map simplification method, however, does not deal with the local concentration of entities on the map. To improve the readability of schematic maps, the map's scale may be locally distorted (for example in the downtown region where concentration of relevant entities is higher than in the outskirts). For this purpose the presented method may be combined with constraint-based methods for graph layout (e.g. Kamps et al., 1995; He & Marriott, 1996).

For the second and more complex example we used an original ARCINFO data set of the Cologne (Germany) conurbation (Fig. 6a). The original data set showing cities, streets and waters of the region (Fig. 6b) comprises approx. 2700 points. Figure 6d shows the maximal simplification that can be performed without violating local spatial arrangement between spatial entities. This maximally simplified data set consists of approx. 600 points. An intermediate simplification level, i.e. a simplification up to a specified relevance measure is shown in Fig. 6c (approx. 750 points).

In this example the point-like entities representing cities have to be treated as *fix points*, i.e. they must neither be deleted nor moved. The points in the linear entities (streets and waters) can generally be considered as removable; i.e., they can be eliminated during the evolution process as far as spatial ordering relations between point-like and linear features are preserved. Observe that no further point in the data set shown in Fig. 6d can be removed without violating local spatial arrangements.

As the different types of linear entities (i.e., streets and waters) play different roles in the interpretation of the resulting map, it may be sensible to use different degrees of schematizations for either of them. For example, streets may be simplified up to the highest possible degree, whereas waters may remain curved to a certain extent to convey their characteristics of being natural entities. For this purpose, different types of geographic objects may be assigned different threshold values for schematization.

6 Conclusion and Outlook

We presented a method for simplifying cartographic shape information based on discrete curve evolution. The basic method of the discrete curve evolution has been extended to meet specific requirements of geographic data depicted in map-like representations.

These qualitative restrictions suitable for geographic data sets can be further extended to make the method usable for more constrained spatial tasks, which may be necessary in some applications. For example, it may be particularly important to preserve the order of objects with respect to their relative distances or with respect to their cardinal directions. Also, concepts for spatial proximity (for example a city being located near the coast) may be intended to be preserved by automatic map simplification processes.

The modified discrete curve evolution presented in this contribution can also be used as a pre-filter for similarity measures of geographic representations. Similar to the shape similarity measure of object contours in (Latecki & Lakämper, 1999b), it seems to be cognitively and computationally desirable to simplify representations before determining their similarity. The fact that the modified discrete curve evolution preserves relevant qualitative properties is of primary importance for this task, since similarity of geographic representations not only depends on metric features but also on qualitative properties (e.g., is a point-like object A to the right or to the left of a line object B?). Similarity measures of geographic representations have many potential applications, e.g., in *geographic information systems* (GISs) or in the automatic matching of satellite images with existing geographic data.

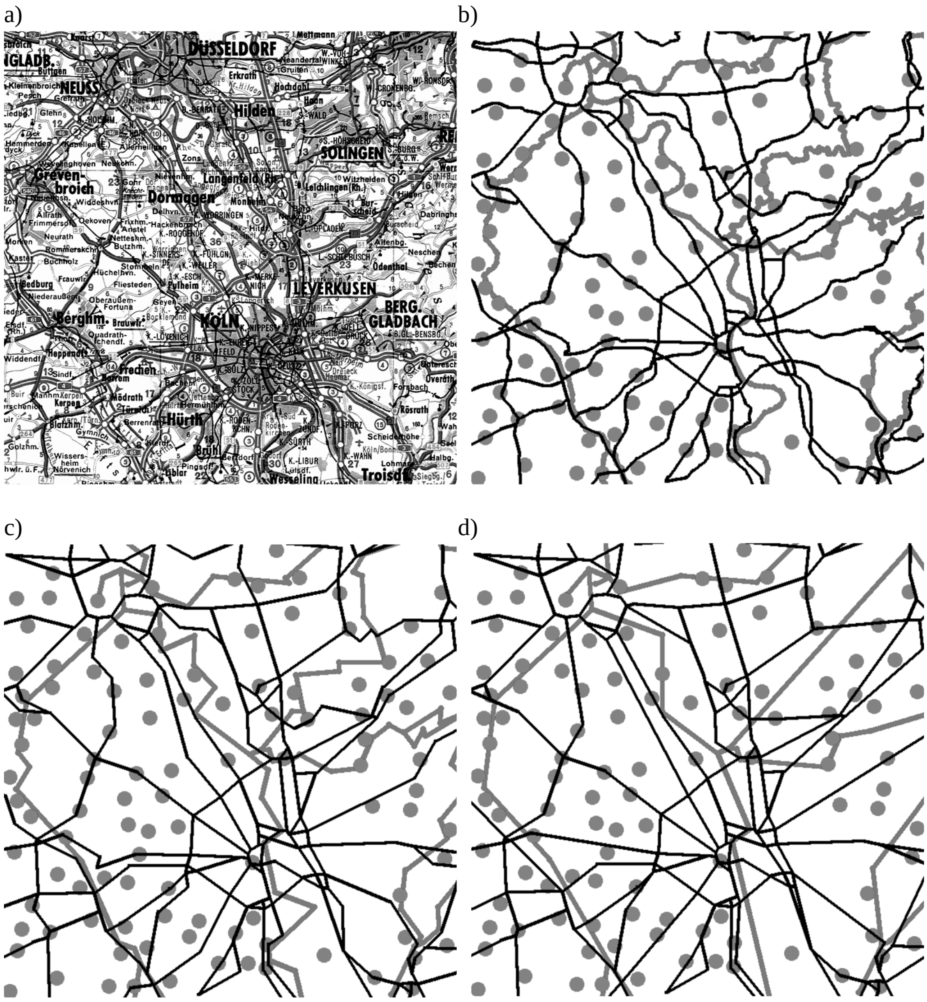

Fig. 6. Example of simplification of an ARCINFO data set: (a) conventional map of the Cologne (Germany) conurbation; (b) the original ARCINFO data set containing cities, streets, and waters; (c) intermediate simplification level; (d) maximal simplification preserving local spatial arrangement

Acknowledgments

We would like to thank Sabine Timpf and an anonymous reviewer for helpful comments on this contribution.

References

Berendt, B., Barkowsky, T., Freksa, C., & Kelter, S. (1998a). Spatial representation with aspect maps. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial cognition - An interdisciplinary approach to representing and processing spatial knowledge* (pp. 313-336). Berlin: Springer.

Berendt, B., Rauh, R., & Barkowsky, T. (1998b). Spatial thinking with geographic maps: an empirical study. In H. Czap, H. P. Ohly, & S. Pribbenow (Eds.), *Herausforderungen an die Wissensorganisation: Visualisierung, multimediale Dokumente, Internetstrukturen* (pp. 63-73). Würzburg: ERGON Verlag.

Freksa, C., Barkowsky, T., & Klippel, A. (1999). Spatial symbol systems and spatial cognition: A computer science perspective on perception-based symbol processing. *Behavioral and Brain Sciences, 22*(4), 616-617.

Hake, G., & Grünreich, D. (1994). *Kartographie*. Berlin, New York: de Gruyter.

He, W., & Marriott, K. (1996). Constrained graph layout. In S. North (Ed.), *Graph drawing (Proceedings of the Symposium on Graph Drawing, GD'96, Berkeley, CA, USA)* (pp. 217-232). Berlin: Springer.

Jones, C. B. (1997). *Geographical information systems and computer cartography*. Essex: Longman.

Kamps, T., Kleinz, J., & Read, J. (1995). Constraint-based spring-model algorithm for graph layout. In R. Tamassia & I. G. Tollis (Eds.), *Graph drawing* (pp. 349-360). Berlin: Springer.

Latecki, L. J., & Lakämper, R. (1999a). Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding, 73*(3), 441-454.

Latecki, L. J., & Lakämper, R. (1999b). Contour-based shape similarity. *Proc. of Int. Conf. on Visual Information Systems*, Amsterdam, June 1999. Berlin: Springer.

Latecki, L. J., & Lakämper, R. (1999c). Polygon Evolution by Vertex Deletion. *Proc. of Int. Conf. on Scale-Space Theories in Computer Vision*, Corfu, Greece, September 1999.

Latecki, L. J., & Lakämper, R. (2000a). Application of planar shape comparison to object retrieval in image databases. *Pattern Recognition (PR)*, to appear.

Latecki, L. J., & Lakämper, R. (2000b). Shape similarity measure based on correspondence of visual parts. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, to appear.

Montello, D. R. (1993). Scale and multiple psychologies of space. In A. Frank & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS. LNCS No. 716* (pp. 312-321). Berlin: Springer.

Morrison, A. (1996). Public transport maps in western European cities. *The Cartographic Journal, 33*(2), 93-110.

Müller, J. C., Lagrange, J. P., & Weibel, R. (Eds.). (1995). *GIS and generalization: Methodology and practice*. London: Taylor & Francis.

Schlieder, C. (1996). Qualitative shape representation. In P. Burrough & A. Frank (Eds.), *Geographic objects with indeterminate boundaries* (pp. 123-140). London: Taylor & Francis.

Siddiqi, K., Kimia, B. B. (1995). Parts of visual form: Computational aspects. *IEEE Trans. PAMI*, 17, 239–251.