An approach for using cubic Bézier curves for schematizations of categorical maps

A. W. Reimer[1] and Chris Volk[2]

[1]University of Heidelberg, Berliner Straße 48, 69120, Germany
Email: andreas.reimer@geog.uni-heidelberg.de

[2]University of Technology Eindhoven, P.O.Box, Hoofdgebouw 5600 MB, Eindhoven, Netherlands
Email: ch.t.volk@gmail.com

1. Introduction

Parametrized curves, and Bezier curves [BC] in particular, are staples of graphic design, font recognition, object-path description in animation and manual cartography. In contrast, modelling of geodata in GI-systems and geodatabases have so far only used curves as intermediary step. In generalization research with its traditional preoccupation with topographic mapping, the graphic complexity of topographic geo-objects was well suited by approximating curvatures with polylines. Elements such as isolines, roads or rivers are shown at such detail, that a copious amount of points is used to form the wanted shape and produce the illusion of smooth curves if that effect is wanted.

If we understand schematization as design-rule driven effort of minimizing non-functional detail according to a target complexity instead of a target scale, then BCs themselves might become desirable or even necessary elements. Most schematization work has been applied to network-forming structures composed of lines. Research into manually produced area-class or categorical sketches and schematizations has shown that subdivisions are often demarcated by BCs (Reimer 2010). We interpret the usage of a very low number of cubic BCs as being part of the aesthetic and conceptual principles, i. e. design-rules of some schematizations. As general quality measures for evaluating schematizations do not exist, using a predefined number of BCs and thereby imitating manual techniques serves as a stand-in until further progress is made in the fundamental question of evaluation. We provide one way of schematizing the internal subdivisions of a geographic region with cubic BCs and discuss the results.

2. Related Work

The problem we address can be described as:

Automatically schematize a categorical map with BCs according to design-rules found in manually schematized maps.

In our running example, the pertinent design rules derived empirically in the related work are:

- *To emphasize the 'container' aspect of a given region of interest, the territorial outline should be depicted differently from the internal subdivision and presented as free-standing entity i. e. as an Inselkarte.*

- *Territorial outlines should have between 8-15 vertices and a high amount of 'parrallelity'.*

- *Subdivision boundaries should be drawn with cubic Bezier splines consisting of one (most common) up to three (seldom) cubic BCs. Some cubic BCs might have two identical control points, effectively representing quadratic BCs.*

Below we provide a short overview of the related work regarding the mentioned sub-problems.

2.1 Territorial outlines

The territorial outline for Italy that is used in our running example has been generated with the method presented by Reimer and Meulemans (2011). In their method, outlines are computed by selecting characteristic points of a given territorial outline. These inputs are used for a simulated annealing process on the vertices and edges that attempts to maximize "parallelity". As with preceding work, it is here applied to a national outline.

2.2 Categorical maps

The general problems concerning choropleth maps as opposed to categorical or area-class maps are well known (Mark and Csillag 1989), but often the actual generalization of categorical geometries remains the realm of manual generalisation (e. g. Bucher and Schlömer 2006), see also Figure 1.

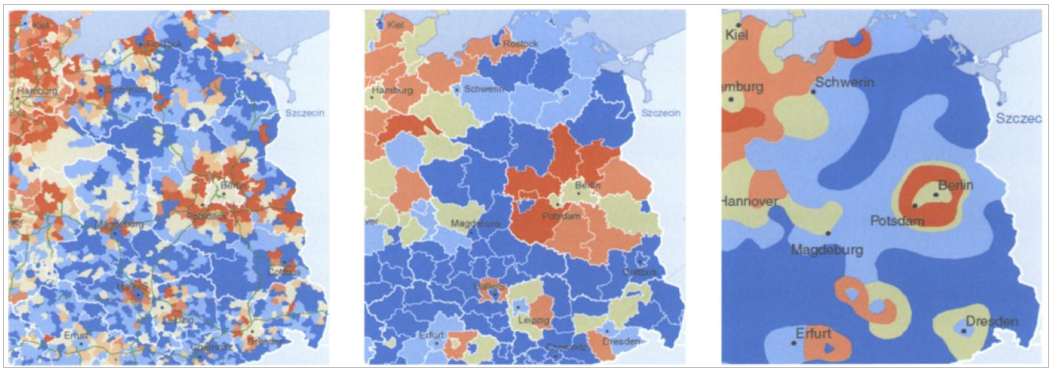

Figure 1. Illustration of the generalization problems with choropleth vs area-class maps; from left to right historical data, choropleth of prognosis data and manually generalized prognosis for east German population (from Bucher and Schlömer 2006).

For the case of a changing base-map or territorial outline, this problem is even heightened. As the general approach in this paper is to treat the internal subdivision different from the outline, we indeed need to map the interior of one area into the interior of another with a schematized shape. Saalfeld (2001) provided an existence proof for area-preserving affine mappings. In essence he proved that a continuous mapping from one polygonal subdivision to another one that preserves area and topology is always possible. Saalfeld (2001) does underline the fact that his procedure is computationally very expensive and basically impractical. He suggests that alternative methods should be developed. To our knowledge, no follow-up work has been presented since.

Galanda (2003) has provided a framework for constraint formulation for polygonal generalization. Steiniger and Weibel 2005's work on horizontal and vertical relations in categorical maps form the basis for considerations leading to our insertion strategies.

2.3 Bezier curves

Computer graphics research, especially in CAD and font/sketch recognition has an ample array of algorithms concerned with BCs. The basic properties of BCs can be found for example in Foley et al. 1990. The properties our approach makes use of are *convex hull, variation diminishing, affine invariance* and *pseudo local control*.

Although there is also a considerable amount of curve fitting techniques, nearly all algorithms fitting BCs to some input line work with piecewise techniques. That is to say that once a cubic curve is found to be unable to match the target polyline a new cubic BC is added, usually while keeping some form of continuity (e.g. Schneider 1990, or Pal et al 2007). Our aim is to find a suitable fit to a given boundary polyline with just a single cubic BC, a task which only few algorithms (e. g. Masood and Ejaz 2010) concern themselves with. While this is known to be difficult (Shao and Zhou 1996), our goal of schematizations allows for a greater degree of geometric error than other applications. While some manual schematizations sometimes use splines of up to three chained BCs, for this paper we only allow a single cubic curve. This provides the starkest schematization and has to our knowledge not been addressed so far. While it is naturally easier to fit a spline to a polyline than just fitting a single curve, this would dilute the purpose of minimizing output complexity. The addition of extraneous bends into schematizations has been established as being bad design (e. g. Roberts 2012). By limiting ourselves to single cubic BCs, we constrain the number of bends to a maximum of two. In lack of other constraints for curve schematization and because of the danger of overfitting the splines to polylines, we chose this route.

3. Approach

The general idea of this approach is to schematize the subdivision by isolating the boundaries between categorically different areas and extract them as polylines. A cubic BC is fitted to each polyline separately which is then projected to the already schematized new outline. The order of insertion of the fitted BCs is crucial for preserving topology. This insertion order is obtained by data preparation and inclusion of some ancillary data and modelled via different classes of anchor points. This general strategy is illustrated by an example of a categorical map of Italy's regional differences in natural and migratory population change.

3.1 Data preparation

The provided example deals with the cartographical schematization. Thus we assume the most important parts of model generalisation to have happened already. Using EUROSTAT/ESPON data on the level of NUTS-3 regions (Figure 2a), these regions were retrieved based on their migratory type. This migratory type already is a strong model generalization provided by the EU as a regional planning device. Administrative enclaves were eliminated and the dataset projected to cartesic UTM coordinates. All islands below a threshold of the smallest mainland NUTS-3 region were eliminated. The topological holes created by the Vatican and San Marino were kept at this time with their centroids later to be projected into the schematized outline as anchors for point symbolization (Figure 2b).

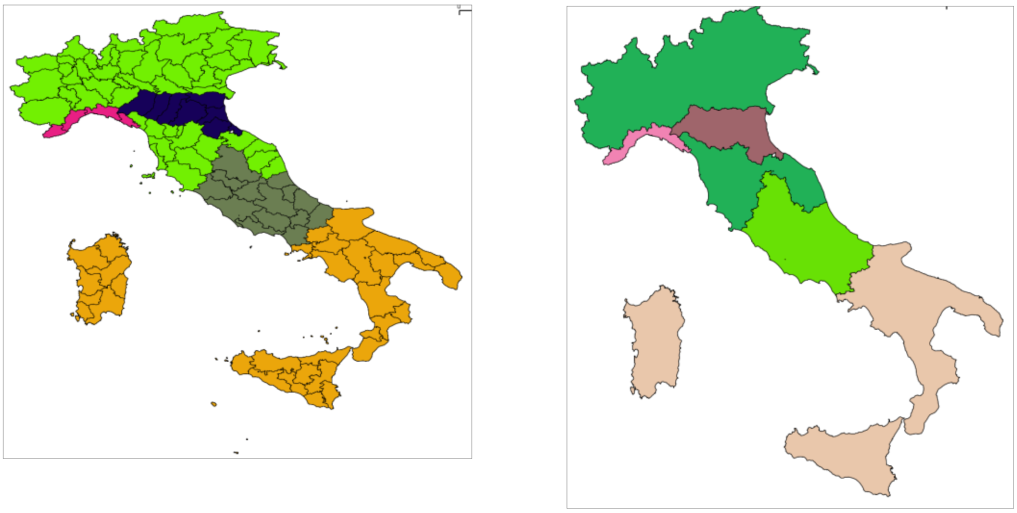

Figure 2. a) Original data on NUTS-3 level b) prepared data for input into the schematization process.

3.2 Anchor points for topology preservation

For approaches that schematize a whole categorical map at the same time, traditional topology preservation techniques and algorithms can be applied during the process. As our approach treats the territorial outline and interior subdivision differently, we use so-called anchor points that inform the reinsertion of schematized boundary lines.

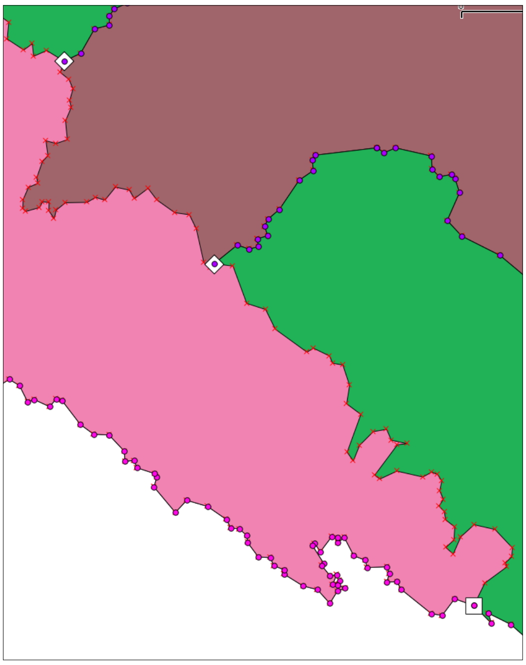

Figure 3. First (square) and second (diamond) order internal anchor points.

In our example, we have two degrees of anchor points that inform the internal schematization (Figure 3). First order internal anchor points are those points that lie on the original territorial outline and are topologically connected to two categorically different areas. Second order anchor points are those points that are connected to three categorically different areas but not the original territorial outline. Assuming the reinserted schematized boundaries do not intersect with outline, themselves or other boundaries, these anchor points ensure that the internal topology is preserved.

Keeping the topology consistent for other neighborhood considerations needs the introduction of ancillary data. For the depiction of countries, it has been observed that the points where a national boundary and an ocean meet are interpreted as being crucial for recognition (Reimer and Meulemans 2011). For the example of Italy, we considered the adjacency to other nations and surrounding areas of the Mediterranean as depicted in Figure 4.

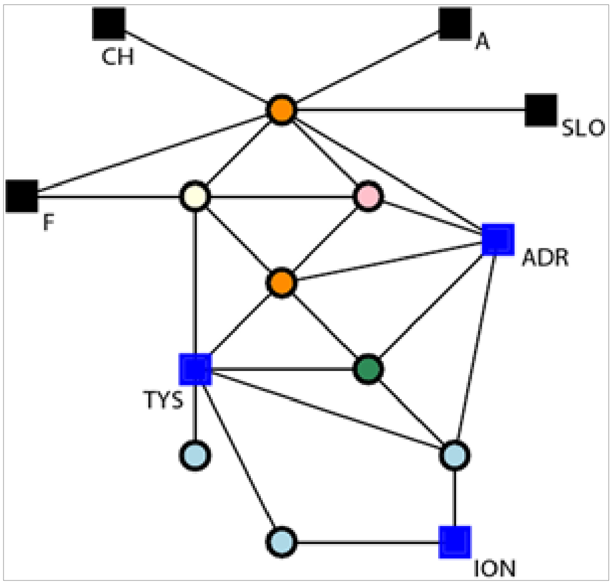

Figure 4. Adjacency graph with ancillary relations. Sea areas are marked as blue squares, outside nations as black squares.

The ancillary information is used here to keep outside relations intact. Every relation depicted in Figure 3 must be assigned to an edge on the schematized outline *O*. Should a characteristic point on the original outline precede an ancillary anchor point, the internal anchor point of the first order must be reassigned to the preceding edge that is now carrier of that external relation. In the example this is the case for the area of Type 3 i.e. Liguria.
The internal anchor points of the first order are projected onto the schematized outline in the following manner (figure 5).

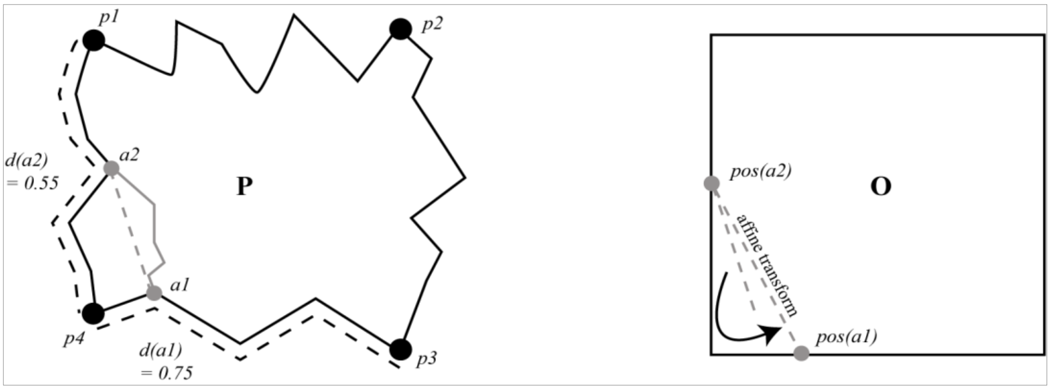

Figure 5. Illustration of the projection process realizing Algorithm step 6.b)

3.3 Polyline schematization with cubic Bezier curves

In this section we describe the algorithm for the schematization, give some explanations and analyse its running time.

Algorithm Schematize (Subdivision *S*)
Input: subdivision *S* that is one connected component with first-order anchor points only
Output: schematization of *S* with cubic Bezier curves for internal boundaries

1. Obtain outline polygon *P* and interior polylines C_1 ... C_n from *S*
2. Shift the indices of *P* such that p_1 (its first vertex) is a diametrical point (that is a point which realizes a Furthest Pair of P)
3. **For each** first order anchor point *a* do
   a. Compute clockwise distance *d(a)* from p_1 to *a* along the boundary of *P*
4. Let *O* be schematized outline of P produced by Reimer-Meulemans algorithm
5. **For each** first order anchor point *a* do
   a. Compute position *pos(a)* along boundary of *O* that has

distance *d(a)* * |*O*| / |*P*| from matched vertex

6. For each polyline C_i do
   a. Fit a Cubic Bezier curve BC_i
   b. Transform BC_i to match $pos(a_1)$ and $pos(a_2)$ where a_1 and a_2 are the first-order anchor points at the ends of C_i
   c. While BC_i intersects a BC or the outline **do**
      c.i. Modify BC_i to move away from intersection
7. **Return** outline O_s and BC_1... BC_n

Explanation and running time analysis

Let |P|=M, $|C_1|+|C_2|+ \ldots +|C_n|=N$ (and hence it is |S|=M+N). Furthermore let |O|=M' with M' ≤ M.

1. Can be performed in O(M+N)
2. Can be performed in O(M logM). As is easily seen a diametrical point has to be in the Convex Hull of P, hence we compute the convex hull of P in O(M log M) first. Given the Convex Hull we can find a diametrical point by a rotating calipers algorithm in O(M) time as done by Shamos (1978).
3. Can be performed in O(M) by scanning P once.
4. Running time for Reimer-Meulemans is $O(M'^2 \cdot K)$, where K is the number of annealing iterations; in our example fixed and thereby constant.
5. Can be performed in O(M+M'). Place these first order anchor points a (in order along P) by scanning over O once. We would like to note the following: This mapping works well when the schematized algorithm preserves relative distances on the outline. While in our exsperiments the mapping results were satisfying, more involved mapping algorithms like Locally Correct Fréchet Matchings by Buchin et al. (2012) could be used.
6. Will be excluded from the running time analysis as it depends on the details of the BC fitting algorithm. Especially 6. c) might result in unbounded running time. Here further investigation is needed. Remark: It is sufficient to guarantee that the control polygons do not intersect with each other (or the outline).

The total running time excluding 6. is $O(M \log M + M'^2 \cdot K + N)$.

Detailed explanation of step 6

The extracted polylines are schematized by redrawing them with a fitted cubic Bezier curves. For our example we used a modified version of the fitting algorithm proposed by Masood and Ejaz (2010) using a part-wise directed Hausdorff distance as error measure for each control point. The underlying idea of the curve fitting algorithm is to work from starting positions for the control points determined from the polyline to be fitted and then do incremental adjustments to improve their positions. Starting and end points of the polyline to be fitted are the starting and end points of the cubic Bezier curves. By measuring the maximum and minimum straight distance from the polyline to the baseline connecting start and end points, we get values d_1 and d_2. The starting positions of control points are obtained by extrapolating the line from the baseline to two times the distance d (Figure 6).

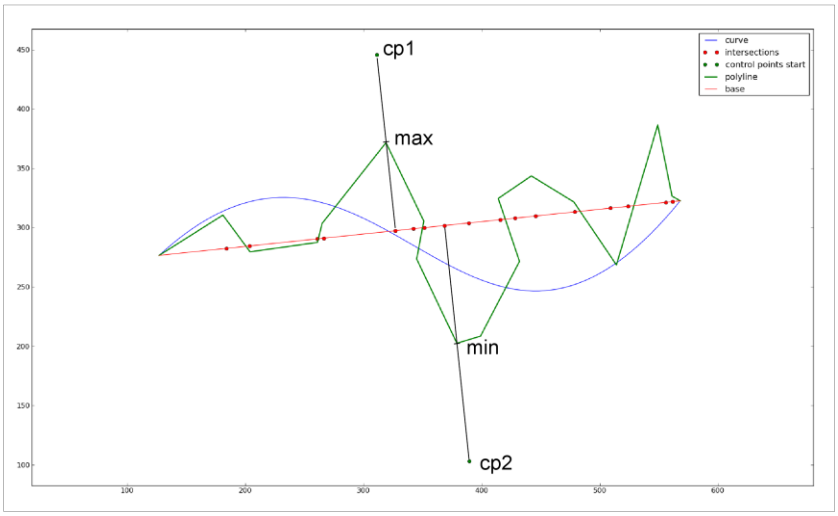

Figure 6. Starting point selection for the control points.

This BC is then scaled to the projected APs and translated with the projected APs as new starting and endpoint via affine transformation. The BC is then checked for intersections with itself, the outline and BCs already inserted. Using the pseudo-local control property of BCs, the control point closest to the intersection is moved stepwise until the intersection is remedied and minimal distances between graphical objects is reached.

4. Results

The final schematization result is depicted in Figure 7. The internal topological situation is the same one as with the input data (Figure 8). In the case of the Type 3 (‘peripheral areas characterised by a very neutral age profile’ depicted in tan) region, which is coterminous with the landscape of Liguria, the implicitly expected neighbourhood to France as well as the sea is kept. The colours correspond to the Type-colour pairings of the ESPON-Atlas. Table 1 shows the absolute and relative areas of output and input regions, ignoring the Vatican and San Marino. The relative areas are provided as percentage of the whole depicted land mass respectively.

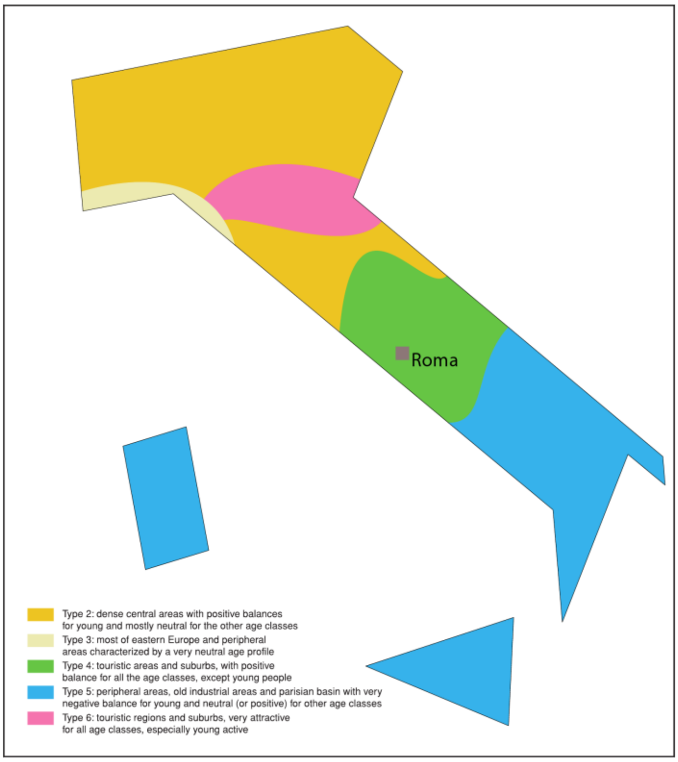

Figure 7. Schematization result.

We note that at least for this example, the relative area sizes changes are all well below 2% with a mean of 0.80% and a standard deviation of 0.46%. For a schematization that is by definition taking great liberties with geometry, and which does not explicitly treat area preservation as constraint, this is remarkably faithful to the original graphical structure.

Table 1. Comparison between output and input area sizes.

Region	Input [km²]	Input [%]	Output [km²]	Output [%]	Δ [km²]	Δ [%]
Type 2	127436.60	42.28	117774.09	41.85	-9661.93	-0.44
Type 3	5411.55	1.80	6387.69	2.27	-976.14	0.47
Type 4	43649.60	14.48	41707.70	14.82	-1941.90	0.34
Type 5	55825.61	18,52	54366.45	19.32	-1459.16	0.80
Type 6	19543.16	6.48	21634.03	7.69	2090.87	1.20
Sicily	25576.41	8.49	19336.80	6.87	-6239.61	-1.62
Sardinia	23938.97	7.94	20235.75	7.19	-3703.22	-0.75
Total	301381.90	100	281443.09	100	-29934.79	0

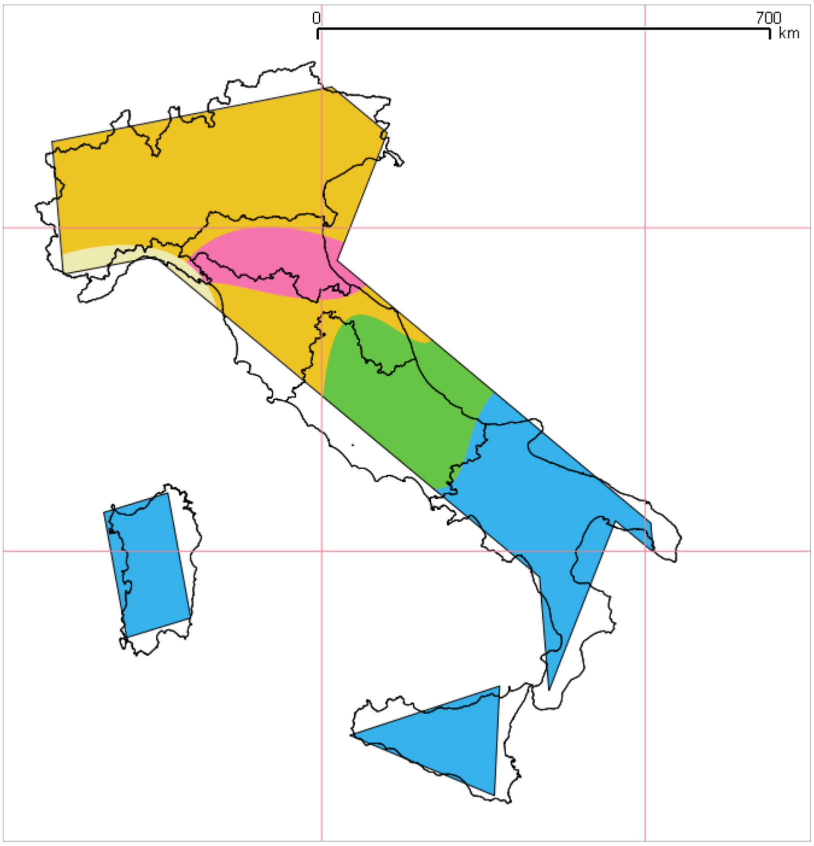

Figure 8. Schematization result overlaid with input geometries.

It has been shown that the approach is able to automatically produce a schematization with BCs that keeps all functional geometric relations while visually highlighting its nature as schematization. For the data used, the computation was fast as there were only two conflicts that needed to be resolved. The convex hull property of BCs states that a BC is contained within the convex hull of its control polygon, i. e. if the control polygons do not overlap, the resulting BCs will be intersection free. This was used to inform the insertion strategy, inserting all BCs with same degree APs without conflicts first and then placing the burden of conflict resolution on the shortest remaining segments, where changes do impact the overall result less. This can be interpreted as that the proposed anchor point/insertion order strategy is able to eliminate many potential conflicts before they even arise. Further development can concentrate on coping with more complex conflicts by such a divide and conquer strategy. We see such a first success as indicator that BCs can be fruitfully used to visually and geometrically characterize subdivisions for categorical maps in automated mapping.

5. Future Work

As has been alluded to above, the insertion strategies above work well with anchor points of low order. For more involved decompositions, i. e. with polygons that have no connection to the outline and consequently with higher order anchor points, strategies are being tested. More generally, the horizontal relations between BCs are going to be addressed via introduction of new similarity measures. Progress in this area will allow to widen BC usage for mixed polygonal-curve outlines as well as tackling gestalt-driven considerations for categorical schematizations and generalizations.

Acknowledgements

This research has been funded the METRIK research training school and the University of Heidelberg and was supported by the GFZ Potsdam. C. Volk was supported by the Netherlands Organisation for Scientific Research (NWO) under project no.~612.001.102. Thanks go to Wouter Meulemans, Herman Haverkort and Kevin Buchin from TU Eindhoven for ongoing discussions about Bezier curves in schematizations.

References

Buchin K, Buchin M, Meulemans W and Speckmann, B, 2012, Locally Correct Fréchet Matchings, In *Proc. 20th European Symposium on Algorithms (ESA)* Springer.

Bucher H and Schlömer C, 2006, Die neue Raumordnungsprognose des BBR, *Raumforschung und Raumordnung* 64 (2006) 3, S. 206-212.

ESPON Atlas, 2006, Mapping the structure of the European territory, edited by the federal office for buidling and regional planning

Foley J, Van Dam S, Feiner S and Hughes J, 1990, *Computer Graphics: Principles and Practice*, Addison-Wesley.

Galanda M., 2003, Automated Polygon Generalization in a Multi Agent System. *PhD thesis*, University of Switzerland.

Mark D and Csillag F,1989 The Nature of Boundaries on `Area-Class' Maps. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 26, 65–77.

Masood A and Ejaz S, 2010, An Efficient Algorithm for Robust Curve Fitting Using Cubic Bezier Curves. *Proceedings of the 6th International Conference on Intelligent Computing, LNAI 6216 (2010)* , 255–262.

Reimer A, 2010, Understanding Chorematic Diagrams: Towards a Taxonomy, *The Cartographic Journal*, Vol. 47 Nr. 4 pp. 330-350.

Reimer, A and Meulemans, W, 2011, Parallelity in Chorematic Territorial Outlines, *Geographic Information on Demand: 14th of the ICA commission on Generalisation and Multiple Representation* , Paris.

Roberts, M, 2012, *Underground Maps Unravelled - Explorations in Information Design*, London.

Saalfeld A, 2001, Area-preserving Piecewise Affine Mappings. *Proceedings of the seventeenth annual symposium on Computational geometry SCG 01 (2001)* , 90–95.

Shamos, M, 1978, Computational Geometry, *PhD thesis*, Yale University.

Shao L and Zhou H, 1996, Curve Fitting with Bezier Cubics, *Graphical models and image processing, Vol. 58, N. 3*, 223-232.

Steiniger S and Weibel R, 2005, Relations and structures in categorical maps, *8th ICA Workshop on generalisation and Multiple Representation, A Coruna (Spain)*, July 7-8th, 2005.