Area-preserving C-oriented schematization

Area-Preserving $\mathcal{C}$-Oriented Schematization*

Kevin Buchin† Wouter Meulemans† Bettina Speckmann†

Abstract

We define an *edge-move* operation for polygons and prove that every simple non-convex polygon P has a non-conflicting pair of complementary edge-moves that reduces the number of edges of P while preserving its area. We use this result to generate area-preserving $\mathcal{C}$-oriented schematizations of polygons.

1 Introduction

A schematic map displays a set of nodes and their connections—for example, highway, train, or metro networks—in a highly simplified form to communicate the connectivity information as effectively as possible. Connections are usually drawn as polygonal paths using few links and few orientations. The set of permissible orientations often contains only the two axis-parallel or the four main orientations. The most general setting is $\mathcal{C}$-oriented schematization where every link has to use one of a set $\mathcal{C}$ of specified orientations.

A substantial part of previous efforts concentrates on the schematization of networks, but it is also often desirable to schematize the boundaries of regions or even complete subdivisions. Whenever exact boundaries are not needed it is preferable to replace them by schematic ones, to reduce visual clutter and to indicate that the map is not a topographic map.

The schematized output should visually resemble the input. But it is not clear how to quantify what the most recognizable schematization of a given input is. Optimization based on standard distance metrics can produce undesirable output for certain input polygons [8]. Hence we focus on area-preserving schematization, that is, the area of the input polygon and its schematization are equivalent. While we do not prove any guarantees on the resulting shapes, experimental results are quite promising and visually pleasing.

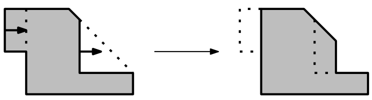

Figure 1: Complementary edge-moves.

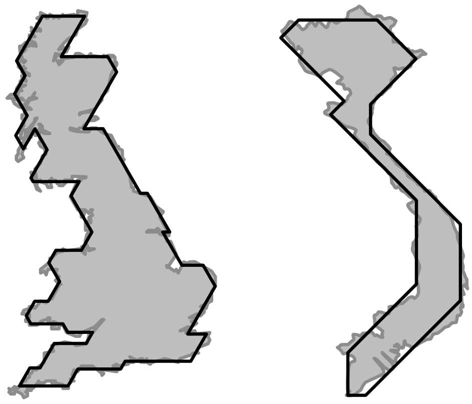

Figure 2: Hexagonal Great Britain using 50 edges; octagonal Vietnam using 15 edges.

Results. We introduce an *edge-move* operation that moves an edge of a polygon inward or outward without changing its orientation. This does change the length of its adjacent edges and potentially itself. A contraction is an edge-move that reduces at least one edge to length zero, hereby reducing the number of edges in the polygon. Since we desire area preservation, we apply this operation in non-conflicting complementary pairs (Figure 1). Theorem 1, our main result, is an immediate consequence of Lemma 6:

Theorem 1 *Every simple non-convex polygon has a non-conflicting pair of complementary edge-moves, one of which is a contraction.*

As edge-moves do not introduce new orientations, we can also use edge-moves to create area-preserving $\mathcal{C}$-oriented schematizations of polygons. Note that a simple polygon can be converted into a simple $\mathcal{C}$-oriented polygon of equal area based on the method presented by Meulemans *et al.* [8].

Corollary 2 *Given a simple $\mathcal{C}$-oriented polygon P and an integer k with $2|\mathcal{C}| \leq k$, an area-preserving $\mathcal{C}$-oriented schematization of P with at most k edges can be generated using only non-conflicting pairs of complementary edge-moves.*

The results in Figure 2 were obtained by repeatedly executing a non-conflicting pair of complementary edge-moves such that the contraction minimizes the area change and the compensating edge-move is as close by as possible. This method can also be used for simple subdivisions restricted to edge-moves that do

*W. Meulemans and B. Speckmann are supported by the Netherlands Organisation for Scientific Research (NWO) under project no. 639.022.707.

†Department of Mathematics and Computer Science, TU Eindhoven, The Netherlands. k.a.buchin@tue.nl, w.meulemans@tue.nl, and speckman@win.tue.nl.

not change the topology. Our algorithm then ensures that each face preserves its area and that adjacencies are correct. In this case we cannot give any guarantees on the reduction in the number of edges. Preliminary experiments suggest that the reduction is significant.

Related work. There is an ample body of work on map schematization and metro map construction, see the surveys by Swan *et al.* [10] and Wolff [11] for an overview. Of particular interest is the work by Merrick and Gudmundsson [7] who describe an algorithm to generate $\mathcal{C}$-oriented metro map layouts. Methods for map schematization can be used to schematize subdivisions, but they usually do not take criteria such as shape and size preservation into account.

The generalization of urban data, specifically building generalization, is closely related to our work. In particular, algorithms for building wall squaring [6] and outline simplification [5] can be used for polygon schematization and vice versa.

Line simplification has been a prominent topic in the GIS literature for many years. Of particular relevance are the work of Delling *et al.* [3] and Neyer [9] dealing with $\mathcal{C}$-oriented schematization of routes or lines and the work of Bose *et al.* [1] on area-preserving line simplification. However, it is generally not advisable to schematize each chain in a subdivision separately. There are some approaches, developed in computational geometry, that preserve the topology of the input subdivision. For example, De Berg *et al.* [2] describe a method that simplifies a polygonal subdivision without introducing intersections or passing over special input points. Unfortunately many subdivision simplification problems that minimize the number of edges in the output are NP-complete [4].

2 Edge-moves and configurations

Definitions and notation. We are given a simple polygon P with vertices $v_1, \ldots, v_n$. We treat the vertices modulo n, e.g. $v_{n+1} = v_1$. The edges are denoted by $e_1, \ldots, e_n$, again treated circularly. The directed edge e_i starts at vertex v_i and ends at vertex v_{i+1}. A vertex is called convex if the angle inside the polygon between its two adjacent edges is at most π, and it is called reflex otherwise. We call an edge convex or reflex if both its vertices are convex or reflex respectively. The exterior angle of a vertex is defined as the angle between one edge and the extension of the other. The exterior angle is sometimes also referred to as turning angle. The angle is negative if and only if the vertex is reflex. The sum of all exterior angles of a simple polygon is always equal to 2π.

We define a chain S as a set of at least three consecutive edges of P. Its edges are denoted by $s_1, \ldots, s_m$ with $m \leq n$. Its vertices are denoted by $u_1, \ldots, u_{m+1}$ and edge s_i is directed from u_i to u_{i+1}. The edges

Figure 3: A positive and negative exterior angle.

s_1 and s_m are the *outer edges* of S, the other edges are its *inner edges*. Likewise, u_1 and u_{m+1} are outer vertices, the other vertices inner vertices. By $\alpha(S)$, we denote the sum of the exterior angles of the inner vertices of S. A *lid* is an open line segment between a point on s_1 and a point on s_m and is fully contained in the interior of P. If S has any lid, it is a *closable chain*. If the open line segment (u_1, u_{m+1}) is a lid, S is a *proper chain*. For a closable chain S and a lid l, we denote by $R_l(S)$ the region enclosed by S and l. For a proper chain, $R(S)$ denotes this region using the lid (u_1, u_{m+1}) implicitly. Due to the lid, we know that for every closable chain, any point on the boundary of P that is inside $R_l(S)$ must be part of S. For any closable chain, $\alpha(S) > 0$ holds. A (not necessarily closable) chain with exactly 3 edges is called a *configuration* G, its edges denoted by g_1, g_2, and g_3.

The outer edges of a configuration G define two tracks, infinite lines through the edges. An *edge-move* on G moves g_2 such that its orientation is preserved and its vertices are on the tracks, making the outer edges longer or shorter. An edge-move is valid if at least one of its vertices remains on its original outer edge and g_2 remains on the same side of or on the intersection point of the tracks (if any). An edge-move which causes one of the edges of G to reach length zero, is a *contraction*. Contractions are extremal edge-moves. An edge-move is *positive* if it adds area to P and *negative* if it removes area.

A configuration supports edge-moves, either positive, negative or both. Let g_2^+ denote the extremal position of g_2 after any valid positive edge-move, i.e. the position after a positive contraction. The *positive contraction region* of a positive configuration, $R^+(G)$, is the region enclosed by g_2, g_2^+, and the tracks. A *feasible positive configuration* is a configuration for which $R^+(G)$ is empty except for G. Similarly, we define the *negative contraction region* $R^-(G)$ and a feasible negative configuration. If a positive or negative configuration is feasible, then any valid positive or negative edge-move respectively is feasible. If a positive configuration is infeasible, then there is some point on δP in $R^+(G) \backslash G$. A point in $\delta P \cap R^+(G) \backslash G$ that is closest to

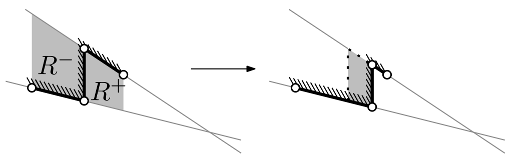

Figure 4: A valid positive edge-move.

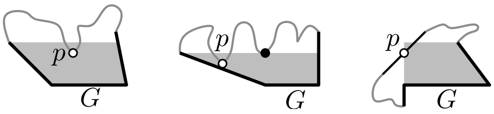

Figure 5: Configurations G and blocking points p.

the line through g_2 is called a *positive blocking point*. Analogously, G can have a *negative blocking point*. If the inner edge of an infeasible G is convex or reflex, there is always a blocking point that is a vertex of P.

Since we desire an approach that preserves the area of the polygon, we combine two *complementary* feasible configurations, one positive and one negative, executing an edge-move on both simultaneously. The one with the smaller contraction region is contracted, while the other is moved just far enough to compensate for the area change. Two configurations *conflict* when they share an edge, unless they share only outer edges and one of these has a convex and a reflex vertex. In this special case the two edge-moves both either shorten or lengthen the shared edge. We call two non-conflicting complementary feasible configurations a *proper configuration pair*.

Completeness. We now prove that any non-convex polygon has a proper configuration pair. First we discuss some properties of a closable or proper chain.

Lemma 3 *If S is a closable chain without a convex inner edge and s_2 is reflex, then S has a feasible negative configuration G with a reflex first inner vertex, $\alpha(G) > 0$ and $R^-(G) \subseteq R_l(S)$ for any lid l of S.*

Proof. As $\alpha(S) > 0$ and there are no convex inner edges, there must a configuration G' in S with a reflex first inner vertex and $\alpha(G') > 0$ (implying that the second inner vertex is convex). Let G' be the first such configuration. A configuration G'' with $\alpha(G'') > 0$ and a convex first inner vertex may occur multiple times before G', but it must always be preceded by an edge g''_0 with $\alpha(\{g''_0\} \cup G'') < 0$ for the chain $\{g''_0\} \cup G''$ as otherwise G' is not the first of its kind along S.

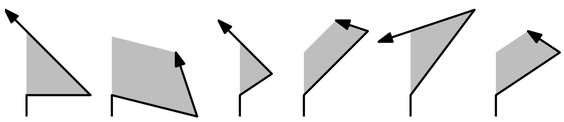

Figure 6: Configurations with $\alpha > 0$ and a reflex first inner vertex.

We prove this lemma by induction. Assume that $m = 3$. Chain S is a configuration G' with a reflex first inner vertex. Any lid l of S must be on one side of the line through s_1, whereas this track enforces a triangular contraction region on the other side. Hence, $R^-(G') \subseteq R_l(S)$ holds and G' is a feasible negative configuration. Examples 1, 3 and 5 in Figure 6 show closable chains, the other examples do not.

For induction, assume that this lemma holds for any closable chain with less than m edges and with a reflex first inner vertex. Let $G' = s_{i-1}s_is_{i+1}$ be the first configuration with $\alpha(G') > 0$ where u_i is reflex. If G' is feasible, we are done. If G' is not feasible, let p denote the corresponding blocking point. Note that p must be part of S as $R^-(G') \subseteq R_l(S)$ for any lid l. Moreover, p must come after s_{i+1} on S as otherwise G' cannot be the first. The closable chain $S' = s_{i+1}, \ldots, p$ (that is, until the edge containing p if p is not a vertex) has less edges than S and thus has a feasible negative configuration by induction. □

Lemma 4 *If a proper chain S has a convex inner edge, then S has a feasible negative configuration G with $R^-(G) \subseteq R(S)$ and $\alpha(G) > 0$. Also, g_2 is convex or starts at a reflex vertex or $g_2 \neq s_{m-1}$.*

Proof. We prove this lemma by induction on m. If $m = 3$, S is a feasible negative configuration, since S is a proper chain and s_2 is a convex edge. For induction, assume that this lemma holds for any suitable chain with less than m edges. Let s_i be a convex inner edge. Hence, $G' = s_{i-1}s_is_{i+1}$ is a negative configuration. If G' is feasible, we are done as a convex edge implies $\alpha(G') > 0$. If G' is not feasible, then let u_j denote the blocking vertex and assume without loss of generality that $i + 1 < j$. Note that u_j must be a vertex of the chain as $R^-(G') \subseteq R(S)$. Consider the proper chain $S' = s_i, \ldots, s_{j-1}$. If S' has a convex inner edge, it must have a feasible negative configuration by induction. However, if it does not, $S'' = s_{i+1}, \ldots, s_{j-1}$ is a closable chain in which s'_2 is reflex. Note that the direction of S'' is reversed when $j < i-1$. By Lemma 3, S'' has a feasible negative configuration G''. By the lemma, the first inner vertex of G'' is reflex in the direction of S''. If S'' was reversed, then $s_m \notin S''$, thus s_{m-1} cannot be the inner edge of G''. □

Lemma 5 *Every simple non-convex polygon P has a feasible positive configuration G with $\alpha(G) < 0$ or all positive configurations are feasible.*

Proof. If P has a reflex edge, then let G be a configuration with a reflex inner edge. If G is feasible, we are done as $\alpha(G) < 0$. If it is not, we can define a chain $\overline{S}$ that is inverted: the interior of $\overline{S}$ is in fact the exterior of the polygon. Using Lemma 3 or Lemma 4, we can now find a feasible negative configuration G^- with $\alpha(G^-) > 0$ in $\overline{S}$. This corresponds to a feasible positive configuration G^+ in P with $\alpha(G^+) < 0$.

If P has no reflex edge, then let G be an infeasible positive configuration. If none exists, all positive configurations are feasible. Otherwise we can define an inverted chain using G. Thus, by Lemma 3, P has a feasible positive configuration G' with $\alpha(G') < 0$. □

Lemma 6 *Every simple non-convex polygon P has a proper configuration pair.*

Proof. By Lemma 5, polygon P has a feasible positive configuration $G^+ = e_{i-1}e_ie_{i+1}$. Assume without loss of generality that the second inner vertex of G^+, v_{i+1}, is reflex. Let v_j denote the first convex vertex after v_{i+1}. Configuration $G^- = e_{j-1}e_je_{j+1}$ of which v_j is the first inner vertex, is negative. We distinguish two cases.

Assume that G^- is feasible. If no edge is shared or if edge $e_{i+1} = e_{j-1}$ is shared (having a convex and a reflex vertex), we have a proper configuration pair. If $e_{i-1} = e_{j+1}$ is shared but the other outer edge is not, then $v_j, v_{j+1} = v_{i-1}, v_i$ are the only convex vertices in P and there is at least one edge in between e_i and e_{j-1}. This edge is the inner edge of a feasible positive configuration, one that does not conflict with G^-.

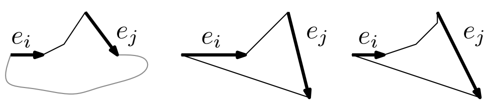

Figure 7: Three cases if G^- is feasible.

Now, assume that G^- is not feasible. The blocking point cannot be in between v_i and v_{j+1}. If G^- is blocked by a vertex v_h, then, depending on the convexity of v_{j+1} and v_{j+2}, either Lemma 3 or Lemma 4 shows that there is a (non-conflicting) feasible negative configuration. If G^- is blocked by an edge e_h, v_{j+1} must be reflex. We distinguish two cases on the closable chain $S = e_j, \ldots, e_h$.

If S does not have a convex inner edge, then we refer to Lemma 3 to find a feasible negative configuration G'. If G' does not conflict with G^+, we are done. If G' does conflict with G^+, we know that $e_h = e_{i-1}$ holds and that e_{h-1} is the inner edge of G'. Moreover, it must now hold that $\alpha(G^+) > 0$ and thus, by Lemma 5, we need to consider this case only when all positive configurations are feasible. Hence, the positive configuration $e_ie_{i+1}e_{i+2}$ is feasible and it does not conflict with G'.

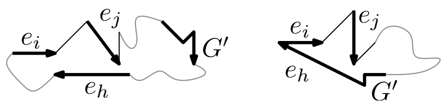

Figure 8: Two cases if G^- is not feasible and S has no convex inner edge. G' may conflict with G^+.

If S has a convex inner edge e_t, then let G' denote the negative configuration $e_{t-1}e_te_{t+1}$. If G' is feasible and not conflicting with G^+, we are done. If G' is feasible but conflicting with G^+, we can argue as above: $e_ie_{i+1}e_{i+2}$ is a feasible positive configuration and it does not conflict with G'. If G' is not feasible, then it must be blocked by some vertex v_b. Without loss of generality, assume that the proper chain $S' = e_t, e_{t+1}, \ldots, e_{b-1}$ does not contain edge e_i. Depending on the convexity of vertex v_{t+2}, Lemma 3 or Lemma 4 shows that there is a feasible negative configuration G'' in S'. We now argue why G'' cannot conflict with G^+. The only way to have a conflict is when $v_b = v_i$. Since v_i is then reflex, the only way to obtain a conflict is when v_{i-1} is reflex as well and v_{i-2} is convex, such that $e_{i-3}e_{i-2}e_{i-1}$ is a negative configuration. However, e_{i-2} is the before-last edge in S' and starts at a convex vertex. Hence, Lemma 3 and Lemma 4 guarantee to find another feasible negative configuration instead. □

References

[1] P. Bose, S. Cabello, O. Cheong, J. Gudmundsson, M. van Kreveld, and B. Speckmann. Area-preserving approximations of polygonal paths. *Journal of Discrete Algorithms*, 4(4):554–566, 2006.

[2] M. de Berg, M. van Kreveld, and S. Schirra. Topologically correct subdivision simplification using the bandwidth criterion. *Cartography and Geographic Information Science*, 25(4):243–257, 1998.

[3] D. Delling, A. Gemsa, M. Nöllenburg, and T. Pajor. Path schematization for route sketches. In *Proc. 12th Scandinavian Workshop on Algorithm Theory*, LNCS 6139, pages 285–296, 2010.

[4] L. Guibas, J. Hershberger, J. Mitchell, and J. Snoeyink. Approximating polygons and subdivisions with minimum-link paths. *International Journal of Computational Geometry and Applications*, 3:383–415, 1993.

[5] J. Haunert and A. Wolff. Optimal and topologically safe simplification of building footprints. In *Proc. 18th International Conference on Advances in GIS*, pages 192–201, 2010.

[6] H. Mayer. Scale-spaces for generalization of 3d buildings. *International Journal of Geographical Information Science*, 19:975–997, 2005.

[7] D. Merrick and J. Gudmundsson. Path simplification for metro map layout. In *Proc. 14th International Symposium on Graph Drawing*, LNCS 4372, pages 258–269, 2006.

[8] W. Meulemans, A. van Renssen, and B. Speckmann. Area-preserving subdivision schematization. In *Proc. 6th International Conference on GIScience*, LNCS 6292, pages 160–174, 2010.

[9] G. Neyer. Line simplification with restricted orientations. In *Proc. 6th International Workshop on Algorithms and Data Structures*, LNCS 1663, pages 13–24, 1999.

[10] J. Swan, S. Anand, M. Ware, and M. Jackson. Automated schematization for web service applications. In *Web and Wireless Geographical Information Systems*, LNCS 4857, pages 216–226, 2007.

[11] A. Wolff. Drawing subway maps: A survey. *Informatik-Forschung und Entwicklung*, 22(1):23–44, 2007.