Exact Solutions for Area-Optimal Simple Polygonization Problems

Natanael Ramos, Pedro J. de Rezende and Cid C. de Souza ({nramos,rezende,cid}@ic.unicamp.br)
Institute of Computing, UNICAMP, Campinas, Brazil

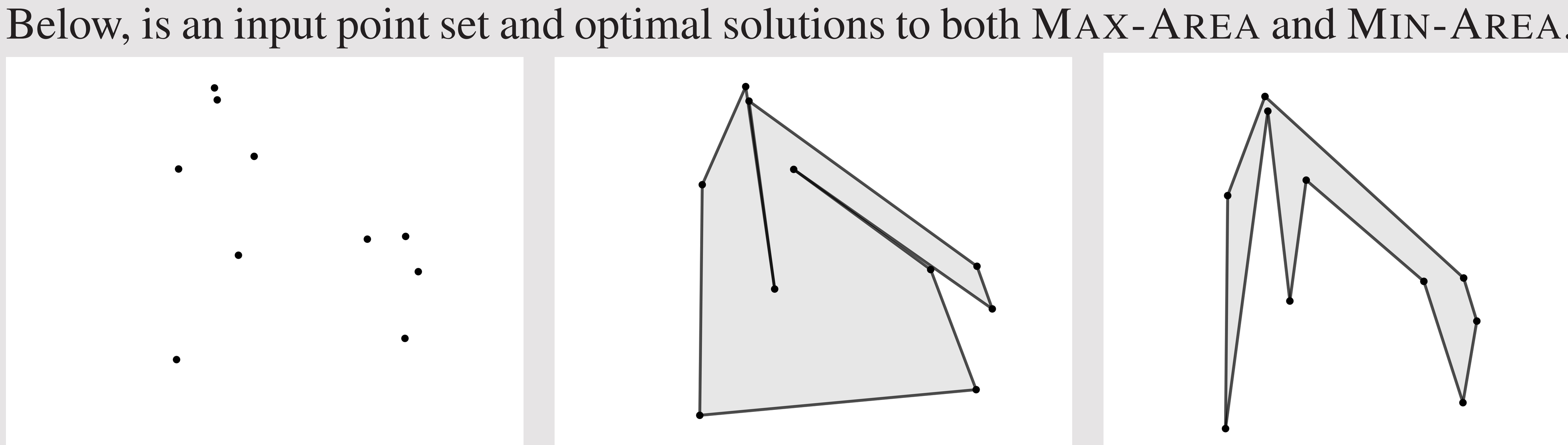
This work was supported in part by Grants from CAPES, Brazil – Finance Code 001; CNPq, Brazil: #140300/2020-1, #309627/2017-6, #313329/2020-6; and FAPESP, Brazil: #2014/12236-1, #2018/26434-0, #2020/09691-0;
Keywords: Computational Geometry; Combinatorial Optimization; Integer Programming; Polygonization.

Problem Statements

Simple Polygonization Given a set S of points in the plane, a *simple polygonization* of S consists of a simple polygon whose set of vertices is precisely S.

When computing such structures, one could want to optimize some given function, like the area of the resulting polygonization. In this vein, we have the *Area-Optimal Simple Polygonization Problems*, in which, given a set of points S we wish to find a simple polygonization with maximum (MAX-AREA) or minimum (MIN-AREA) area.

Below, is an input point set and optimal solutions to both MAX-AREA and MIN-AREA.

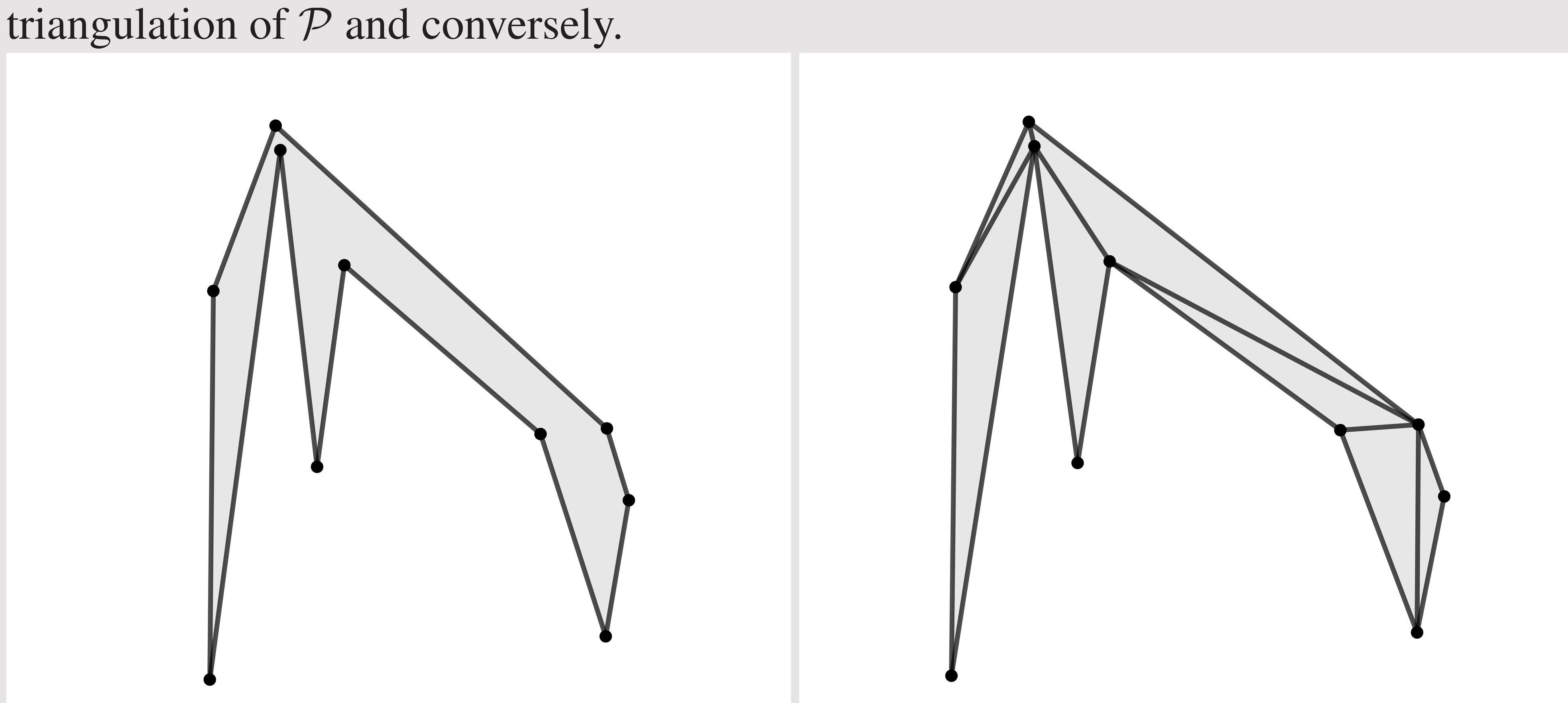


Such problems come up in areas such as Pattern Recognition, Image Reconstruction and Clustering [1–3, 5]. Both problems were proved to be $\mathbb{NP}$-complete and there are exact, heuristic and approximation algorithms for them in the literature.

Here, we seek to find *exact* solutions to both problems, MAX-AREA and MIN-AREA, through the resolution of a novel Integer Linear Programming (ILP) formulation.

Polygon Triangulation

Our model relies on the fact that, given a simple polygon $\mathcal{P}$ of n vertices, $V(\mathcal{P})$, any set of $n-2$ interior-disjoint triangles in $\mathcal{P}$, having vertices in $V(\mathcal{P})$, constitutes a triangulation of $\mathcal{P}$ and conversely.



References

[1] E. Althaus and K. Mehlhorn. "Traveling Salesman-Based Curve Reconstruction in Polynomial Time". In: *SIAM Journal on Computing* 31.1 (2001), pp. 27–66.

[2] T. K. Dey, K. Mehlhorn, and E. A. Ramos. "Curve reconstruction: Connecting dots with good reason". In: *Computational Geometry* 15.4 (2000), pp. 229–244.

[3] S. P. Fekete. "On Simple Polygonalizations with Optimal Area". In: *Discrete & Computational Geometry* 23.1 (2000), pp. 73–110.

[4] S. P. Fekete et al. "Computing Area-Optimal Simple Polygonalizations". In: *Journal of Experimental Algorithmics* (2020), To appear.

[5] J. Giesen. "Curve Reconstruction, the Traveling Salesman Problem, and Menger's Theorem on Length". In: *Discrete & Computational Geometry* 24.4 (2000), pp. 577–603.

ILP Formulation

Denote a segment between points $i, j \in S$ by ij and a triangle by ijk when its vertices are $i, j, k \in S$. The set of all segments between points of S is indicated by $E(S)$, and the set of all empty triangles in S by $\Delta(S)$. Also, $\mathcal{X}(E(S))$ is the set of pairs of segments that cross each other. The area of a given triangle ijk is written $\mathcal{A}(ijk)$. Given a subset of points $U \subset S$, $\delta(U)$ is the set $\{ij : i \in U \text{ and } j \in S \setminus U\}$.

We employ two sets of binary variables $\{x_{ij}, ij \in E(S)\}$ and $\{t_{ijk}, ijk \in \Delta(S)\}$.

(max/min total area) $$\max / \min \sum_{ijk \in \Delta(S)} \mathcal{A}(ijk) t_{ijk}$$

(# triangles) $$\sum_{ijk \in \Delta(S)} t_{ijk} = n - 2,$$

(no crossings) $$x_{ij} + x_{kl} \leq 1, \ \forall (ij, kl) \in \mathcal{X}(E(S))$$

(x- and t-var. binding) $$x_{ij} + x_{ik} + x_{jk} - t_{ijk} \leq 2, \ \forall ijk \in \Delta(S)$$

(x- and t-var. binding) $$x_{ij} \leq \sum_{k \in S \setminus \{i,j\}} t_{ijk} \leq 2x_{ij}, \ \forall ijk \in \Delta(S)$$

(connectivity) $$\sum_{ij \in \delta(U)} x_{ij} \geq 2, \ \forall U \subset S$$

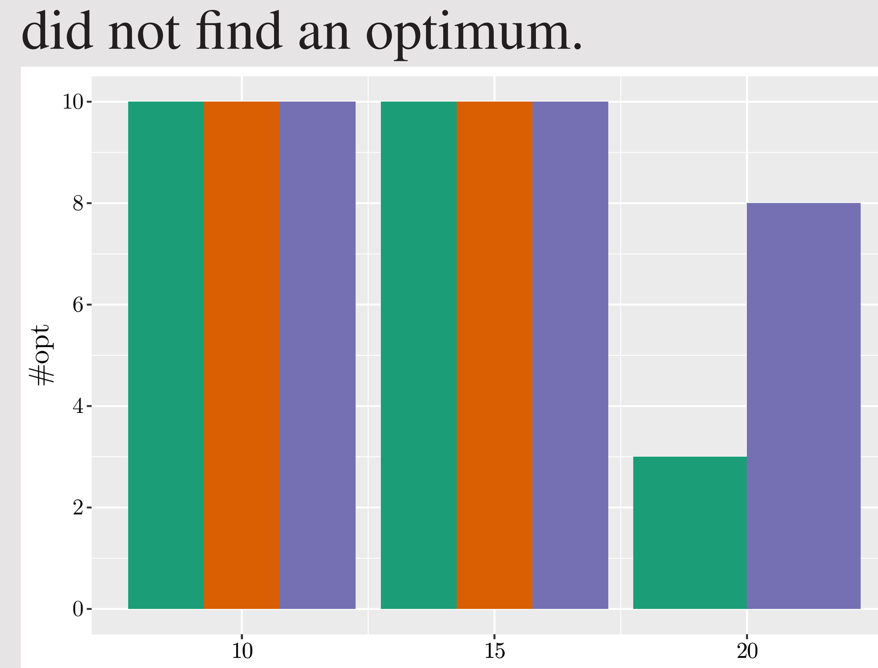
(no interior points) $$\sum_{jk \,|\, ijk \in \Delta(S)} t_{ijk} = \sum_{j \in S \setminus \{i\}} x_{ij} - 1, \ \forall i \in S$$

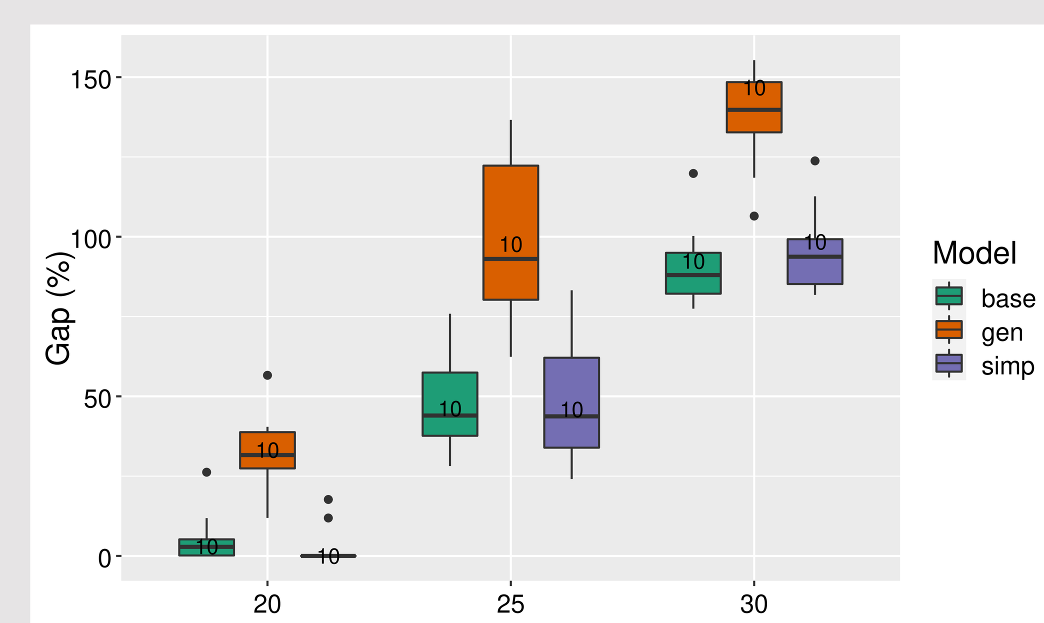
Note that a polygon may have several different triangulations, leading to symmetry related performance issues while solving our formulation. To cope with this, we derived two sets of inequalities that impose a single triangulation for any polygon.

Results

Instances We generated a set of 50 uniformly random instances: 10 per value of $n = 10, 15, 20, 25, 30$. We use "base" to refer to the above formulation; "simple" and "gen" indicate the addition of each of the sets of symmetry-breaking inequalities.

The left chart shows the number of instances in which each model found the optimum. The right one shows the optimality gap for the instances where at least one of the models did not find an optimum.





As one can see, the addition of the "simp" inequalities improved the overall performance, while including "gen" deteriorated it. Observed that, although the latter inequalities are more powerful, since they generalize the previous ones, their separation is computationally too expensive.

Newer Results We designed other formulation, based on the geometric dual perspective of a polygon triangulation. With this new model, we are able to solve instances with up to 25 points to optimality and in many of them, faster than the best known method from the literature [4].