On some connection problems in straight-line segment arrangements*

Helmut Alt†‡ Sergio Cabello§ Panos Giannopoulos†‡ Christian Knauer¶

Abstract

We study the complexity of some problems of the following type: Given a set of straight-line segments in the plane and a set of cells in the induced arrangement, compute the minimum number of segments one needs to remove so that the cells become connected. We show that the problems of connecting two cells and connecting all cells are both NP-hard. We also discuss several polynomial-time solvable and fixed-parameter tractable cases.

1 Introduction

Let S be a set of straight-line segments in $\mathbb{R}^2$ and $\mathcal{A}(S)$ be the arrangement induced by S.

In the 2-Cells-Connection problem we are given two points $s, t \in \mathbb{R}^2$ and we want to compute a set $S' \subseteq S$ of minimum possible size, with the property that s and t belong to the same cell of $\mathcal{A}(S \setminus S')$. In other words, we want to compute an s-t path that *crosses* the minimum number of segments of S counted without multiplicities. The *cost* of a path is the total number of segments crossed by it.

In the All-Cells-Connection problem we want to compute a set $S' \subseteq S$ of minimum possible size such that $\mathcal{A}(S \setminus S')$ consists of one cell only.

Without loss of generality, we assume that every segment in S intersects at least two other segments and that both endpoints of a segment are intersection points. We say that two segments cross if and only if they intersect at a common interior point (a segment crossing).

Results. We show that 2-Cells-Connection is NP-hard even when no three segments intersect at a point. When any three segments may intersect only at a common endpoint, the problem is fixed-parameter tractable with respect to the number of segment crossings. We also consider an application variant of the problem where the segments lie inside a polygon with holes P and have their endpoints on its boundary and the s-t path must also stay inside P, and show it to be fixed-parameter tractable with respect to the number of holes of P. On the other hand, we show that All-Cells-Connection is NP-hard even if no three segments intersect at a point and there are no segment crossings.

Related work. Bereg and Kirkpatrick [1] studied the problem of computing the so called barrier resilience in wireless sensor networks, i.e., the minimum number of disks whose removal connects two given cells in an arrangement of unit disks (sensors), and gave a 3-approximation algorithm based on a shortest path computation in the dual of the arrangement. The complexity of the problem though remains open.

2 Connecting two cells

We show that 2-Cells-Connection is NP-hard by a reduction from MAX-2-SAT, a well studied NP-complete problem [6]: Given a propositional CNF formula Φ with m clauses on n variables and at most two variables per clause, decide whether there exists a truth assignment that satisfies at least k clauses, for some given $k \in \mathbb{N}$, $k \leq m$. Let ℓ be the maximum number of occurrences of a variable in Φ. For simplicity we can assume that $\ell \leq 3$ since the restricted version of MAX-2-SAT where any variable occurs in at most 3 clauses is still NP-complete [9].

Using Φ we construct an instance consisting of a set of segments S and two points s and t as follows.

Abusing the terminology slightly, by a *segment* we actually mean a set of identical single segments stacked on top of each other. The cardinality of the set is the *weight* of the segment. Either all or none of the single segments in the set can by crossed by a path.

There are three different types of segments, T_1, T_2, and T_3, according to their weight, see Fig. 1. Segments of type T_1 have weight 1 (single or light segments), while segments of type T_2, T_3 have weight $(m+1)$ and $18n(m+1)$ respectively (heavy segments).

First, we construct a polygon, called the *tunnel*, with heavy boundary segments of type T_3, see Fig. 1. There are 21 boundary segments in total. The tunnel has a 'zig-zag' shape and can be seen as having 8 corridors, $C_1, \ldots, C_8$. It starts with C_1, the *main* corri-

*Part of this research was conducted at the 9th McGill - INRIA Barbados Workshop on Computational Geometry, 2010.

†Institut für Informatik, Freie Universität Berlin, Takustraße 9, D-14195 Berlin, Germany, {alt, panos}@mi.fu-berlin.de

‡This research was supported by the German Science Foundation (DFG) under grant Kn 591/3-1.

§Department of Mathematics, IMFM and FMF, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia, sergio.cabello@fmf.uni-lj.si

¶Institut für Informatik, Universität Bayreuth Universitätsstraße 30 D-95447 Bayreuth Germany, christian.knauer@uni-bayreuth.de

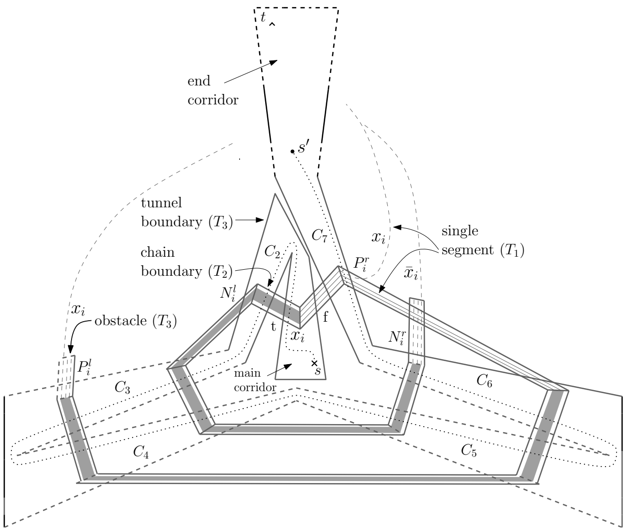

Figure 1: Variable chain and a possible path from s to s'; a shaded area represents a set of ℓ single segments.

dor (at the center of the figure), which contains point s, then it turns left to C_2, then right, etc., gradually turning around to C_7 and then to the *end* corridor C_8 (at the top). The latter contains point t. The total size of the tunnel is $21 \cdot 18n(m+1) = \mathcal{O}(nm)$.

The rest of the construction will force any s-t path of some particular cost (to be given shortly) to stay always in the interior of the tunnel.

Each variable of Φ is represented by a *chain* of segments, see Fig.1. The chain has 12 pieces and each piece is separated from its neighbors by (or ends at, when it is the last one) a 'short' heavy segment of type T_3, referred to as an *obstacle*; with the exception of the two ends of the chain, all obstacles lie in the interior of the tunnel. Each piece consists of:

(i) 2 *chain boundary* segments of type T_2; every endpoint of such a segment lies on an obstacle.
(ii) ℓ single segments (type T_1).

It has $2(m+1)+\ell$ segments in total and it is said to be crossed by a path if and only if all of its segments are crossed.

Consider the example of the variable chain x_i in Fig. 1, and its part in the main corridor. From the obstacle, the chain extends to the left towards corridor C_2 and to the right towards corridor C_7. Consider the right half of the chain. In the first piece all ℓ single segments lie between the two boundary segments: they start at the obstacle in the main corridor and end at the obstacle in C_7. In the second piece, denoted by P_i^r (positive-right), all single segments start at this latter obstacle. Some (in the example just one) end inside the end corridor at intersection points that represent clauses with a positive occurrence of x_i; in Fig. 1, these segments are drawn by dashed lines, which can be made straight by stretching the

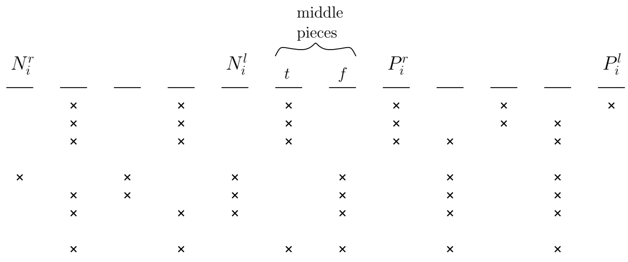

Figure 2: All possible combinations of six pieces of a variable chain crossed by a minimum-cost s-s' path.

end corridor sufficiently. The rest continue down to C_6, always lying between the boundary segments of the piece, and end at an obstacle inside. The third, fourth, and fifth piece are similar to the first one: all single segments of a piece (represented by a shaded area) lie between its respective boundary segments. The chain makes three consecutive right turns, from C_6 to C_5, then to C_4, and then to C_3. In the last piece, P_i^l (positive-left), all single segments start at an obstacle in C_3. Again, some go up to the end corridor, to clauses with a positive occurrence of x_i, while the rest end at an obstacle outside the tunnel. Note that any single segment may intersect a chain boundary segment only outside the tunnel. The left half of the chain is constructed in an analogous fashion. Its second and sixth piece are denoted by N_i^l (negative-left) and N_i^r (negative-right) respectively; some single segments coming from these pieces go to clauses with a negative occurrence of x_i.

Each clause of Φ is represented by an intersection of two single segments inside the end corridor; see Fig. 3 for an example of the overall construction. Each segment corresponds to some literal x_i or $\bar{x}_i$ in the clause: in the first case the segment comes from either P_i^r or P_i^l, while in the second one it comes from either N_i^r or N_i^l. For the construction, these choices for a each clause can be made arbitrarily, provided that one segment intersects the tunnel from the left side and the other one from the right. In this way, the end corridor is obstructed by m pairs of intersecting segments such that any path from the intermediate point s' to point t staying inside the tunnel must intersect at least one segment from each pair.

Observe that any minimum-cost path from s to s' that stays in the interior of the tunnel crosses only 6 pieces from each variable chain (and no obstacles), where at least one piece from every two consecutive ones is crossed; see Fig. 1 for an example of such a path. There are 7 such possible sets of pieces, see Fig. 2, and the choice is made independently for each chain. Consider the two *middle* pieces, parts of which lie in the main corridor. When only the left one is crossed (first three sets), P_i^r is crossed as well, while none of N_i^l, N_i^r is crossed: effectively, x_i is set to *true*.

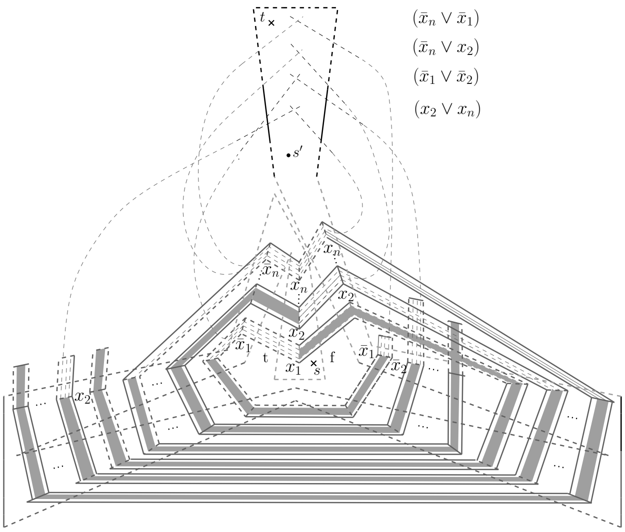

Figure 3: Example of overall construction.

Symmetrically, when only the right piece is crossed (next three sets), x_i is set to *false*. The last set does not correspond to a valid truth assignment. The total cost for such a path is $6n(2(m+1)+\ell)$.

Thus, for an s-t path to cost $6n(2(m+1)+\ell)+k$, for some $0 \le k < m$, its subpath from s' to t in the end corridor must cross at most k single segments that have not been crossed before. This implies that there are at least $(m-k)$ clauses from each of which at least one segment has been already crossed by the s-s' subpath. For such a clause, let $u_i \in \{x_i, \bar{x}_i\}$ be the literal that the crossed segment corresponds to. By construction, if $u_i = x_i$, at least one of P_i^r, P_i^l has been crossed, while if $u_i = \bar{x}_i$, at least one of N_i^r, N_i^l has been crossed. From the discussion above, x_i is set to true in the first case and to false in the second one, and, hence, the clause is satisfied.

Lemma 1 *There exists an s-t path with a cost of at most $6n(2(m+1)+\ell)+k$ if and only if there exists a truth assignment that satisfies at least $(m-k)$ clauses of Φ.*

We can modify our construction by replacing every heavy segment with a set of distinct parallel single segments in a way such that every single segment in S intersecting the original heavy segment now intersects all the segments in the new set and making sure that no three segments have a point in common.

Theorem 2 2-Cells-Connection *is NP-hard even when no three segments intersect at a point.*

We can reduce 2-Cells-Connection to the minimum color path problem (MCP): Given a graph G with colored (or labeled) edges and two of its vertices, find a path between the vertices that uses the minimum possible number of colors. We color the edges of the dual graph G of $\mathcal{A}(S)$ as follows: two edges of G get the same color if and only if their corresponding edges in $\mathcal{A}(S)$ lie on the same segment of S. Then, finding an s-t path of cost k in $\mathcal{A}(S)$ amounts to finding a k-color path in G between its two vertices that correspond to the cells which s, t lie in.

However, MCP is NP-hard [2] and W[1]-hard [5] (with respect to the number of colors in the path) even for planar graphs, it has a polynomial-time $\mathcal{O}(\sqrt{n})$-approximation algorithm and is non-approximable within any polylogarithmic factor [8].

2.1 Tractable cases

Consider the colored dual graph G of $\mathcal{A}(S)$ as defined above. A face of G (except the outer one) corresponds to a point of intersection of some $r \ge 2$ segments and has r colors and, depending on the type of intersection, from r to $2r$ edges. For example, for $r = 2$ we can get two multiple edges, a triangle, or a quadrilateral, with two distinct colors.

When any three segments may intersect only at a common endpoint and no two segments cross, G can only have multiple edges, bi-chromatic triangles, and arbitrary large faces where all edges have different colors. In this case, since two segments can intersect only at one point, each color induces a connected subgraph of G, in fact a tree for there can be no monochromatic cycle in G. Then, 2-Cells-Connection reduces to a shortest path problem in an uncolored modification of G, where each monochromatic tree is contracted into a star.

Generalizing this, if we allow k segment crossings, i.e., bi-chromatic quadrilaterals in G, we can easily find a minimum-cost s-t path as follows. Let $C \subseteq S$ be the set of the (at most $2k$) segments participating in these crossings. For every possible subset C' of C, first contract every edge of G corresponding to a segment in C', and then assign an infinite weight to every edge corresponding to a segment in $C \setminus C'$. For a fixed subset, a solution can be found by computing a shortest path in an uncolored weighted graph.

Theorem 3 2-Cells-Connection *is fixed-parameter tractable with respect to the number of segment crossings if any three segments may intersect only at a common endpoint.*

2.2 An application

Let P be a polygon with h holes and S be a set of n segments lying inside P with their endpoints on its boundary; see Fig. 4, where, for clarity, the boundary of P is drawn by a set of simple closed curves. We consider the restricted 2-Cells-Connection problem on $P \cup S$ where the s-t path may not cross the boundary. This version is also NP-hard by a simple reduction from the general one.

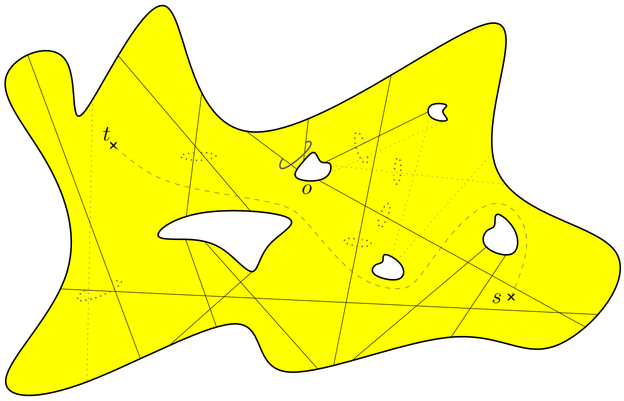

Figure 4: Some segment clusters in a polygon with holes and a minimum-cost s-t path.

We partition S into clusters using homotopies as follows: two segments belong in the same cluster if and only if there is a continuous transformation from one to the other, during which the endpoints stay on the boundary of P; for this, s and t are treated as special holes. In Fig. 4 for example, the segments that touch hole o give four clusters. Using simple topological arguments we prove the following:

Lemma 4 *S can be partitioned into $O(h^4)$ clusters with the property that either all or none of the segments in a cluster are crossed by a minimum-cost s-t path.*

A minimum-cost s-t path can now be easily found by testing all possible subsets of clusters.

Theorem 5 *The restricted* 2-Cells-Connection *problem in a polygon with holes is fixed-parameter tractable with respect to the number of holes.*

3 Connecting all cells

We prove that All-Cells-Connection is NP-hard by a reduction from the well-known NP-hard problem of feedback vertex set (FVS) in planar graphs [6]: Given a planar graph G, find a minimum-size set of vertices X such that $G - X$ is acyclic.

First, we subdivide every edge of G obtaining a planar bipartite graph G'. It is clear that G' has a feedback vertex set of size k if and only if G has one. Next, we use the result by de Fraysseix et al. [4] (see also Hartman et al. [7]), which states that every planar bipartite graph is the intersection graph of horizontal and vertical segments, where no two of them cross. Let S be the set of segments whose intersection graph is G'; it can be constructed in polynomial time. Since G' has no triangles, no three segments of S intersect at a point. Finally, observe that all cells in $\mathcal{A}(S)$ become connected by removing k segments if and only if G' has a feedback vertex set of size k.

Theorem 6 All-Cells-Connection *in NP-hard even if no three segments intersect at a point and there are no segment crossings.*

It is also easy to see that a k-size solution to All-Cells-Connection corresponds to a k-size solution of FVS in the intersection graph of the input segments. For general graphs, FVS is fixed-parameter tractable when parameterized with the size of the solution [3], and has a polynomial-time 2-approximation algorithm [10].

Corollary 7 All-Cells-Connection *is fixed-parameter tractable with respect to the size of the solution and has a polynomial-time* 2*-approximation algorithm.*

Acknowledgments

We would like to thank Primož Škraba for bringing to our attention some of the problems studied in this abstract.

References

[1] S. Bereg and D. G. Kirkpatrick. Approximating barrier resilience in wireless sensor networks. In *Proc. of the 5th ALGOSENSORS*, volume 5804 of *LNCS*, pages 29–40. Springer, 2009.

[2] H. Broersma, X. Li, G. Woeginger, and S. Zhang. Paths and cycles in colored graphs, p.299. *Australasian J. Combin.*, 31:299–312, 2005.

[3] J. Chen, F. V. Fomin, Y. Liu, S. Lu, and Y. Villanger. Improved algorithms for feedback vertex set problems. *J. Comput. Syst. Sci.*, 74:1188–1198, 2008.

[4] H. de Fraysseix, P. O. de Mendez, and J. Pach. Representation of planar graphs by segments. *Intuitive Geometry*, 63:109–117, 1991.

[5] M. R. Fellows, J. Guo, and I. Iyad. The parameterized complexity of some minimum label problems. *J. Comput. Syst. Sci.*, 76:727–740, 2010.

[6] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman & Co., 1979.

[7] I. B. Hartman, I. Newman, and R. Ziv. On grid intersection graphs. *Discrete Mathematics*, 87:41–52, 1991.

[8] R. Hassin, J. Monnot, and D. Segev. Approximation algorithms and hardness results for labeled connectivity problems. *J. Comb. Optim.*, 14(4):437–453, 2007.

[9] V. Raman, B. Ravikumar, and S. S. Rao. A simplified NP-complete MAXSAT problem. *IPL*, 65(1):1–6, 1998.

[10] V. V. Vazirani. *Approximation Algorithms.* Springer, 2001.