The Logic in Computer Science Column

BY

Yuri Gurevich

Microsoft Research
One Microsoft Way, Redmond WA 98052, USA
gurevich@microsoft.com

On the rectilinear Steiner problem

Andreas Blass
University of Michigan
ablass@umich.edu

Yuri Gurevich
Microsoft Research
gurevich@microsoft.com

Abstract

This is a gentle introduction to the Rectilinear Steiner Problem. The interest in the Rectilinear Steiner Problem is related to the Very Large-Scale Integration (VLSI) technology that combines thousands of transistors into a single chip. Our note is based on the pioneering paper of Maurice Hanan. The main result of Hanan is an algorithm reducing any solution to a solution within a so-called Hanan grid. We simplify Hanan's algorithm.

1 Introduction

In combinatorics, vertices of a graph are abstract entities, and edges are unordered pairs of vertices. In applications, the nature of vertices may be important, and edges may be more informative.

In the case of interest to us here, graphs are geometric objects residing on a plane with a fixed cartesian coordinate system. The vertices of such a graph are points on the plane, and there are only finitely many of them; an edge is a continuous line connecting two vertices and composed of finitely many horizontal and vertical segments. No two edges cross; a point that belongs to distinct edges is an end point of each of them. The total length of the edges of a geometric graph is the *cost* of the graph. In the sequel, graphs are by default such geometric graphs.

For any set P of points on the plane, the *Rectilinear Steiner Problem* for P, in short RSP(P), is to find a minimal-cost connected graph whose vertices include P. We say that a graph G is a *candidate solution* or simple a *solution* for RSP(P) if G is connected and includes all points of P as vertices; the P vertices are the *terminals* of G, and the remaining vertices are *auxiliary*. A solution of minimal cost is a *minimal solution* for RSP(P).

Note that auxiliary vertices of degree 2 are optional in the following sense. For any RSP(P) and any solution G for RSP(P), removing a degree 2 auxiliary vertex (and combining the incident edges into one) or putting a degree 2 auxiliary vertex onto an existing edge (and thus breaking it into two incident edges) produces another solution for RSP(P) of the same cost.

The interest in the Rectilinear Steiner Problem is related to the Very Large-Scale Integration (VLSI) technology that combines thousands of transistors into a single chip. The wires on a chip run horizontally or vertically.

Maurice Hanan wrote an influential pioneering work on the Rectilinear Steiner Problem [1]. The *Hanan grid* or just *grid* of a point set P on the plane comprises points (a, b) such that the vertical line $x = a$ and the horizontal line $y = b$ each hosts at least one point of P. Hanan's main theorem asserts that, for any RSP(P), there is a minimal solution with all vertices on the grid. The proof of the theorem is more informative. It provides an algorithm supporting the following theorem.

Theorem 1 (Grid Theorem)**.** *For any RSP(P), every solution with at least one vertex or edge corner off the Hanan grid for P can be transformed to a smaller-or-equal cost solution with all vertices and edge corners on the grid.*

Since there are only finitely many solutions with all vertices and edge corners on the grid, the Grid Theorem implies the existence of a minimal solution for any RSP(P). Another obvious consequence of the Grid Theorem is that, for any solution for RSP(P) that bulges out of the enclosing axis-aligned rectangle of P, there is a smaller-or-equal cost solution within the rectangle. In fact, by Corollary 6 below, there is always a smaller-cost solution within the rectangle.

This paper is a gentle introduction to the Rectilinear Steiner Problem based on Hanan's paper. We simplify Hanan's algorithm for the Grid Theorem and make various additional improvements.

It will be convenient to use geographical language: the x-axis goes west to east, the y-axis goes south to north, x-coordinates are longitudes, y-coordinates are latitudes, lines parallel to the x-axis are parallels, lines parallel to the y-axis are meridians.

2 Junctions vs. true terminals

As in graph theory, our trees are acyclic connected graphs. The vertices of a tree are often called nodes.

There is a simple transformation of any non-tree solution G_1 to a smaller-cost tree solution. Remove one edge of a cycle in G_1; the remaining graph G_2 is connected. If G_2 still has a cycle, remove one cycle edge in G_2; the remaining graph G_3 is connected. And so on. Since the number of edges in G_1 is finite, the process terminates and produces a tree solution.

According to Hanan [1], in every minimal-cost solution for any RSP(P) with $t \geq 2$ terminals, the number of auxiliary vertices of degree ≥ 3 is $\leq t - 2$ [1].

In fact, a stronger claim is valid for every tree solution for RSP(P). For brevity, auxiliary vertices of degree ≥ 3 will be called *junctions*, and terminals of degree 1 will be called *true terminals*.

Theorem 2. *In every tree solution for any RSP*(P) *with $t \geq 2$ true terminals, the number of junctions is $\leq t - 2$.*

We will use the standard (and easy to prove) fact that every tree with N nodes has exactly $N - 1$ edges.

Proof. First we prove the theorem in a special case where every terminal is a true terminal and there are no degree 2 vertices. Let n be the number of nodes in our tree solution. Then the tree has altogether $n + t$ nodes and $n + t - 1$ edges. Each terminal has one edge adjacent to it and each junction has at least three, so there are $\geq t+3n$ ends-of-edges. That's twice the number of edges, so $2(n+t-1) \geq t+3n$. Transposing some terms in this inequality gives $t - 2 \geq n$, as required.

Now we prove the theorem in the general case. Given a tree solution S for RSP(P) with n junctions and $t \geq 2$ true terminals, let $P_0 = \{p \in P : \text{Degree}(p) = 1\}$. Treat the non-true terminals of S as auxiliary nodes, and remove all degree 2 nodes from S. This transforms S into a tree solution for RSP(P_0), with $n_0 \geq n$ junctions and the same number t of true terminals. By the special case of the theorem, $n \leq n_0 \leq t - 2$. □

3 The case of three terminals

In the case of three terminals, the Rectilinear Steiner Problem is rather simple and instructive. Our exposition follows Hanan's but we provide all the proofs. This section will not be used in the sequel. If x_1, x_2, x_3 are real numbers with $x_1 \leq x_2 \leq x_3$, the *median* $\text{Med}\{\!\{x_1, x_2, x_3\}\!\}$ of the multiset $\{\!\{x_1, x_2, x_3\}\!\}$ is x_2.

Lemma 3. $\min_x \sum_{i=1}^{3} |x_i - x| = \max x_i - \min x_i$, *and the minimum is attained if and only* $x = \mathit{Med}\{\!\{x_1, x_2, x_3\}\!\}$.

Proof. Without loss of generality, $x_1 \leq x_2 \leq x_3$ so that $x_1 = \min x_i$, $x_2 = \text{Med}\{\!\{x_1, x_2, x_3\}\!\}$ and $x_3 = \max x_i$. Then

$$\begin{aligned}\sum |x_i - x| &= |x_2 - x| + (|x_3 - x| + |x - x_1|) \\ &\geq |x_2 - x| + (x_3 - x_1),\end{aligned}$$

and the minimum $x_3 - x_1$ is attained if and only $x = x_2$. □

For points $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$ on the plane, the *rectilinear distance* $d(p_1, p_2)$ between p_1 and p_2 is $|x_1 - x_2| + |y_1 - y_2|$. The rectilinear distance is also known as Manhattan or L_1 distance.

Lemma 4. *Let $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$, $p_3 = (x_3, y_3)$ and $q = (x, y)$ be arbitrary points on the plane. Then*

$$\min_q \sum_{i=1}^{3} d(p_i, q) = (\max x_i - \min x_i) + (\max y_i - \min y_i\})$$

and the minimum is attained if and only if $q = (Med\{\!\{x_1, x_2, x_3\}\!\}, Med\{\!\{y_1, y_2, y_3\}\!\})$.

Proof.

$$\begin{aligned}\min_q \sum_{i=1}^{3} d(p_i, q) &= \min_q \Big(\sum_{i=1}^{3} |x_i - x| + \sum_{i=1}^{3} |y_i - y| \Big) \\ &= \min_x \sum_{i=1}^{3} |x_i - x| + \min_y \sum_{i=1}^{3} |y_i - y|.\end{aligned}$$

The claim follows by applying the previous lemma twice. □

Theorem 5. *Let $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$, $p_3 = (x_3, y_3)$ be distinct points and $q_0 = (x_0, y_0) = (Med\{\!\{x_1, x_2, x_3\}\!\}, Med\{\!\{y_1, y_2, y_3\}\!\})$. Ignoring vertices of degree 2, $RSP\{p_1, p_2, p_3\}$ has a unique minimal solution G, and $Cost(G) = \sum_{i=1}^{3} d(p_i, q_0)$, and either G has no junctions or it has a single junction q_0.*

Proof. Let $c = \sum_{i=1}^{3} d(p_i, q_0)$. In any solution G for RSP$\{p_1, p_2, p_3\}$, there is a path connecting a westmost terminal with an eastmost one. The path goes from longitude $\min\{x_1, x_2, x_3\}$ to longitude $\max\{x_1, x_2, x_3\}$. Accordingly, the total length of the horizontal edge segments of the path is $\geq \max\{x_1, x_2, x_3\} - \min\{x_1, x_2, x_3\}$. Similarly G has a path with the total length of the vertical edge segments $\geq \max\{y_1, y_2, y_3\} - \min\{y_1, y_2, y_3\}$. By Lemma 4, Cost$(G) \geq c$. So any solution of cost c is minimal.

To construct the desired minimal solution, we consider two cases. If q coincides with one of the terminals, say $q = p_2$, then the desired minimal solution has two edges, an edge between p_1 and p_2 of length $d(p_1, p_2)$ and an edge between p_2 and p_3 of length $d(p_2, p_3)$. Otherwise, the desired minimal solution has three edges, between q and terminals p_1, p_2, p_3 of lengths $d(p_1, q), d(p_2, q), d(p_3, q)$ respectively. In either case, the cost is c.

It remains to prove the uniqueness claim. The fact that G has at most one junction follows from Theorem 2. The fact that the junction, if present, should be q_0 follows from the previous lemma. □

4 Proof of Grid Theorem

In this section we prove the Grid Theorem formulated in the introduction. In the process we construct a simplified algorithm that transforms a given solution for RSP(P) into a smaller-or-equal cost solution within the grid.

Recall that degree-2 vertices are optional. In this proof, all corner points of a graph — and only corner points — will be treated as degree-2 vertices. It follows that every edge is horizontal or vertical.

For any solution G, call a parallel or meridian *G-wrong* if it hosts an edge of G but no terminals. Define the *deficit* of G to be the total number of G-wrong parallels and meridians.

We prove the theorem by induction on the deficit. So let G be a positive-deficit solution for RSP(P). It suffices to transform G into a smaller-deficit solution G' of smaller or equal cost.

By the parallel/meridian symmetry, without loss of generality we may assume that there is a G-wrong meridian. For any G-wrong meridian V, let $w(V)$ and $e(V)$ be the numbers of edges connecting vertices on V with vertices to the west or east of G respectively. By the west/east symmetry, without loss of generality we may assume that there is a G-wrong meridian V with $w(V) \leq e(V)$.

Let V' the westmost meridian east of V that hosts a vertex of G. Move V eastward until it merges with V'. The merged meridian needs to be cleaned up. Do this in two stages.

Stage 1. For any two vertices $v \in V$ and $v' \in V'$ of the same latitude, merge v with v'. If v' is terminal then the merged vertex is terminal as well. If v' is auxiliary and the merging results in a vertex of degree 2 that isn't a corner, delete the merged vertex (and combine the two incident edges into one).

Stage 2. List south to north the vertices of the merged meridian resulting from stage 1: $p_1, p_2, \ldots$ If p_i and p_{i+1} are connected by a portion e of a vertical edge that used to reside on V as well as by a portion e' of an edge residing on V', merge e and e'.

The result is a solution G' for RSP(P). The deficit of G' is smaller by at least 1 than that of G because we've eliminated the G-wrong meridian V. There could also have been a G-wrong parallel hosting a single edge of G, namely a horizontal edge between V and V' which would disappear when its ends are brought together. Such a parallel would be eliminated as well.

The total length of horizontal edges of G' is less than or equal to that of G depending on whether $w(V) < e(V)$ or $w(V) = e(V)$. The total length of vertical edges of G' is less than or equal to that of G depending on whether any portions of vertical edges were merged or not. Thus $\mathrm{Cost}(G') \leq \mathrm{Cost}(G)$. This concludes the proof of the Grid Theorem.

Corollary 6. For any RSP(P), every solution for RSP(P) that bulges out of the enclosing rectangle of P can be transformed to a smaller-cost solution within the rectangle.

Proof. The proof is that of the Grid Theorem except a parallel (resp. meridian) is declared G-wrong if it is north or south (resp. west or east) of the enclosing rectangle. Again, without loss of generality there is a G-wrong meridian. By the west/east symmetry, there is a G-wrong meridian west of the enclosing rectangle. If V is the westmost meridian, we have $w(V) = 0 < e(V)$. Accordingly moving V eastward results in smaller total length of horizontal edges and therefore in a smaller cost. □

References

[1] Maurice Hanan, "On Steiner's problem with rectilinear distance," SIAM J. Applied Math., 14 (1966), 255–265.