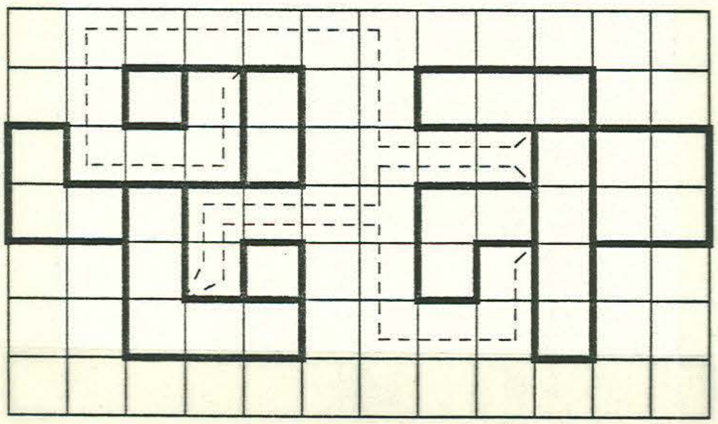
MIT/LCS/TM-237

PLANAR EMBEDDING OF PLANAR GRAPHS



D. Dolev, F. T. Leighton, and H. Trickey

February 1983

Planar Embedding of Planar Graphs

Danny Dolev† IBM Research San Jose Laboratory

Frank Thomson Leighton‡ Mathematics Dept. MIT

Howard Trickey* Computer Science Dept. Stanford University

Key words: graph embedding, planar graphs, outerplanar graphs, crossover-free, $\mathcal{NP}$-complete

Abstract

Planar embedding with minimal area of graphs on an integer grid is an interesting problem in VLSI theory. Valiant [V] gave an algorithm to construct a planar embedding for trees in linear area; he also proved that there are planar graphs that require quadratic area.

We fill in a spectrum between Valiant's results by showing that an N-node planar graph has a planar embedding with area $O(NF)$, where F is a bound on the path length from any node to the exterior face. In particular, an outerplanar graph can be embedded without crossovers in linear area. This bound is tight, up to constant factors: for any N and F, there exist graphs requiring $\Omega(NF)$ area for planar embedding.

Also, finding a minimal embedding area is shown to be $\mathcal{NP}$-complete for forests, and hence for more general types of graphs.

† Part of the work was done while the author was at Stanford University and supported by a Chaim Weizmann postdoctoral fellowship and DARPA contract MDA903-80-C-0102. Current address: Institute of Mathematics and Computer Science, Hebrew University, Givat Ram 91904, Jerusalem, Israel.

‡ Work supported by DARPA contract N00014-80-C-0622, Air Force contract OSR-82-0326, and a Bantrell Fellowship.

* Work supported by an NSERC scholarship and NSF grant MCS-80-12907.

1. Introduction

VLSI design motivates the following class of problems: given a graph, map its vertices onto a plane and its edges onto paths in that plane between the corresponding mapped vertices. Normally there are some restrictions that the mappings must obey, such as a minimum distance between mapped vertices. The maps give a *layout*, and the problem is to find a layout with a small cost. The mapping restrictions and the cost function together specify a particular member of the class of layout problems.

Embedding of graphs has been extensively studied during the last few years [L80, V, FP, BK, CS, R, RS, L81, L82]. In this paper we consider the layout problem when the layouts are rectilinear embeddings *without crossovers* and the cost is the area of a box bounding the layout. To avoid complications, we assume that graphs are restricted to have vertices of degree 4 or less.

In [V], Valiant looked at the layout problem for rectilinear embeddings (both with and without crossovers), using the bounding box area cost. He proved that a tree of vertices with maximum degree 4 can be laid out without crossovers in an area that is linear in the number of edges (or vertices). He also showed how to get a such an embedding for any planar graph using quadratic area, and proved that there are planar graphs requiring quadratic area.

Definition: A planar graph has *width* F if there is a planar embedding of the graph such that every node of the graph is linked to the external face of the embedding by a path of at most F vertices.

We shall show that any N-node planar graph of width F can be laid out in $O(NF)$ area. Special cases of this include linear area embeddings for trees and outerplanar graphs, and quadratic area embeddings for graphs of width $O(N)$. Furthermore, the area bound is tight up a to constant factor. This fills in a spectrum between Valiant's results. The graph in Fig. 1.1 has N nodes and width F, and each component requires $\Omega(F^2)$ for a planar embedding (see [V]), so the entire graph requires $\Omega(NF)$ area.

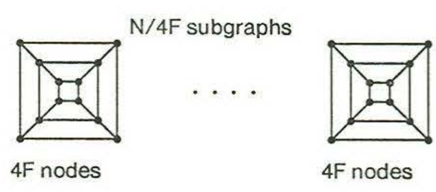


Figure 1.1 Graph needing $\Omega(NF)$ area

We shall also show that finding an optimal embedding for a forest is $\mathcal{NP}$-complete.

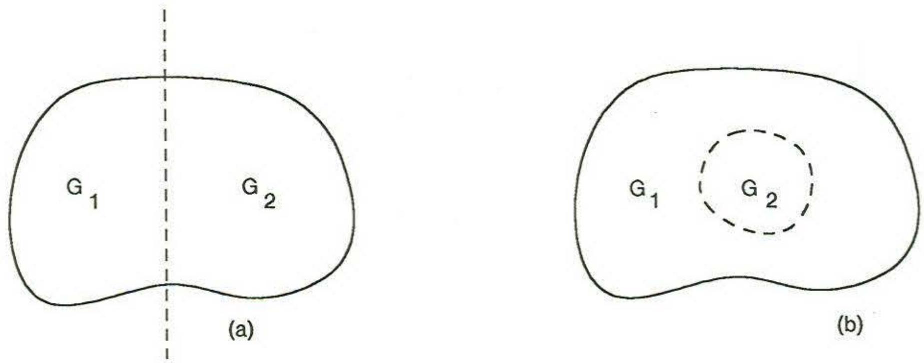


Figure 2.1 (a) good separation (b) bad separation

2. A Planar Graph Separator

The layout method is basically that used by Valiant [V] and Leiserson [L80] to get embeddings with crossovers allowed: the graph is split into two by removing edges, each subpart is recursively laid out, and then the subproblem layouts are "married" by embedding the edges that were removed.

The key to methods like these are separator theorems, which guarantee that one can always split up a graph as needed without having to remove too many edges. Lipton and Tarjan [LT] investigated planar graph separators, and showed that any planar graph of N nodes can be split into approximately equal-sized parts by removing $O(\sqrt{N})$ edges. However, sometimes these separations are unsuitable for a divide-and-conquer layout strategy. The edges removed by their method divides the graph either as shown in Fig. 2.1(a) or Fig. 2.1(b), and only the former can be used in our layout method. The following theorem helps to characterize when the "good" separations occur.

Theorem 1. *A planar graph with $N > 2$ vertices of degree at most 4 and width F can be separated into two subgraphs by removing $O(F)$ edges, such that each subgraph has at least $\frac{1}{3}$ of the vertices. Given a planar drawing of the graph, the separation can be made as shown in Fig. 2.1(a) rather than Fig. 2.1(b) (assuming the given drawing actually has width F or less).*

Proof: If necessary, add dummy edges to the graph until the given drawing has a simple cycle as the outer face, and there are only triangles as interior faces. This can always be done, keeping the graph planar and without increasing the width. Call this graph G.

Define the *distance* of a vertex in G to be the number of nodes in the shortest path from the vertex to the outer face. Let a *separating path* in G be a simple path from a vertex on the outer face to another one, such that the distances of the vertices on the path go like 1, 2, ..., $k-1$, k, k, $k-1$, ..., 2, 1 or 1, 2, ...,

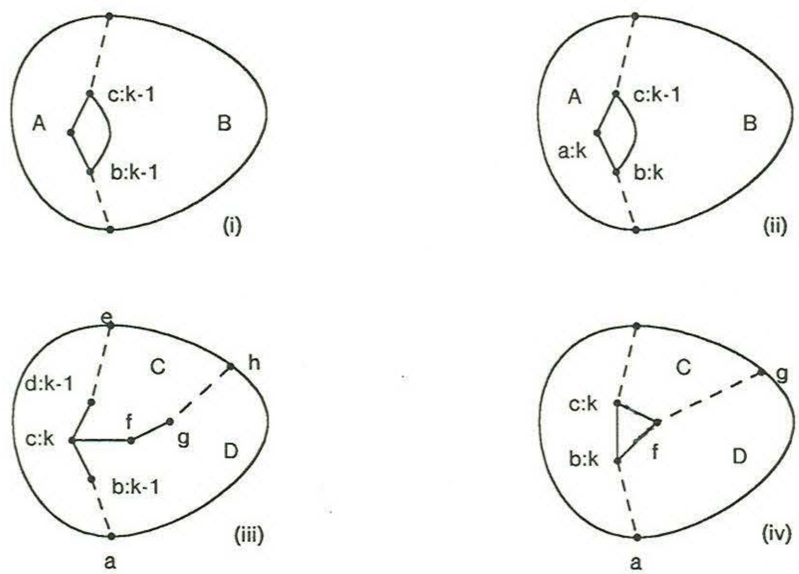


Figure 2.2 Cases for Separator Theorem

$k-1$, k, $k-1$, ..., 2, 1. We will find a separating path with $k \leq F$ such that no more than two thirds of G's vertices are on either side of it (where "side" refers to one of the two regions that the path divides the plane into if line are drawn from the path ends to infinity). Then G can be separated as required in the theorem statement by removing at most $4 \times 2k$ edges from the vertices in the separating path. The vertices on the path itself can be divided between the two sides so that neither side ends up with more than two thirds of G.

Start out with any outer-face edge as the separating path. Assume, in general, that we have a situation with A vertices on one side of the path, B vertices on the other, and $N - A - B$ vertices on the path itself. If $A \leq \frac{2}{3}N$ and $B \leq \frac{2}{3}N$ then we are done, so assume that $B > \frac{2}{3}N$.

The cases that arise are shown in Fig. 2.2 (where vertex distances are shown after colons). Figures 2.2(i) and 2.2(ii) are degenerate cases that are handled by using the right b–c path instead of the left one. The process continues with the new path.

In Fig. 2.2(iii), vertex f is not the same as b or d. By the definition of distance of a vertex, there is an exit path, g–$\cdots$–h, with vertices of distances j, $j-1$, ..., 1 or j, $j+1$, ..., $m-1$, m, $m-1$, ..., 1 where $j \leq k+2$ and $m \leq F$. This exit path may coincide wholely or in part with d–$\cdots$–e or b–$\cdots$–a, but it never need cross over them because it can merge with the rest of whichever path it touches. Also, the path should not go back through f; this can always be avoided in a triangulated graph.

Most of the B vertices that were on the right side of the original separating path are now divided into pieces of sizes C and D. Assuming $D \geq C$, the new separating path is a–$\cdots$–b–c–f–g–$\cdots$–h. Clearly, this

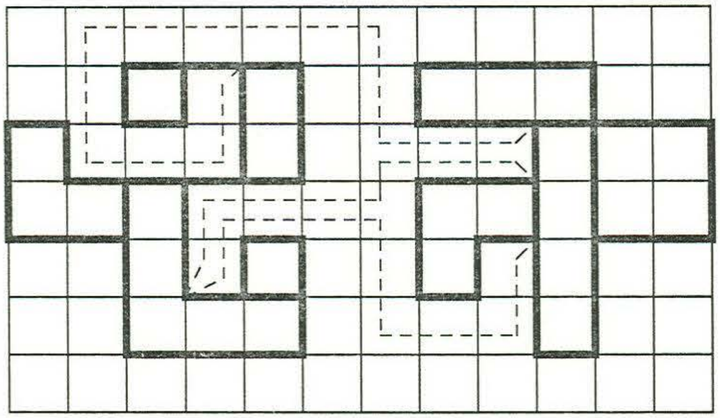


Figure 3.1 Marrying two embeddings

new path is of the required form. If $D \leq \frac{2}{3}N$ then we are done, otherwise repeat the process. The vertices f and g are part of the B vertices that were on one side of the the old separating path, so we must have $D < B$. This means that progress has been made towards the stopping condition, since we have decreased the number of vertices on the big side of the path, a process that cannot go on forever. Note that the new separating path may have a bigger maximum distance, but this is irrelevant as far as progress towards stopping is concerned.

The situation of Fig. 2.2(iv), where the path has two vertices of maximum distance in the middle, is handled just like case (iii). If $D \geq C$, the new separating path $a\text{-}\cdots\text{-}b\text{-}f\text{-}\cdots\text{-}g$ is of the required form. Progress towards the stopping condition has been made, because we will have lost vertex c at the very least.

The above operations can be repeated until a separating path has been found with no more than $\frac{2}{3}N$ vertices on either side, proving the theorem. ∎

3. Planar Embedding Algorithm

The layout method used in [V] and [L80] for embedding with crossovers allowed almost works for planar embeddings. The difference is in the marrying step.

In order for the layout method to work recursively, it has to be able to embed a graph so that it is topologically equivalent to a given planar drawing. Suppose G is separated into G_1 and G_2 using the separator theorem of the previous section, and then the subparts are embedded, respecting topology. Then the removed edges can be drawn in the plane without crossovers, because they are attached to vertices that are still on the outer faces of G_1 and G_2, in the same order. For example, see Fig. 3.1, where the separating edges are shown dotted.

To turn such a drawing into a grid embedding, insert a new grid line for every dotted straight line segment. For the diagonal lines making the connections, at most two new horizontal and two new vertical grid lines may be needed. (The existing edges may have to be shifted so they make their final approach from a different direction.) Let K be the number of "kinks", i.e., horizontal and vertical grid lines that need to be added to connect any exterior face vertex of a given embedding to somewhere completely outside. It is easy to see that K increases only by $O(1)$ at each marrying step, because the added edges needn't wrap around the layout more than once. Thus, if i is the maximum of the number of marrying stages involved in laying out G_1 and G_2, then they can be married by added $O(iF)$ horizontal and vertical grid lines to embed the $O(F)$ separating edges.

Theorem 2. *Any planar graph G with N vertices of degree at most 4, and width at most F, has a planar embedding in a grid of area $A(N) = O(FN)$.*

Proof: Other than the separation and marrying methods, the layout algorithm is the same as the one in [V]. It has to be able to produce an embedding in an $H \times W$ grid, as long as $\frac{1}{3} \leq H/W \leq 3$, and HW is sufficiently large. Suppose by induction that $A(N)$ is sufficient area for an N-vertex graph. Also, suppose that $K(N)$ is a bound on the number of kinks.

G is separated into G_1 and G_2 by removing $O(F)$ edges, with $|G_1| = x|G|$, $\frac{1}{3} \leq x \leq \frac{2}{3}$. Then an $(H - cFK(N)) \times (W - cFK(N))$ grid is divided in two by a cut parallel to the shorter side in the ratio $x : (1-x)$. By a theorem in [V], the aspect ratios of the two pieces will be in the range $[\frac{1}{3}, \frac{2}{3}]$. If G_1 and G_2 can be laid out in these pieces, then the embedding can be completed as described above, inserting at most $cFK(N)$ horizontal and vertical grid lines, for some constant c. So the theorem is true if (assuming $H \leq W$)

$$(H - cFK(N))(xW - xcFK(N)) \geq A(xN), \qquad \forall x, \quad \frac{1}{3} \leq x \leq \frac{2}{3}.$$

Using $HW \geq A(N)$ and $(H+W)/\sqrt{A(N)} \leq 4/\sqrt{3}$, this will be true if

$$x(A(N) - \frac{4}{\sqrt{3}}\sqrt{A(N)}\,cFK(N)) \geq A(xN), \qquad \forall x, \quad \frac{1}{3} \leq x \leq \frac{2}{3}.$$

After $\log_{2/3} N/F$ separation steps the graph pieces are no larger than F, so if we stop the recursion at that point we have $K(N) = O(\log N/F)$. It is easily verified by substitution that

$$A(N) = \alpha NF - \beta N^{\frac{1}{2}} F^{\frac{3}{2}} \log \frac{N}{F}$$

satisfies the recurrence, for some α and β independent of N and F. In the base case, with $N = F$, an $O(N^2)$ embedding (see [V]) can be used. One has to be careful to get an embedding that preserves the topology of a given planar drawing, but it is easy to see how to do this. ∎

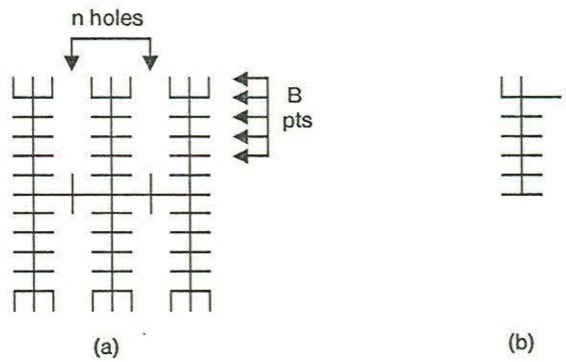


Figure 4.1 Frame Tree for $n = 2$, $B = 5$

4. $\mathcal{NP}$-Completeness of Optimal Forest Embedding

Given a forest and an integer A, the *forest layout problem* is to find whether or not there is a planar rectilinear embedding with area less than or equal to A. In this section we will show that the forest layout problem is $\mathcal{NP}$-complete. This will be done by transforming the 3-partition problem to it.

In the 3-partition problem there is a set of integers $x_1, \ldots, x_{3m}$ such that

$$\sum_{i=1}^{3m} x_i = mB$$

and $B/4 < x_i < B/2$ for $1 \leq i \leq 3m$. The question is whether the set can be partitioned into m disjoint sets such that each set sums to B. This problem is known to be strongly $\mathcal{NP}$-complete [GJ].

Consider the tree in Fig. 4.1(a). Call it the *frame tree*. There are vertices at every grid point except for $m = 2n$ holes of size B. (The case for m odd will be considered later; it is just a trivial modification.)

Lemma 3. *The only embeddings of the frame tree with a bounding box area of* $(4n+3) \times (2B+3)$ *or less and leaving* mB *free grid points are either exactly like that shown in Fig. 4.1(a) (possibly after point relabelling), or modifications of that diagram where some of the tops of the vertical spines are changed as in Fig. 4.1(b) and its various reflections.*

Proof: The tree has $(4n+3) \times (2B+3) - mB$ vertices, so the embedding is required to use every grid point for a vertex or else leave it free. This means that no edge of the tree can be stretched to a path of 2 units, for that would take up a grid point in the middle that is not used for embedding a graph vertex.

Any layout using only 1-unit edges must have all of the degree-4 vertices of a vertical spine one on top of the other, as in the diagram. For otherwise there would have to be two degree-4 vertices at opposite corners of a 1-unit square, which is impossible (one of the other corners would have to be shared between two vertices).

Therefore, the only possible changes to the given diagram, other than point renaming, are ones at the degree-2 vertices, such as in Fig. 4.1(b). ∎

Notice that if the frame tree is embedded using an allowed folding near the top of a spine, this cuts a hole into two pieces of sizes 2 and $B-2$. There cannot be more than one fold into a hole. From now on, use to term "hole" to mean either a B-point vertical slot or one of these $2+(B-2)$-point aggregates.

Theorem 4. *The forest layout problem is $\mathcal{NP}$-complete.*

Proof: Given an instance of the 3-partition problem, construct the frame tree and add $3m$ other pieces, unconnected to that tree: for each x_i there is a piece consisting of x_i vertices joined into a line by $x_i - 1$ edges. If m is odd, use the frame graph for the next higher even number and fill in one of the vertical holes.

Now we claim that the 3-partition problem instance has a solution iff there is an embedding of this forest with a bounding box area of $(4n+3)\times(2B+3)$. For, by the lemma, if there is such an embedding then it must be as shown in Fig. 4.1(a) with the extra pieces filling up the holes. Since all the grid points are to be used, this gives a solution to the 3-partition problem, because the size restrictions on the x's imply that there must be exactly three pieces in each hole. Conversely, given a solution to the 3-partition problem, a suitable embedding can be found by filling the holes in the frame tree with the pieces corresponding to the partitioned sets.

This is not a polynomial reduction, since the frame tree has a number of vertices of the order of the numbers involved in the 3-partition problem, rather than the number of bits required to represent those numbers. This does not matter, however, since the 3-partition problem is *strongly* $\mathcal{NP}$-complete. The layout problem is in $\mathcal{NP}$ because one can simply guess a mapping of all the vertices to grid points and then verify that the edges can all be put along the connecting lines. Therefore, the forest layout problem is $\mathcal{NP}$-complete. ∎

References

[BK] R. P. Brent and H.T. Kung, "On the area embedding of binary tree layouts," Australian Nat. Univ. Rep. TR-CS-79-07, 1979.

[CS] M. Cutler and Y. Shiloach, "Permutation layout," *Networks* 8, pp. 253–278, 1978.

[FP] M. J. Fischer and M. S. Paterson, "Optimal tree layout," *Proc. 12th ACM Symp. Theory Comput.*, 1980.

[GJ] M. R. Garey and D. S. Johnson, *Computers and Intractability*. San Francisco: Freeman, 1979.

[L81] F. T. Leighton, "New lower bound techniques for VLSI," *Proc. 22nd IEEE Symp. Foundations Comput. Sci.,* 1981, pp. 1–12.

[L82] F. T. Leighton, "A layout strategy for VLSI which is probably good," *Proc. 14th ACM Symp. Theory Comput.,* 1982, pp. 85–97.

[L80] C. E. Leiserson, "Area-efficient graph layouts (for VLSI)," *Proc. 21st IEEE Symp. Foundations Comput. Sci.,* 1980.

[LT] R. J. Lipton and R. E. Tarjan, "A separator theorem for planar graphs," *A Conf. on Theoretical Comput. Sci.,* University of Waterloo, 1977, pp. 1–10.

[R] A. L. Rosenberg, "On embedding graphs in grids," IBM Watson Research Center Tech. Rep. RC 7559 (# 2668), 1979.

[RS] W. L. Ruzzo and L. Snyder, "Minimum edge length planar embeddings of trees," *CMU Conf. on VLSI Systems and Computations,* October 1981.

[V] L. G. Valiant, "Universality considerations in VLSI circuits," *IEEE Trans. Comput.,* vol. C-30, pp. 135–140, 1981.