SIAM J. APPL. MATH.
Vol. 32, No. 4, June 1977

THE RECTILINEAR STEINER TREE PROBLEM IS *NP*-COMPLETE*

M. R. GAREY AND D. S. JOHNSON†

Abstract. An optimum rectilinear Steiner tree for a set A of points in the plane is a tree which interconnects A using horizontal and vertical lines of shortest possible total length. Such trees correspond to single net wiring patterns on printed backplanes which minimize total wire length. We show that the problem of determining this minimum length, given A, is *NP*-complete. Thus the problem of finding optimum rectilinear Steiner trees is probably computationally hopeless, and the emphasis of the literature for this problem on heuristics and special case algorithms is well justified. A number of intermediary lemmas concerning the *NP*-completeness of certain graph-theoretic problems are proved and may be of independent interest.

1. Introduction. Let A be a finite set of points in the (oriented) plane. A *rectilinear Steiner tree* (RST) for A is a tree structure, composed solely of horizontal and vertical line segments, which interconnects all the points in A. An *optimum* RST for A is one in which the line segments used have the shortest possible total length. It is important to note that, in contrast to the usual notion of a "spanning tree", an RST is permitted to have three or more line segments meeting at a point that does *not* belong to A, called a *Steiner point.* In fact, it is frequently the case that *every* optimum RST for A contains one or more Steiner points. The "rectilinear Steiner tree problem" is, given A, to find an optimum RST for A.

The RST problem has received attention from a number of authors [1], [4], [5], [8], [9], [10], [12], [15], [16], motivated primarily by potential applications to wire layout for printed circuit boards. Efficient algorithms for several special cases have been described in [1]. However, no efficient algorithm for constructing optimum RST's in general has yet been found. In this paper we present strong evidence for the impossibility of such an efficient general algorithm, by proving that the general RST problem belongs to the infamous class of *NP*-complete problems.

The widely held belief that all *NP*-complete problems are computationally intractable is based on two important properties of this class:

(A) There is no known polynomial-time algorithm that solves any single problem in the class.

(B) Despite the wide variety and large number of problems in the class, the existence of a polynomial-time algorithm for any one of them would imply that *every* problem in the class could be solved with a polynomial-time algorithm.

For a more detailed discussion of the class of *NP*-complete problems and its members, see [2], [13], [14]. Suffice it to say that until now the closest relative of the RST problem known to be *NP*-complete was the Steiner problem in graphs [13], a much more general and abstract problem. It had been hoped that the highly restricted geometric nature of the RST problem might render it more tractable. Our result, however, shows that this is not the case.

The reader is referred to [2] for a thorough description of the formal requirements for a proof of *NP*-completeness. Basically, the two steps required in proving that a particular problem X is *NP*-complete are:

* Received by the editors June 18, 1976.

† Bell Laboratories, Murray Hill, New Jersey 07974.

(a) Prove that X can be solved in polynomial time by a "nondeterministic" Turing machine.

(b) Prove that some known NP-complete problem X' can be "polynomially transformed" into X, in such a way that any polynomial-time algorithm for solving X could be used to solve X' in polynomial time.

The first requirement is rather technical, but trivial for all the problems we discuss here. Thus we shall omit verification of (a) from our proofs, leaving the details to the interested reader. Our proofs will focus on the transformation required by (b).

2. Overall strategy. In order to prove the NP-completeness of the RST problem, we first prove a sequence of auxiliary NP-completeness results. Given a graph $G=(V, E)$, a *node cover* for G is any subset $V^* \subseteq V$ that contains at least one endpoint from every edge. We begin with the following problem which was shown to be NP-complete in [6]:

Node cover in planar graphs. Given a planar graph $G=(V, E)$ and an integer k, does there exist a node cover V^* for G satisfying $|V^*| \leqq k$?

We then transform "node cover in planar graphs" into the following more restricted version of itself:

Node cover in planar graphs with maximum degree 3. Given a planar graph $G=(V, E)$ with no vertex degree exceeding 3 and an integer k, does there exist a node cover V^* for G satisfying $|V^*| \leqq k$?

Next we transform this node cover problem into yet another restricted node cover problem, which requires that the node cover itself be of a special form:

Connected node cover in planar graphs with maximum degree 4. Given a planar graph $G=(V, E)$ with no vertex degree exceeding 4 and an integer k, does there exist a node cover V^* for G satisfying $|V^*| \leqq k$ and such that the subgraph of G induced by V^* is connected?

This last problem will be transformed into the RST problem, stated as follows:

Rectilinear Steiner tree. Given a finite set A of integer coordinate points in the plane and an integer l, does there exist an RST for A with total length less than or equal to l?

3. The proofs. We now describe the required transformations.

LEMMA 1. "*Node cover in planar graphs with maximum degree* 3" *is NP-complete.*

Proof. Given a planar graph G and an integer k, we construct a planar graph G' with no vertex degree exceeding 3 and an integer k' such that G' has a node cover of size k' if and only if G has a node cover of size k.

Let $G=(V, E)$ where $V=\{v_1, v_2, \cdots, v_n\}$. The construction begins with a fixed planar representation of $G=G_0$. For each integer i, from 1 up to n, we construct a planar representation for a graph G_i from that for G_{i-1} as follows (see Fig. 1):

(i) Let $\{v_i, w_1\}, \{v_i, w_2\}, \cdots, \{v_i, w_p\}$ be the edges leaving v_i in the order that they occur around v_i in the planar representation of G_{i-1}.

(ii) Replace v_i with a cycle consisting of the new vertices $u_i(j)$, $v_i(j)$, $1 \leqq j \leqq n$, and the new edges $\{u_i(j), v_i(j)\}$, $1 \leqq j \leqq n$, $\{v_i(j), u_i(j+1)\}$, $1 \leqq j \leqq n-1$, and $\{v_i(n), u_i(1)\}$.

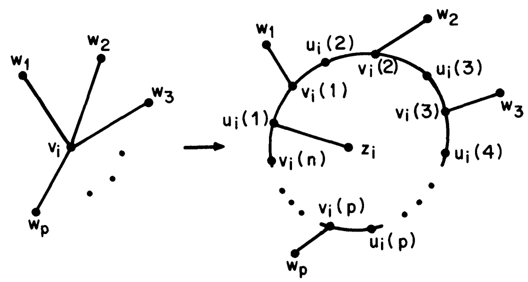

FIG. 1. *Vertex substitute for Lemma* 1

(iii) Replace each edge $\{v_i, w_j\}$ by the edge $\{v_i(j), w_j\}$, add a new vertex z_i, and add the edge $\{u_i(1), z_i\}$.

Finally we set $G' = G_n$ and $k' = n^2 + k$. Observe that G' has no vertex with degree exceeding 3.

Now suppose V^* is a node cover for G satisfying $|V^*| \leqq k$. Then there is a node cover V_1^* for G' satisfying $|V_1^*| \leqq k'$, namely

$$V_1^* = \{v_i(j) : v_i \in V^*, 1 \leqq j \leqq n\} \cup \{u_i(1) : v_i \in V^*\} \cup \{u_i(j) : v_i \notin V^*, 1 \leqq j \leqq n\}.$$

It is easy to check that V_1^* has the required properties.

Conversely, suppose V_1^* is a node cover for G' satisfying $|V_1^*| \leqq k'$. Since the only vertices of G' that cover edges corresponding to edges of G are the $v_i(j)$ vertices, we immediately know that the set

$$V^* = \{v_i : \text{for some } j, 1 \leqq j \leqq n, v_i(j) \in V_1^*\}$$

must form a node cover for G. We shall show that $|V^*| \leqq k$. First we note that we may assume that $u_i(1) \in V_1^*$ for every i, since the edge $\{u_i(1), z_i\}$ must be covered and z_i only has degree 1. Define, for $1 \leqq i \leqq n$, $S_i = V_1^* \cap \{u_i(j), v_i(j) : 1 \leqq j \leqq n\}$. In order to cover all $2n$ edges in the cycle for v_i we must have $|S_i| \geqq n$. Since $k' = n^2 + k$, this implies that at most k values of i can satisfy $|S_i| > n$. Furthermore, since $u_i(1) \in S_i$, the only set of exactly n vertices that covers all $2n$ edges in the cycle for v_i is $\{u_i(j) : 1 \leqq j \leqq n\}$. Thus if there exists a j for which $v_i(j) \in S_i$, we must have $|S_i| > n$. Since this occurs for at most k values of i, we have $|V^*| \leqq k$, and V^* is the desired node cover for G.

Since G' can clearly be constructed in time a polynomial in the size of G, and has the desired node cover if and only if G does, our transformation works as required, and the restricted problem is *NP*-complete. □

LEMMA 2. *"Connected node cover in planar graphs with maximum degree* 4" *is NP-complete.*

Proof. Given a planar graph G with no vertex degree exceeding 3 and an integer k, we construct a planar graph G' with no vertex degree exceeding 4 and an integer k' such that G has a node cover of size k if an only if G' has a "connected" node cover of size k'.

Let $G = (V, E)$ where $V = \{v_1, v_2, \cdots, v_n\}$. The construction again begins with a fixed planar representation for G and then performs the following operations (see Fig. 2):

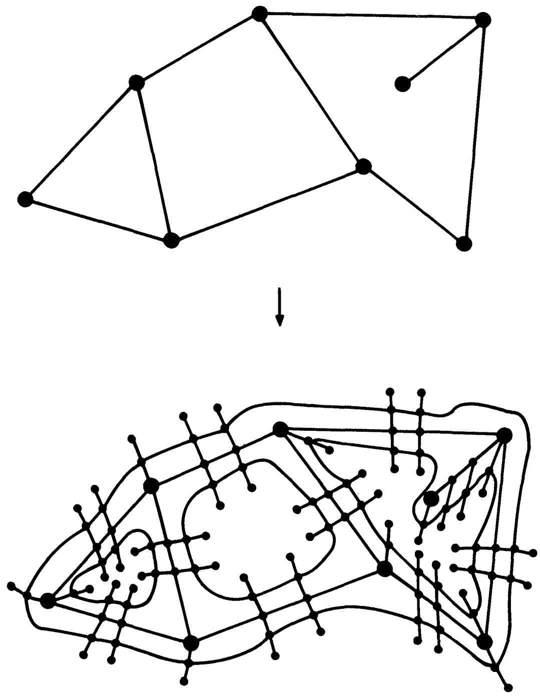

FIG. 2. *Connected node cover construction*

(i) Replace each edge $\{v_i, v_j\} \in E$ by three edges $\{v_i, x_i(j)\}$, $\{x_i(j), x_j(i)\}$, $\{x_j(i), v_j\}$ where $x_i(j)$ and $x_j(i)$ are new vertices.

(ii) Consider in turn each region R of the planar representation of the graph resulting from step (i). Let W be the set of all those vertices on the boundary of R which have degree less than 4 (including edges added for previously considered regions). For each $w \in W$, introduce two new vertices w', w'' in the interior of R and add the edges $\{w, w'\}$, $\{w', w''\}$. (Notice that each original vertex v_i will act as such a w at least once, since it initially has degree at most 3, and each x-type vertex will act as such a w for every one of the (one or two) regions on whose boundary it occurs.)

(iii) For each region, join all the w' vertices in that region into a single cycle in such a way that the graph remains planar. (This is easy to do, for example, by joining them up in essentially the same order as their neighbors on the original boundary of the region.)

Let $G' = (V', E')$ be the resulting graph. Let r be the total number of w' vertices introduced in step (ii). Then $|V'| = n + 2 \cdot |E| + 2 \cdot r$. Set $k' = k + |E| + r$.

The graph G' and integer k' can clearly be constructed in polynomial time. To complete the proof of NP-completeness, we show that G' has its desired node cover if and only if G does.

First suppose that V^* is a node cover for G with $|V^*| \leqq k$. Define the set $V_1^* \subseteq V'$ by

$$V_1^* = V^* \cup \{\text{all } w'\text{-type vertices}\}$$
$$\cup \{x_i(j) \in V' \colon v_j \in V^*, \text{and either } v_i \notin V^* \text{ or } i < j\}.$$

We claim that V_1^* is the desired connected node cover for G'. Clearly every edge of the form $\{w, w'\}$, $\{w', w''\}$ or which joins two w'-type vertices is covered, since all w'-type vertices belong to V_1^*. Now consider any edge $\{v_i, v_j\} \in E$. By construction of V_1^* we have

$$\{v_i, v_j\} \cap V_1^* = \{v_i, v_j\} \cap V^* \neq \varnothing.$$

If $\{v_i, v_j\} \cap V^* = \{v_j\}$, then v_j and $x_i(j)$ belong to V_1^*, covering the three edges $\{v_i, x_i(j)\}$, $\{x_i(j), x_j(i)\}$, $\{x_j(i), v_j\}$. If $\{v_i, v_j\} \cap V^* = \{v_i, v_j\}$ and $i < j$, then v_i, v_j, and $x_i(j)$ all belong to V_1^*, again covering those three edges. Thus V_1^* is a node cover for G'. Furthermore, since exactly one of each pair $\{x_i(j), x_j(i)\}$ belongs to V_1^* for each $\{v_i, v_j\} \in E$, we have $|V_1^*| \leqq k + r + |E| = k'$. It remains to show that the subgraph induced by V_1^* is connected. All the w'-type vertices that were placed in the same region are connected by their common cycle, and each vertex in $V^* \cap V_1^*$ is joined to at least one such cycle for a region on whose boundary it occurs. Finally the w'-cycles for adjacent regions are connected together through their common edge $\{v_i, v_j\}$ (as viewed in G) via either $x_i(j)$ or $x_j(i)$. Thus V_1^* is the desired connected node cover for G'.

Conversely suppose that V_1^* is a connected node cover for G' satisfying $|V_1^*| \leqq k'$. Since each w''-type vertex is adjacent only to the corresponding w'-type vertex, we know that all w'-type vertices belong to V_1^* and may assume that no w''-type vertices belong to V_1^*. We also may assume that exactly one of each pair $\{x_i(j), x_j(i)\}$ belongs to V_1^*, by replacing $x_i(j)$ by v_i in V_1^* whenever both belong. (At least one *must* belong.) With these assumptions on V_1^* we immediately have

$$|V_1^* \cap V| \leqq k' - r - |E| = k.$$

We claim that $V^* = V_1^* \cap V$ forms the desired node cover for G. Consider any edge $\{v_i, v_j\} \in E$. Without loss of generality suppose $x_i(j)$ is the single member of $\{x_i(j), x_j(i)\}$ that belongs to V_1^*. Then, in order to cover the edge $\{x_j(i), v_j\}$, we must have $v_j \in V_1^*$ and hence $v_j \in V_1^* \cap V$. Thus $V_1^* \cap V$ contains at least one endpoint of every edge in E and is the desired node cover for G. □

THEOREM 1. *"Rectilinear Steiner Tree" is NP-complete.*

Proof. Given a planar graph G with no vertex degree exceeding 4 and an integer k, we construct a set A of points in the oriented plane and an integer l such that G has a connected node cover of size k if and only if there is an RST for A with total length l or less.

Let $G = (V, E)$ where $V = \{v_1, v_2, \cdots, v_n\}$ and $E = \{e_1, e_2, \cdots, e_m\}$. Consider a discrete grid of squares imposed on the oriented plane, consisting of all line segments having the form $[(6in^2, 6jn^2), (6(i-1)n^2, 6jn^2)]$ or $[(6in^2, 6jn^2), (6in^2, 6(j-1)n^2)]$ where i and j are integers. The construction begins by obtaining a planar representation of G which uses only horizontal and vertical line segments

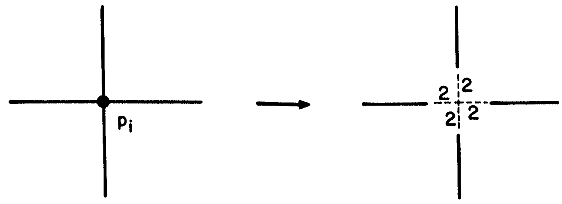

FIG. 3. *Vertex deletion for Theorem* 1

chosen from the above grid,[1] each vertex of G being mapped into a point of the form $(6in^2, 6jn^2)$ for some integers i and j. In order to do this, we need the property that G has no vertex with degree exceeding 4, since no point in the grid has more than 4 incident line segments. Given this property, there are a number of ways in which the desired representation can be constructed in low-order polynomial time. For instance, we could use the methods of [11] to obtain a description of an ordinary planar representation, and from this construct a list of the regions, each given as a sequence of vertices representing its boundary cycle. We can then build up a "rectilinear" representation from an initial region, by successively adding adjacent regions, one at a time.

The set A of points will be constructed by removing portions of the line segments that make up the planar representation of G. For each $v_i \in V$, let p_i denote the point corresponding to v_i in this representation. Considering each p_i in turn, delete p_i and the portions of all incoming line segments within distance 2 of p_i (see Fig. 3). Let L denote the total length of the remaining line segments. Finally replace each remaining line segment by the set of all points on that line segment which have integer coordinates. These points form the set A. We set $l = L + 2m + 2(k-1)$.

Suppose G has a connected node cover V^* with $|V^*| \leqq k$. The corresponding RST for A contains all the line segments joining pairs of points from A that are exactly one unit apart (i.e., those line segments deleted in the last step of the construction for A). Each connected component of the resulting structure corresponds to an edge of G and we call the component corresponding to edge $e_s \in E$ the *e_s-component.* In addition, for each $v_i \in V^*$, we shall also select some of the line segments incident with p_i that were deleted in the second to last step of the construction for A. To do this we first choose a spanning tree for the subgraph of G induced by V^*. This spanning tree exists by the connectivity property of V^* and contains $|V^*|-1$ edges. For every edge $e_s = \{v_i, v_j\} \in E$ that does not belong to the spanning tree, select one endpoint v_i that belongs to $e_s \cap V^*$ and select the length 2 line segment joining p_i to the e_s-component. For every edge $e_s = \{v_i, v_j\}$ that does belong to the spanning tree, select both length 2 line segments joining the e_s-component to p_i and to p_j. Since V^* is a connected node cover for G, the resulting collection of selected line segments forms an RST for A. For each edge in the spanning tree we added two length 2 line segments and for each edge not in the spanning tree we added a single length 2 line segment. Thus the total length of all the selected line segments is at most $L + 2m + 2(k-1)$ as required.

[1] Each edge of G will be a path composed of a sequence of one or more elementary grid segments.

Conversely, suppose there exists an RST for A with total length l or less. We shall show that there must exist such an RST having a similar form to that constructed in the previous portion of the proof. By a result of Hanan [9], since all points in A have integer coordinates, we may restrict our attention to RST's composed only of line segments whose endpoints have integer coordinates. Let T be an optimum RST for A having this form and which maximizes, among all such RST's, the number of unit length line segments with both endpoints belonging to A. Recall that the only possible such line segments are those deleted in the last step of the construction for A, which we shall call *edge segments*. We claim that T must contain all of the edge segments. Suppose T fails to contain some one of the edge segments. Then adding that edge segment to T must form a cycle. However, by our construction there are no cycles composed only of edge segments, so this cycle must contain some unit length segment which is not an edge segment. Deleting that nonedge segment results in an alternative optimum RST for A which contains one more edge segment than T does, contradicting the assumption that T contains a maximum possible number of edge segments. Thus, T contains all the edge segments deleted in the last step of the construction of A.

Using our terminology introduced earlier, we may now think of T as composed of the "edge component" for each edge of G plus some additional line segments joining these edge components together. It is convenient to think of this collection of additional line segments in the most elementary form, as a collection of unit-length segments, having endpoints with integer coordinates, which we call *supplementary segments*. Observe that, since T contains all the edge segments and since the total length of T is at most l, the number of supplementary segments in T is at most $2m+2(k-1)$.

We shall now show that the supplementary segments that belong to T come from a very restricted set. Any supplementary segment in T must form part of some path, composed entirely of supplementary segments, that joins two edge components. (Of course, some of the points on that path may be Steiner points, from which additional paths branch off.) For each point p_i, define the *active region* for p_i to be the set of all points reachable from p_i by a "rectilinear" path of length less than $3n^2$ (see Fig. 4). Consider any supplementary segment that is not contained in any active region. By our construction of A and the definition of active region, any path containing that segment which joins two edge components must contain at least $3n^2 > 2m+2(k-1)$ supplementary segments. Since we already know that T cannot contain that many supplementary segments, it must therefore be the case that all supplementary segments in T are contained within active regions. Supplementary segments within a particular active region, say for point p_i, can only serve the very limited purpose of joining together edge components representing edges that met at vertex v_i in G. It is not hard to see, by considering cases, that all such connections in the active region for p_i can be made with minimum possible length simply by joining each edge component involved in such a connection to the point p_i with a length 2 line segment. Thus we may assume that T consists of all the edge components plus various length 2 line segments joining certain edge components to certain points p_i. This is essentially the same form as the RST constructed in the first half of the proof.

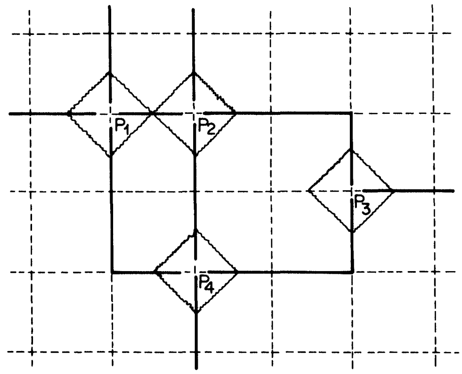

FIG. 4. *Four active regions in original grid*

Now we must use this RST to determine a connected node cover for G. We claim that

$$V^* = \{v_i \in V\text{: some edge component is joined to } p_i\}$$

forms the desired node cover. First, since every edge component is joined to some p_i (for which $v_i \in V^*$) and since an edge component can only be joined to p_i if the corresponding edge in G has v_i as an endpoint, V^* is indeed a node cover. Furthermore, since the total length of T does not exceed l, there can be at most $k-1$ edge components joined to two points p_i. If we delete from T those edge components that are joined to only one p_i, the resulting structure is connected and contains at most $k-1$ edge components. Therefore, if we perform the corresponding edge deletions in G, the resulting subgraph has at most $k-1$ edges, is connected, and contains exactly those vertices which belong to V^*. Since the subgraph is connected and contains exactly those vertices in V^*, it follows that V^* is a connected node cover. Furthermore, since a connected graph with $k-1$ edges can have at most k vertices, we see that $|V^*| \leqq k$. Thus V^* is a connected node cover for G with $|V^*| \leqq k$, as desired. □

This completes the series of reductions showing that the rectilinear Steiner tree problem in *NP*-complete. One might also ask about the computational complexity of the related problem of finding optimal *Euclidean* Steiner trees [3], [7]. Here we are again given a finite set A of points in the plane and wish to connect them up with a tree structure made of straight line segments, possibly using additional "Steiner" points. However, in this case the line segments can run in any direction, not just horizontally and vertically. The authors, with R. L. Graham, have recently shown that this problem, too, is *NP*-complete. The proof, although it shares some common ideas with the one given here, is considerably more involved, and will be presented separately.

REFERENCES

[1] A. V. AHO, M. R. GAREY AND F. K. HWANG, *Rectilinear Steiner trees: Efficient special case algorithms*, Networks, to appear.

[2] A. V. AHO, J. E. HOPCROFT AND J. D. ULLMAN, *The Design and Analysis of Computer Algorithms*, Addison-Wesley, Reading, MA, 1974, Chap. 10.

[3] S. CHANG, *The generation of minimal trees with a Steiner topology*, J. Assoc. Comput. Mach., 19 (1972), pp. 699–711.

[4] S. E. DREYFUS AND R. A. WAGNER, *The Steiner problem in graphs*, Networks, 1 (1972), pp. 195–207.

[5] Y. FU, *Applications of linear graph theory to printed circuits*, Proc. 1st Asilomar Conf. on Systems and Circuits, 1967, pp. 721–729.

[6] M. R. GAREY, D. S. JOHNSON AND L. STOCKMEYER, *Some simplified NP-complete graph problems*, Theoret. Comput. Sci., 1 (1976), pp. 237–267.

[7] E. N. GILBERT AND H. O. POLLAK, *Steiner minimal trees*, this Journal, 16 (1968), pp. 1–29.

[8] S. L. HAKIMI, *Steiner's problem in graphs and its implications*, Networks, 1 (1972), pp. 113–135.

[9] M. HANAN, *On Steiner's problem with rectilinear distance*, this Journal, 14 (1966), pp. 255–265.

[10] ———, *A counterexample to a theorem of Fu on Steiner's problem*, IEEE Trans. Circuit Theory, CT-19 (1972), p. 74.

[11] J. HOPCROFT AND R. E. TARJAN, *Efficient planarity testing*, J. Assoc. Comput. Mach., 21 (1974), pp. 549–568.

[12] F. K. HWANG, *On Steiner minimal trees with rectilinear distance*, this Journal, 30 (1976), pp. 104–114.

[13] R. M. KARP, *Reducibility among combinatorial problems*, Complexity of Computer Computations, R. E. Miller and J. W. Thatcher, eds., Plenum Press, New York, 1972, pp. 85–104.

[14] ———, *On the computational complexity of combinatorial problems*, Networks, 5 (1975), pp. 45–68.

[15] Y. Y. YANG AND O. WING, *An algorithm for the wiring problem*, Digest of IEEE Internat. Symp. on Electrical Networks, 1971.

[16] ———, *Optimal and suboptimal solution algorithms for the wiring problem*, Proc. IEEE Internat. Symp. on Circuit Theory, 1972, pp. 154–158.