A SOLVABLE CASE OF THE OPTIMAL LINEAR ARRANGEMENT PROBLEM ON HALIN GRAPHS

by

Todd Easton, Steve Horton, and R. Gary Parker
School of Industrial and Systems Engineering
The Georgia Institute of Technology
Atlanta, GA 30332

Abstract

The optimal arrangement problem is one of the best known of various labeling problems on graphs. The problem is hard in general but is known to be solvable in certain special cases among which are paths, cycles, and trees. In this note we add to this list by giving a fast algorithm for the problem when instances are confined to a particular type of *Halin* graph. So far as we know, the status of the problem on Halin graphs, in general, is open. We conclude by examining some related issues.

Key Words: Halin graph, linear arrangement, polynomial algorithms

CONGRESSUS NUMERANTIUM 119 (1996), pp.3-17

1 INTRODUCTION

Let $G = (V, E)$ be a finite graph of order n. Then the *optimal linear arrangement* problem (OLA) seeks a vertex labeling $f : V \rightarrow \{1, 2, \ldots, n\}$ such that $\sum_{(u,v)\in E} |f(u) - f(v)|$ is minimum over all such labelings. The problem is well-known to be hard as is its popular variant, *bandwidth* where we are interested in a labeling that minimizes the largest value $|f(u) - f(v)|$. On the other hand, OLA is solvable on trees following work in Shiloach (1979) and more recently, Chung (1984). (Observe that bandwidth remains hard, even on trees with maximum degree three.) For a review of results including ones involving other labeling problems as well as results on particular graphs, the reader is directed to Chung (1981).

In this note, our interest is in OLA confined to a particular subset of the class of *Halin* graphs. The latter are planar graphs, with the property that the edge set can be partitioned into a tree no vertex of which has degree 2 and a cycle C on only and all pendant vertices of the tree. So far as we know, the status of OLA on arbitrary Halin graphs remains open. On the other hand, we will show here that it can be solved by a fast algorithm on the subclass of Halin graphs where the underlying tree is a *caterpillar, i.e.*, a tree such that the removal of degree-1 vertices leaves a path. Representative of this set of restricted Halin graphs is the structure shown in Figure 1. The caterpillar is given in bold.

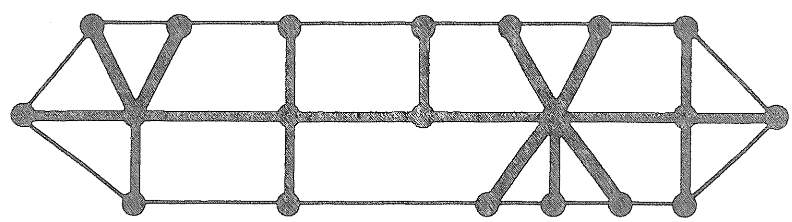

Figure 1: Halin graph with tree component a caterpillar

2 OLA ON ARBITRARY CATERPILLARS

In this section, we give an algorithm for solving OLA on the special tree class of caterpillars. As it turns out, this is actually all that we need in order to solve the problem on the corresponding Halin graph. First, we state some well-known properties of optimal linear arrangements on arbitrary trees. Observe that we denote the vertex and edge sets of a graph G by $V(G)$ and $E(G)$ respectively.

Property 1. An optimal linear arrangement, f^* of a tree T, maps $V(T)$ onto a set of consecutive integers.

(We will hereafter assume that vertex labels are drawn from the integers, $\{1, 2, \ldots, n\}$.)

Property 2. The vertices u and v with $f^*(u)$ and $f^*(v)$ labeled as 1 and n respectively both are pendants, *i.e.*, $deg(u) = deg(v) = 1$.

Property 3. Let P be the path in T which connects the pendants labeled 1 and n. Denote P by $\{i_0, i_1, \ldots, i_t\}$. Then the labelings of P are "monotone" in that

$$f^*(i_x) < f^*(i_{x+1}) \text{ for } i = 0, \ldots, t-1,$$

or

$$f^*(i_x) > f^*(i_{x+1}) \text{ for } i = 0, \ldots, t-1.$$

The next property is specific to caterpillars.

Property 4. Suppose P is a path connecting a pair of pendants in T and moreover, let this be a longest path in T. Then the graph formed by $E(T) \backslash E(P)$ is a vertex disjoint collection of stars each of which is labeled by consecutive integers.

Note that, necessarily, the path P just described will include every vertex on the *spine* of the caterpillar, *i.e.*, all vertices with degree at least two in T.

Following, we state an easy lemma which establishes a lower bound on the value ν of any labeling of a caterpillar and hence for the value of an optimal one. We then state an equally simple algorithm for labeling the vertices of a caterpillar which achieves this value and is thus optimal. We have

Lemma 1: Let T be an arbitrary caterpillar on n vertices and denote by $h_i, i = 1, 2, \ldots, t$, the vertices on the spine of the underlying path P of the caterpillar. Then

$$\nu(T) \geq n - 1 + \sum_{i=1}^{t} \lfloor \frac{(deg(h_i) - 1)^2}{4} \rfloor$$

Proof: Consider any labeling f of T. It is easy to see that the path from $f(i) = 1$ to $f(j) = n$ in T has value at least $n - 1$. Now, if the edges of this path are removed from T then the subgraph of T that results can

be expressed as a (not necessarily vertex disjoint) union of at most t stars each with order at least $deg(h_i) - 1$. But the value of an optimal labeling of a star of order p is well known to be $\lfloor \frac{p^2}{4} \rfloor$ and we are done.
□

We now state an algorithm for caterpillars. First, denoting the caterpillar by T, we find a path P in T as defined in Property 4. Label the end vertices of the stated P by 1 and n respectively. Now, partition the integers $\{2, 3, \ldots, n-1\}$ as $\{2, \ldots, k_1\}, \{k_1+1, \ldots, k_2\}, \ldots, \{k_q+1, \ldots, n-1\}$ and label each of the $q+1$ stars formed by $E(T) \backslash E(P)$ in an optimal way with the integers in the respective components of the stated partition.

Clearly, the labeling of the vertices in P satisfies the monotonicity attribute of Property 3 and, moreover, has a value of exactly $n-1$. Each of the stars formed by the removal of $E(P)$ are labeled by consecutive integers and the labeling is optimal in each case. (Note that the optimal labeling of stars is well known and we take no space here for a description of the strategy.) We have then that the value of the total labeling is exactly the bound value of the lemma and is thus an optimal labeling.

We can demonstrate the procedure by operating on the tree instance from Figure 1. Accordingly, let us select a pair of pendants and label these by 1 and n as indicated. Removal of the edges on the path connecting these vertices leaves the forest of stars shown in Figure 2. As suggested, stars are easy to label in general and in this case the corresponding result is indicated as shown.

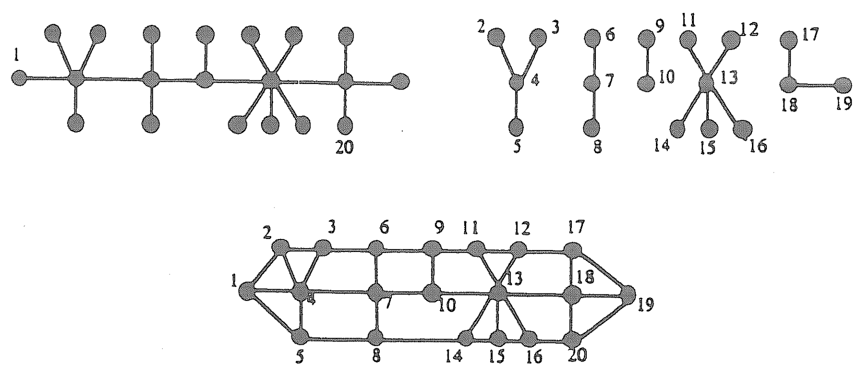

Figure 2: OLA on a caterpillar

3 A SOLVABLE CASE OF OLA ON HALIN GRAPHS

We are now in a position to establish that OLA is solved for the class of Halin graphs where the tree component of the decomposition is a caterpillar. In fact, the algorithm is already at hand: we solve OLA on the caterpillar with the cycle labeling induced directly by the labeling of the pendants of the tree. The bottom graph in Figure 2 illustrates the notion. Note that in this regard, some care is required in labeling the pendants of each star. Specifically, we want the induced labeling on the cycle to be such that the label monotonicity property is satisfied for each of the paths (defining the cycle) connecting vertices labeled 1 and n. For a given embedding of stars, this is a trivial task.

We now establish that this overall strategy is correct. Let us begin with a pair of results, the first of which is easy (and applies to any graph).

Lemma 2: Let G be a finite graph and let $G^1, G^2, \ldots, G^k$ be any set of edge-disjoint subgraphs of G. Then

$$\nu^*(G) \geq \sum_{i \in S} \nu^*(G^i) \text{ for all } S \subseteq \{1, 2, \ldots, k\}$$

□

That is, the value of an optimal labeling for G is at least as large as the sum of optimal values for independent labelings on subgraphs of G. This sum is defined over any set of subgraphs; clearly it is strengthened by judicious choices of the latter.

The next lemma is particularly important and is specific to the stated class of Halin graphs.

Lemma 3: Let G be a Halin graph with cycle C and tree component T which is a caterpillar. Then

$$\nu(G) \geq 3(n-1) + \sum_{i=1}^{t} \lfloor \frac{n_i^2}{4} \rfloor$$

where n_i denotes the order of the ith star defined as per Lemma 1.

Proof: Suppose the instance is defined on graph G and that the labeling algorithm has been applied resulting in $f(G)$. We consider two cases: (1) where vertices u and v with $f(u) = 1$ and $f(v) = n$ are in $V(C)$ but at least one is not a vertex on a longest path in T; (2) where given u and/or v labeled as 1 and n is/are nonpendant vertices in T, *i.e.,* not in $V(C)$.

Consider case (1) first. Since Halin graphs are 3-connected, there must exist in G three internally vertex-disjoint paths connecting every pair of distinct vertices. But any path with termini labeled as say x and y has value at least $|y - x|$ and if the labels on the path satisfy Property 3, this value will be exact. Hence, G will have at least one subgraph consisting of the stated three paths connecting the vertices labeled 1 and n and this subgraph has total label value at least $3(n-1)$. Now, remove this subgraph from G and denote the result by G'. Then the degree of each vertex on the spine of T is reduced by either 2 or 0. But then each of the stars described in the proof of Lemma 1 is either isomorphic to a component in G' or is isomorphic to a subgraph of a component in G'. In either case, we have from Lemma 2 that $\nu(G')$ is at least as large as the optimal values of labelings of the stars. Adding this value to $3(n-1)$ produces the bound of the lemma.

Now, consider case (2). Here, we assume that one or both of the vertices u and v with $f(u) = 1$ and $f(v) = n$ are not pendant vertices in T. For ease, we will consider only the case of the vertex labeled 1; the case for only n as well as for both can be treated in identical fashion and are not presented here. Now, the same argument regarding the formation of G' can be employed where the aforementioned 3-path subgraph, when removed, contributes at least $3(n-1)$ as before. But now the vertex on the spine of T which is labeled 1 has its degree reduced by 3. However, in comparing the total label value of G', we need only examine the effect of the stated degree reduction at the vertex with label 1. In this regard, it is easy to see that G' contains edge-disjoint subgraphs that are either isomorphic to, or that require at least the label value of the stars defined earlier, or else G' will contain a star of order one less but which, by hypothesis, has its hub vertex label fixed at 1, in turn yielding a greater overall label value than that of an optimal star of order one greater. In all cases, G' costs at least $\sum_{i=1}^{t} \lfloor \frac{n_i^2}{4} \rfloor$ and again, we have produced the stated bound. These are the only cases we need to consider and the proof is complete.
□

For ease, both cases described in the proof of Lemma 3 are demonstrated in Figure 3. At the top right in the figure, we show the stars that result *vis-a-vis* the application of the proposed algorithm and specifically relative to the labeling of the caterpillar. For reference, let us code these by the letters a, b, c, d and e as indicated. Now, case 1 is demonstrated at the middle/bottom left where both vertices labeled by 1 and n are on the cycle but not as termini of a longest path in the caterpillar. The subgraph in bold represents a choice of three vertex-disjoint paths connecting the vertices labeled 1 and n; its removal, yielding G' is shown directly below. Further, subgraphs of G' which are isomorphic to the respective stars above are denoted as indicated.

For case 2, we will demonstrate a pair of possibilities for the sake of clarification. The first, which we denote by subcase 2.1, assumes a label of

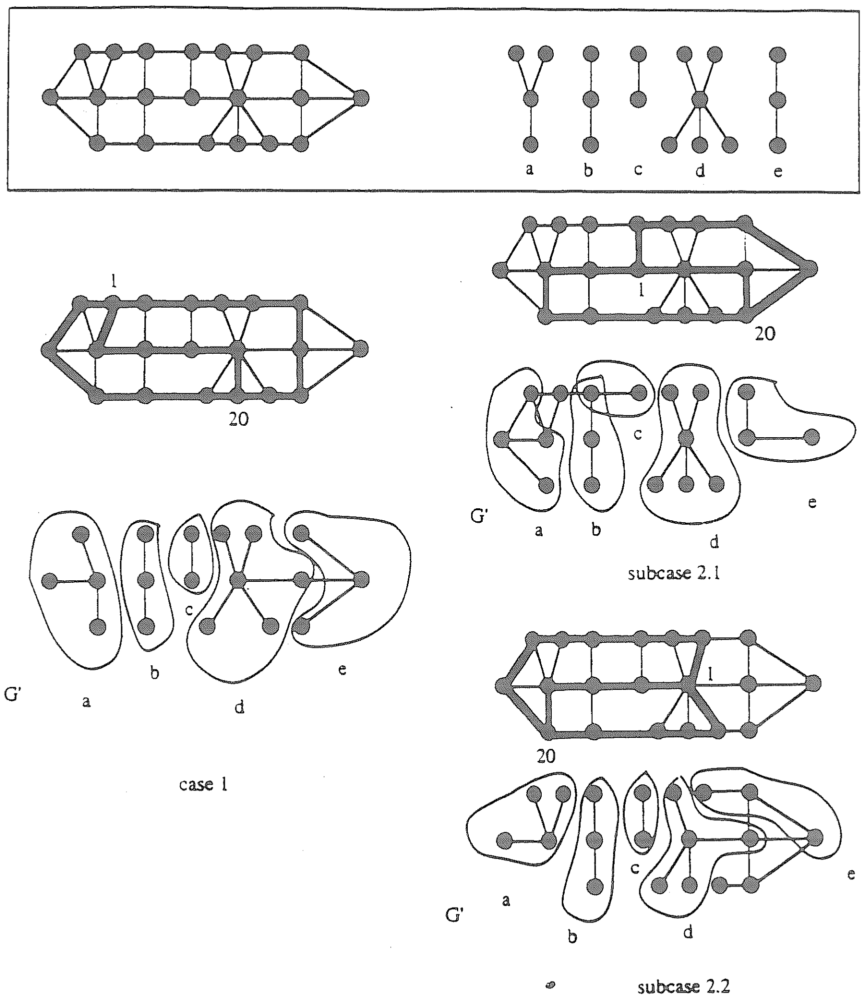

Figure 3: Cases described in proof of Lemma 3

1 at a nonpendant vertex of the caterpillar (with the vertex labeled n on the cycle as shown). Again, the subgraph of $1-n$ paths is denoted in bold and G' is given below. Instructive here is that relative to the given vertex labeled 1, the concomitant removal of its incident edges eliminates *all* of the star specified as component c at the top. However, it must be that somewhere else in G' there is a subgraph isomorphic to c (in this case, an edge) and so that the result of Lemma 2 is applicable. Such an alternative is shown in the figure. On the other hand, the subcase 2.2 shows the outcome when the label 1 is assigned to a high degree, nonpendant vertex of the caterpillar. Here, when G' is formed, a star is left having hub-vertex of degree 4 and moreover, there is no subgraph (anywhere in G') which is isomorphic to d at the top. However, since the hub of the star formed relative to G' has label 1, no labeling of its adjacent vertices can result in a value overall that is strictly better than the optimal labeling of the star of order 6 shown by component d above. Indeed, this outcome generalizes to any pair of stars, one of order t and the other of $t+1$ for all t greater than 3. (Observe that orders less than this leave edges or paths of length two and the argument for case 2.1 applies.)

We now come to the desired result.

Theorem 4: Let G be a Halin graph with cycle C and tree component T which is a caterpillar. Then,

$$f^*(T) = f^*(G)$$

where $\nu^*(G) = \nu^*(T) + 2(n-1)$.

(*i.e.*, the pendant labels relative to T remain optimal when $E(C)$ are appended)

Proof: The result of Lemma 3 specifies a lower bound on the cost of any labeling which most surely holds for an optimal one. Moreover, the application of the stated algorithm will always produce a labeling with exactly this value and is thus optimal.

□

4 DISCUSSION

This short note raises some interesting questions. Not the least of these, indeed probably the foremost, is the issue regarding OLA on the broader class of arbitrary Halin graphs. If found to be polynomially solvable, such an outcome would be particularly worthwhile. This follows since it is known that Halin graphs are contained in the class of 3-terminal recursive graphs and interestingly, labeling problems such as OLA are notoriously resistant to (fast) resolution on recursive structures whereas this is not the case with the vast majority of optimization problems when confined to such graphs (*cf.* Borie *et al.* (1992)). By the same token, if an intractability result could be established for the problem (on Halin graphs), the corresponding status on the other recursive graph classes might be similarly resolvable. Prominent among the latter is the class of partial 2-trees or series-parallel graphs where at this point, OLA appears to be open.

On the other hand, there is an interesting subclass of series-parallel graphs upon which OLA is solvable. Recall that outerplanar graphs are characterized by the absence of subgraphs homeomorphic to K_4 *and* $K_{2,3}$ (the only forbidden subgraph relative to series-parallel graphs is the first of these) and accordingly, following a result in Frederickson and Hambrusch (1988), OLA can be solved on this restricted class. Topologically, outerplanar graphs are structures embeddable in the plane in such a way that all vertices lie in the outer face. The graphs in Figure 4 illustrate some examples.

The Frederickson-Hambrusch results notwithstanding, it should be apparent that for some outerplanar structures, the simple labeling tactic described for the restricted Halin graphs will work. Predicated upon the simple notion that if a graph class decomposes into substructures which can be (easily) labeled in optimal ways and that if this labeling is consistent when substructrures are "composed," then the labeling for the original graph must be optimal. Suppose $G = (V, E)$ to be outerplanar and let us form edge set $F = E\backslash E(C)$. Let us also assume that G is not simply a cycle or a path. First, consider the case where F is either a tree or a forest spanning vertices in $V\backslash\{x, y\}$ where x and y are the two degree-2 vertices in G (there must be at least two of these vertices in any outplanar graph). Either way, we need only label vertices x and y as 1 and n respectively and it's easy to see that there exists optimal independent labelings for the vertices of the alluded to tree/forest components and which form the desired optimal labeling property for the vertices on the cycle C, *cf.* the labeling described by Property 3. The labeling in Figure 5 helps demonstrate.

Now, if the subgraph induced by removal of $E(C)$ does not span as indicated and/or is not even acyclic, it is easy to see that the "fast" labeling strategy employed to this point will not always work, *i.e.*, optimal labels for components of F induce suboptimal labels for C and conversely. For example, in the simple outerplanar graph in part (a) of Figure 6 we could label the subgraph F, shown in bold, as indicated but the outcome is suboptimal for any cycle labeling. On the other hand, the alternative in part (b) labels C in an optimal way and the corresponding labeling for F, while suboptimal accordingly, yields a less harmful outcome overall. We have not bothered to refine cases which work out easily as opposed to those that do not but it is clear that most troublesome are homeomorphs of the sort described by the graph to the right in Figure 7(a). Structures like that in

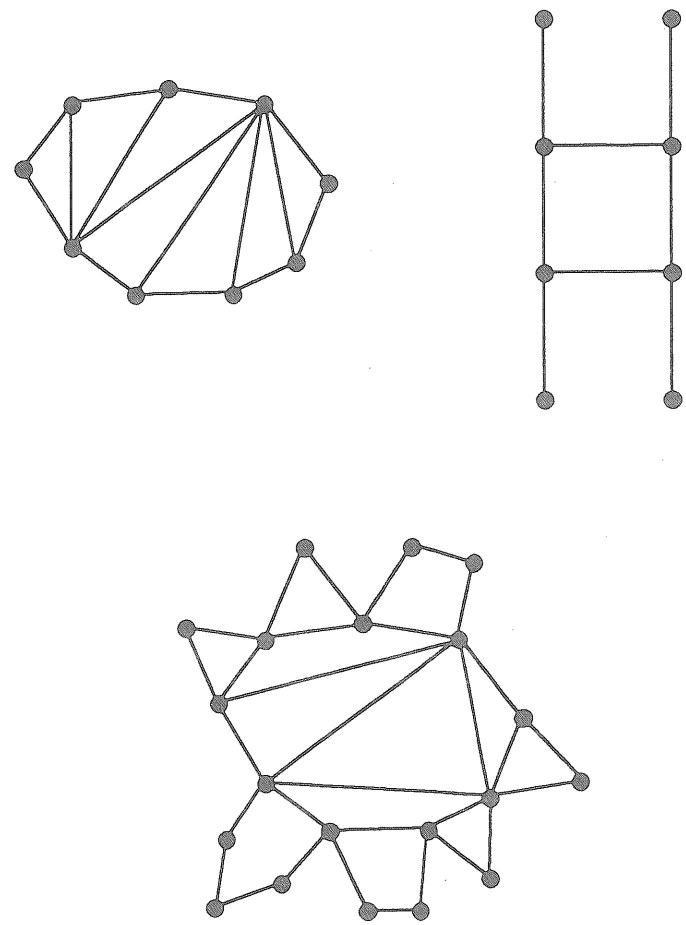

Figure 4: Some outerplanar graphs

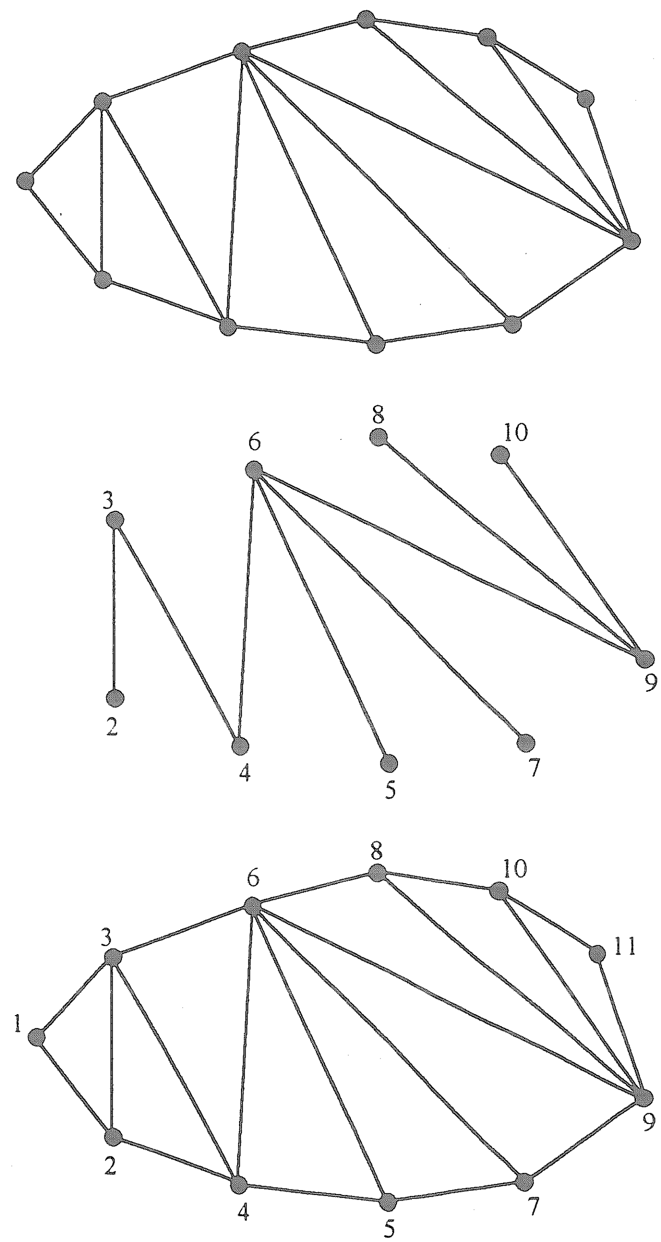

Figure 5: Optimal labeling of certain outerplanar graphs

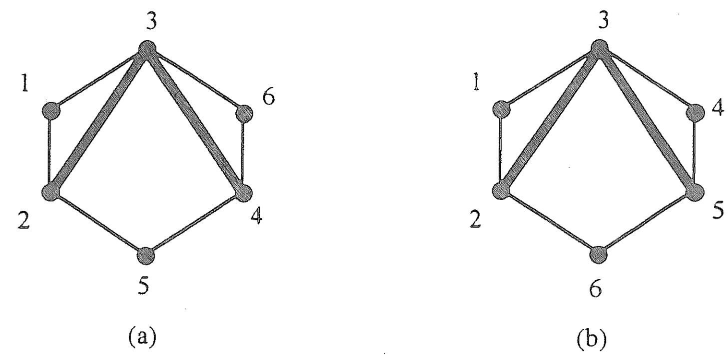

Figure 6: Illustration of potential "problem" cases

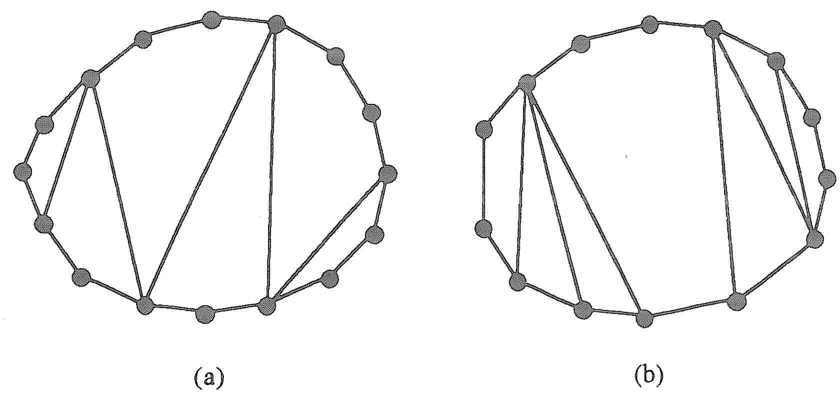

Figure 7: General form of some easy and difficult cases

7(b) are easy; the previous labeling strategy works as before.

Returning to a point raised at the outset, it was remarked that the general difficulty of OLA was well known. In fact, the (recognition) problem remains NP-Complete even when instances are restricted to the class of 2-*outerplanar* graphs (refer to Frederickson and Hambrusch (1988)). Crudely stated, a graph G is k-outerplanar if there exists an embedding of G having disjoint cycles properly nested at most k deep (*cf.* Baker (1983)). In this regard, outerplanar graphs are simply "1-outerplanar" and it is easy to see that all Halin graphs are 2-outerplanar. On the other hand, series-parallel graphs exist which are k-outerplanar for arbitrarily high k. The graph in Figure 8 makes the point.

Regarding the notion of problem complexity, the present work has also exposed a particularly interesting albeit slightly modified version of the primary problem. For ease, we might call this version the *partial OLA problem* where now we assume that as part of the instance, some (possibly empty) subset of vertices have been labeled and the aim is to map the remaining labels (from $\{1, 2, \ldots, n\}$) to the other vertices and to do so in an optimal way overall. Indeed, it is not clear that even for graph classes where OLA is solved, that this modified version would submit as well. In fact, for other problems, we know that analogous "completion" problems are, in fact, hard. Classic in this regard is the so-called *4-color completion problem* on planar graphs. Well known, of course, is that 4-colorability is decidable (trivially) on planar graphs; however, if vertices (of a planar graph) are preassigned any of at most four colors, deciding if the remaining vertices of the graph can be properly colored using no more than four colors overall is NP-Complete. The proof of this is only an exercise; the reduction is from planar graph 3-colorability. Nonetheless, it seems worthwhile to consider the aforementioned partial OLA. Presently, its status on even primitive graph classes such as paths is not clear to us.

Finally, one might consider the following problem: Given a graph G, find a smallest, spanning connected subgraph of G say H having the property that the optimal linear arrangement of H is also optimal for G. Phelps (1995) has referred to this as a "critical subgraph" version of OLA. Clearly, the problem is well-defined in that every graph exhibits a candidate subgraph–namely, the graph itself. More importantly, it is easy to see that the problem possesses interest. Consider Figure 9. Assuming G to be the graph on the left, it is easy to see that a critical subgraph results in the form of H_1 with its labeling as shown to the right. Alternately, the subgraph H_2 must be labeled as indicated which, of course, is not optimal for G.

The observant reader will be quick to recognize that the results of this note also produces an easy, critical subgraph outcome. Indeed, if G is a Halin graph with tree component, a caterpillar, then the following is immediate:

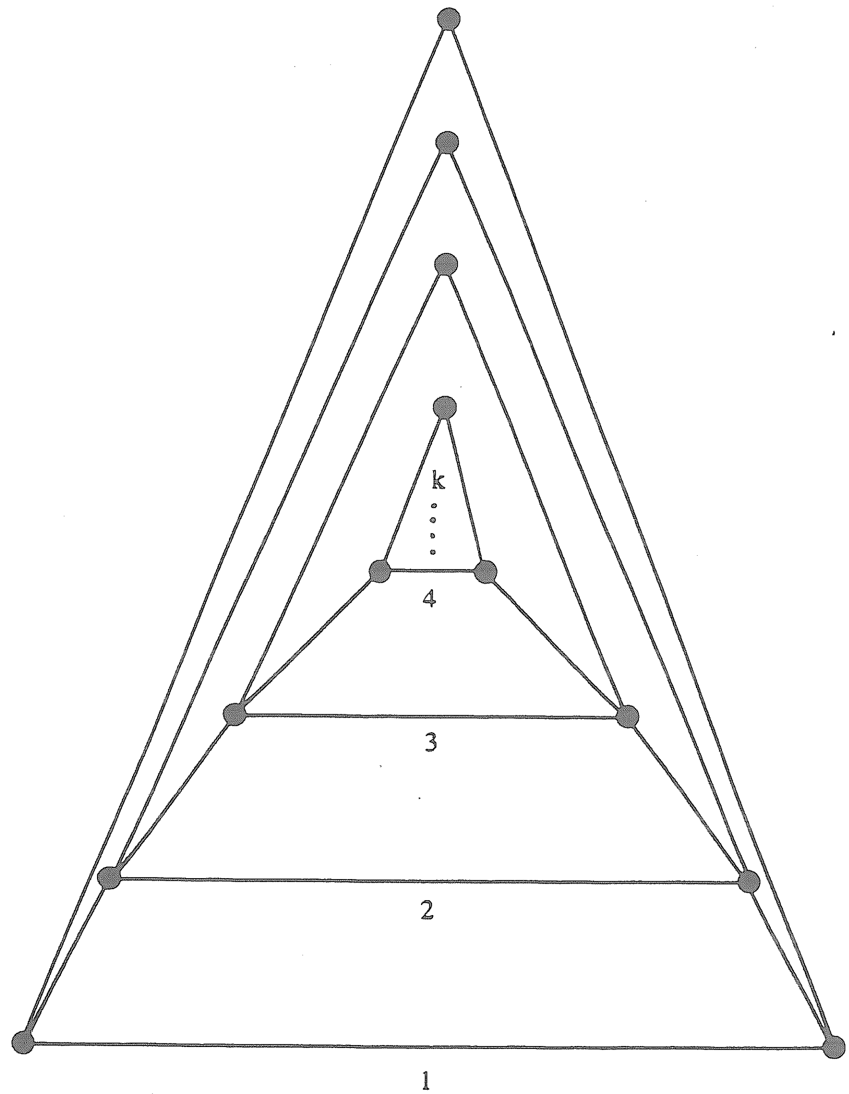

Figure 8: Series-parallel graph that is k-outerplanar

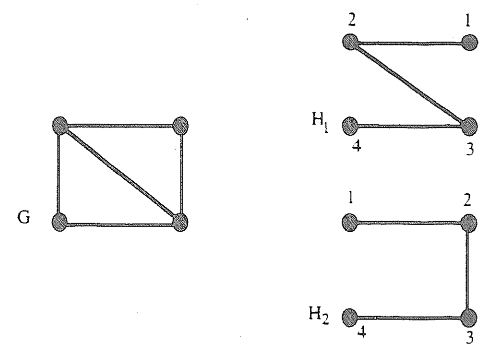

Figure 9: Critical subgraph concept

Corollary 5: Let G be a Halin graph with tree component T a caterpillar. Then T a critical subgraph of G.
□

5 REFERENCES

1. Baker, B., "Approximation Algorithms for NP-Complete Problems on Planar Graphs," unpublished manuscsript (1983).
2. Borie, R. B., R. G. Parker and C. A. Tovey, "Automatic Generation of Linear-Time Algorithms from Predicate Calculus Descriptions of Problems on Recursively Constructed Graph Families," *Algorithmica*, 7, 555-581, 1992.
3. Chung, F. R. K., "Some Problems and Results in Labelings of Graphs," in *The Theory and Applications of Graphs*, G. Chartrand, Y. Alavi, D. Goldsmith, L. Lesniak-Foster, and D. Lick, eds., John Wiley & Sons, New York, 255-264, 1981.
4. Chung F. R. K., "Oon Optimal Linear Arrangements of Trees," *Comp. and Math. with Appls.*, 10, 43-60, 1984.
5. Frederickson, G. N. and S. E. Hambrusch, "Planar Linear Arrangements of Outerplanar Graphs," *IEEE Transactions on Circuits and Systems*, 35, 323-332, 1988.
6. Phelps, K., personal communication, 1995.
7. Shiloach, Y., "A Minimum Linear Arrangement Algorithm for Undirected Trees," *SIAM J. Comp.*, 8, 15-32, 1979.