

# The Minimum Linear Arrangement Problem on Proper Interval Graphs

Ilya Safro\*

August 2, 2006

## Abstract

We present a linear time algorithm for the minimum linear arrangement problem on proper interval graphs. The obtained ordering is a 4-approximation for general interval graphs.

## 1 Preliminaries

Let  $\mathcal{F}$  be a family of nonempty sets. The *intersection graph* of  $\mathcal{F}$  is obtained by representing each set in  $\mathcal{F}$  by a vertex and connecting two vertices by an edge if and only if their corresponding sets intersect. The intersection graph of a family of intervals on a linearly ordered set (like the real line) is called an *interval graph*. If these intervals are constructed such that no interval properly contains another then such graph is called a *proper interval graph*. The families of interval and proper interval graphs are widely studied and used in different fields. In this chapter we present an algorithm which produces an optimal solution of the MinLA on proper interval graphs.

Let us construct graph  $G = (V, E)$  in a following way (algorithm  $A$ ):

- set  $n$  as number of vertices in a graph
- drop  $n$  vertices on an axis with integer coordinates from 1 to  $n$
- take a subset of successive vertices and make a clique from them
- return to the previous step  $t$  times

As a result of this construction we obtain a graph with the representation like on Figure 1. If we have a situation with nested cliques, we can ignore the clique that is placed inside of some other clique. We solve the problem for a family of graphs obtained by applying the algorithm  $A$  and then show that there is an algorithm which produces such representation for proper interval graphs.

In the following claims we will work with a graph  $G = (V, E)$  that is a chain of  $k$  cliques  $C_1 \dots C_k$  constructed using algorithm  $A$ . In all following orders we index the vertices from 1 to  $n$ , where  $|V| = n$ .

The orders of the vertices that preserve the order of cliques  $C_1, C_2, \dots, C_k$  and full rotation  $C_k, C_{k-1}, \dots, C_1$  will be called 'natural orders' (or  $N$ -order) of  $G$ .

Denote by  $d_v$  the degree of vertex  $v$ .

---

\*This work is done as a part of M.Sc. thesis [3]

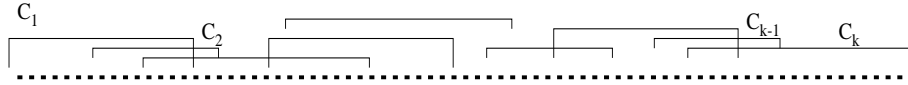


Figure 1: Schematic example of the chain of cliques

## 2 The Algorithm

Let us formulate and prove two claims that will serve as the basis of induction for Claim 2.3.

Claims 2.1 and 2.2 refer to a graph  $G = (V, E)$  that is a union of two cliques  $C_1$  and  $C_2$  whose intersection is not empty and no clique contains the other (all other cases of union of two cliques are trivial).

**Claim 2.1** *Given a graph  $G = (V, E)$ . If  $G$  is a union of two cliques  $C_1$  and  $C_2$  then in every optimal linear order of  $G$  the last vertex will not be from the intersection of  $C_1$  and  $C_2$ .*

**Proof:** Take any optimal order  $\varphi$ . Assume that the claim is false and the last vertex  $u$  in optimal arrangement  $\varphi$  comes from the intersection. Let us take vertex  $v$  that is not in the intersection and  $\varphi(v) = \max_{x \notin (C_1 \cap C_2)} \varphi(x)$ . Suppose that  $\varphi(v) = n - k$  and w.l.o.g.  $v \in C_1 \setminus (C_1 \cap C_2)$ . Exchange  $u$  and  $v$ . Since

$$N(v) \subset N(u),$$

by exchanging  $u$  and  $v$ , the connections to  $N(v)$  do not increase the cost of  $\varphi$  after the flip. However  $u$  is connected to the vertices from  $C_2 \setminus (C_1 \cap C_2)$  too. By moving  $u$  to the position  $n - k$  we reduce the cost of these edges, and does not change the cost of edges from  $u$  to the vertices at  $[n - k + 1, \dots, n - 1]$  positions (because they are from  $C_1 \cap C_2$ ). So  $u \rightleftharpoons v$  flip must reduce the cost of the optimal arrangement and this is a contradiction to the assumption.

**Claim 2.2** *Given a graph  $G = (V, E)$ . If  $G$  is an union of two cliques  $C_1$  and  $C_2$  then every optimal linear order of  $G$  has the following structure*

$$\langle \{v \mid v \in C_1 \setminus (C_1 \cap C_2)\}, \{u \mid u \in (C_1 \cap C_2)\}, \{w \mid w \in C_2 \setminus (C_1 \cap C_2)\} \rangle$$

or

$$\langle \{v \mid v \in C_2 \setminus (C_1 \cap C_2)\}, \{u \mid u \in (C_1 \cap C_2)\}, \{w \mid w \in C_1 \setminus (C_1 \cap C_2)\} \rangle$$

Call the order of this type  $N_{2C}$  - order.

**Proof:** Let us prove the claim by induction on a number of vertices in  $G$ .

- The basis of induction is when  $|V| = 1$  and the claim in this case is true.
- Suppose that if  $|V| = n - 1$  the claim is true.
- Let  $|V| = n$ . Take some optimal order  $\varphi$  of  $G$ . Suppose, that w.l.o.g. the last vertex in  $\varphi$  is  $v \in C_1 \setminus (C_1 \cap C_2)$  (as was proved in Claim 2.1). Remove  $v$  from  $G$  with all adjacent edges. We obtain a new graph  $G'$  with  $n - 1$  vertices. Look at its arrangement  $\psi$  such that  $\psi(u) = \varphi(u)$ . There are two possible cases :  $\psi$  is optimal and  $\psi$  is not optimal.

If  $\psi$  is optimal then, by induction hypothesis, it has  $N_{2C}$  - order type and, after returning  $v$  to  $G$ , we save an  $N_{2C}$  - order type. Since we started with the optimal order then  $\varphi$  has  $N_{2C}$  - order type.

Assume that  $\psi$  is not optimal. Then there exists some optimal order  $\rho$  for  $G'$  and  $C_{G',\psi} > C_{G',\rho}$ . Thus

$$C_{G,\varphi} \geq C_{G',\psi} + \alpha > C_{G',\rho} + \frac{d_v(d_v + 1)}{2}$$

Recall that  $\varphi$  is optimal for  $G$ . But  $C_{G',\rho}$  produces a cost of linear order which is  $N_{2C}$  - order on  $G'$  and, by adding vertex  $v$ , we save  $N_{2C}$  - order increasing  $C_{G',\rho}$  by  $\frac{d_v(d_v+1)}{2}$ . So we obtain a strict inequality for given optimal order  $\varphi$  and this is a contradiction. Thus  $\psi$  cannot be not optimal.

Let us switch to the main claim.

**Claim 2.3** *Given a graph  $G = (V, E)$  constructed using algorithm A then every minimum linear order of  $G$  is an N-order.*

**Proof:**

Suppose that after removing all nested cliques, the graph will remain with  $k$  different cliques. We can assume that  $G$  is connected, otherwise the problem can be divided into the similar subproblems per connected component. Let us prove it by induction on  $k$ .

- When  $k = 0$  or  $1$  the claim is trivial and when  $k = 2$  we prove it in claims 2.1 and 2.2.
- Suppose that the claim is true for  $k - 1$  and let us prove it for  $k$  cliques chain. Here we use the second induction on the number of vertices in  $T = C_k \setminus \bigcup_{i=1}^{k-1} C_i$  (call it  $l$ ).
  1. When  $l = 0$  there are  $k - 1$  cliques in  $G$  and the claim is true by induction hypothesis.
  2. If  $l = 1$  there exist unique vertex  $v$  in  $T$ . We will call vertex  $u \in G$   $p$ -vertex if it belongs to the intersection of  $C_k$  with some other clique. Assume that we have some optimal order  $\phi$ . If  $\phi$  is N-order then the claim is true. Suppose that  $\phi$  is not an N-order. Then there exist two possible cases of  $\phi$ 's structure depicted in Figure 2.

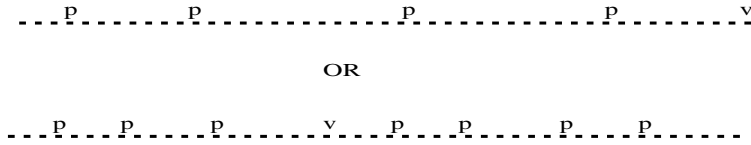


Figure 2: Possible cases of structure of  $\phi$

**First case :** If  $\phi(v) > \phi(p_i)$  (or less) for any  $i$  then

$$C_{G,\phi} = C_{G \setminus v, \phi} + \alpha$$

where  $\alpha \geq \frac{d_v(d_v+1)}{2}$ . From the other hand we have

$$C_{G, N\text{-order}} = C_{G \setminus v, N\text{-order}} + \frac{d_v(d_v + 1)}{2} \quad (1)$$

Then

$$C_{G \setminus v, \phi} + \alpha \leq C_{G \setminus v, N\text{-order}} + \frac{d_v(d_v + 1)}{2}$$

$$C_{G \setminus v, \phi} + \alpha - \frac{d_v(d_v + 1)}{2} \leq C_{G \setminus v, N\text{-order}}$$

So we obtain that  $\phi$  is at least good at  $G \setminus v$  as  $N\text{-order}$ . The situation when it is better is impossible by induction hypothesis and if it is not better then it is exactly  $N\text{-order}$  and then on  $G$  it will be also  $N\text{-order}$ .

**Second case :**  $v$  is placed between  $p$ -vertices. Suppose w.l.o.g. that there are  $L$   $p$ -vertices  $p_i$  such that

$$\forall p_i \phi(p_i) < \phi(v), \quad 1 \leq i \leq L$$

and  $R$   $p$ -vertices  $q_i$  such that

$$\forall q_i \phi(q_i) > \phi(v), \quad 1 \leq i \leq R$$

then

$$C_{G, \phi} \geq C_{G \setminus v, \phi} + \frac{R(R + 1)}{2} + \frac{L(L + 1)}{2} + LR.$$

From the other hand we have equation 1. Combining them we obtain

$$\begin{aligned} C_{G \setminus v, N\text{-order}} + \frac{d_v(d_v + 1)}{2} &= \\ &= C_{G \setminus v, N\text{-order}} + \frac{(L + R)(L + R + 1)}{2} \geq \\ &\geq C_{G \setminus v, \phi} + \frac{R(R + 1)}{2} + \frac{L(L + 1)}{2} + LR \quad (2) \end{aligned}$$

and then

$$C_{G \setminus v, N\text{-order}} \geq C_{G \setminus v, \phi}. \quad (3)$$

If inequality 3 is strict, this is a contradiction to the induction hypothesis and if there is an equality,  $\phi$  should be an  $N\text{-order}$  by induction. However, placing  $v$  in the middle of the  $N\text{-order}$  cannot allow the optimal order and this is a contradiction too.

3. Suppose the claim is true for  $|T| = l - 1$  and now we need to prove it for  $l$ . Let  $v$  be the  $l$ -th vertex in  $T$ . Assume that  $\phi$  is some optimal order. If  $\phi$  is an  $N\text{-order}$  then the claim is true. Suppose that  $\phi$  is not an  $N\text{-order}$ . If  $v$  is the last (or first, the same proof in this case) vertex in  $\phi$  then

$$C_{G, \phi} = C_{G \setminus v, \phi} + \alpha$$

where  $\alpha \geq \frac{d_v(d_v + 1)}{2}$ . From the other hand we have

$$C_{G, N\text{-order}} = C_{G \setminus v, N\text{-order}} + \frac{d_v(d_v + 1)}{2}$$

Then

$$C_{G \setminus v, \phi} + \alpha \leq C_{G \setminus v, N\text{-order}} + \frac{d_v(d_v + 1)}{2}$$

$$C_{G \setminus v, \phi} + \alpha - \frac{d_v(d_v + 1)}{2} \leq C_{G \setminus v, N\text{-order}}$$

So we obtain that  $\phi$  is at least good at  $G \setminus v$  as  $N\text{-order}$ . If  $\phi$  is better then it is impossible by the induction hypothesis. Otherwise, if it is not then  $\phi$  is exactly  $N\text{-order}$  and then on  $G$  it will be also  $N\text{-order}$ .

Suppose that  $v$  is not last (or first) and no other vertex from  $T$  is (in this case we can flip them and obtain the same optimal order). Then there are two indexes  $i$  and  $j$  s.t.  $1 < i < j < n$ , where  $i$  is a first occurrence of some vertex from  $T$  in  $\phi$  and  $j$  is a last. Let us flip  $v$  with  $j$ -th vertex and call the new order  $\phi'$ . The cost of  $\phi'$  is remained the same as it was in  $\phi$  because the flipped vertices have the same set of neighbors. Like in a previous case, there are  $L$   $p$ -vertices to the left of  $v$  and  $R$  to the right. If  $R = 0$ , we remove  $v$  like in the case of  $l = 1$  and obtain a contradiction to the induction on  $l$ . Otherwise, if  $R > 0$  then

$$C_{G, \phi'} \geq C_{G \setminus v, \phi'} + \frac{R(R+1)}{2} + \frac{(l+L-1)(L+l)}{2} + (L+l-1)R.$$

From the other hand we have equation 1. Combining them we obtain

$$\begin{aligned} C_{G, N\text{-order}} &= C_{G \setminus v, N\text{-order}} + \frac{d_v(d_v + 1)}{2} = \\ &= C_{G \setminus v, N\text{-order}} + \frac{(l-1+L+R)(l+L+R)}{2} \geq \\ &\geq C_{G, \phi'} \geq C_{G \setminus v, \phi'} + \frac{R(R+1)}{2} + \frac{(l+L-1)(L+l)}{2} + (L+l-1)R \end{aligned} \tag{4}$$

This case is similar to the case of  $l = 1$  but while removing  $v$  we use the induction on  $l$  and the claim is true.

Now we can define a process of calculation of the MinLA on proper interval graphs in polynomial time.

**Claim 2.4** *The minimum linear order of proper interval graph is calculated in polynomial time.*

**Proof:** Proper interval graphs can be recognized by a linear order: their vertices can be linearly ordered such that the vertices contained in the same clique are consecutive. The recognition works in linear time and produces such order [2]. Graphs with such possible order may be created with algorithm  $A$ . Then the minimum linear arrangement of proper interval graphs is calculated in polynomial time.

Thus, the algorithm for calculation of the MinLA is following :

1. Apply any polynomial algorithm for proper interval graph recognition which produces an  $N\text{-order}$  of vertices,
2. This order is a minimum linear arrangement.

### 3 4-Approximation for interval graphs

The complexity of the MinLA problem on general interval graphs is not known. The order of vertices by their start (or finish) time is not optimal already on the example at Figure 3. In this chapter we introduce 4-approximation algorithm for

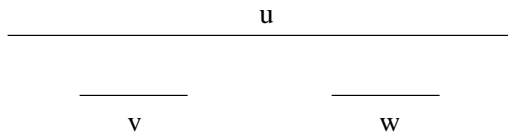


Figure 3: Counterexample of interval graph.

the problem on interval graphs. Before we will start with the algorithm let us look at some preliminary facts :

**Theorem 3.1 (Gilmore-Hoffman [1964])** *The following are equivalent :*

1.  $G$  is an interval graph.
2. It is possible to order the maximal cliques in  $G$  so that for every  $v \in V$  the cliques that contain  $v$  appear consecutively in this order.

The polynomial algorithms that construct such orders are introduced in many sources (for example in [1]) given an interval graph (possibly with no interval representation). Suppose we have such clique order  $\phi$

$$C_1, C_2, \dots, C_k.$$

Then for every vertex  $v \in V$  it is possible to define a segment  $[C_i, \dots, C_{i+l}]$  in  $\phi$  that contains the maximal cliques of  $v$ . Let us call the corresponding  $i$  and  $i+l$  by  $s_v$  and  $f_v$  respectively (start and finish). Now we can order the vertices of the graph by their  $s_v$  values (call the order by  $\pi$  and ordered vertices by  $v_1, \dots, v_n$ ).

**Theorem 3.2** *Linear order  $\pi$  is a 4-approximation for the Minimum Linear Arrangement problem on general interval graphs.*

**Proof:** Let us estimate the cost of arbitrary  $\pi$ -order. We call an incident to  $u$  edge  $e = (uv)$  in order  $\pi$  – “right oriented edge” if

$$\pi(v) > \pi(u).$$

The similar definition will be for “left oriented edges”. Every vertex in  $\pi$  has “right oriented edges” and/or “left oriented edges” as well. Every edge in  $\pi$  is right oriented and left oriented for its different ends. We will pass vertex by vertex in order  $\pi$  and estimate the total cost of right oriented edges for every vertex and this will give the estimation of the total cost of the order  $\pi$ . Denote by  $R_j$  the set of right oriented edges of vertex  $v_j$  in order  $\pi$ .

$$C_{G,\pi} = \sum_{v_i \in V} \sum_{e \in R_i} cost_\pi(e).$$

Since the vertices are ordered by their  $s_v$  values

$$\sum_{e \in R_i} cost_\pi(e) \leq \frac{d_{v_i}(d_{v_i} + 1)}{2}.$$

Thus

$$C_{G,\pi} \leq \sum_{v \in V} \frac{d_v(d_v + 1)}{2}.$$

From the other hand for any optimal order  $\psi$  the following is true :

$$C_{\psi^*} \geq \frac{\sum_v \left( 2^{\frac{d_v(d_v+1)}{2}} \right)}{2} = \frac{\sum_v \left( \frac{d_v}{2} (d_v + 1) \right)}{2}. \quad (5)$$

Combining previous results we obtain

$$A = \frac{\sum_v \left( \frac{d_v}{2} (d_v + 1) \right)}{2} \leq C_{\psi^*} \leq C_\pi \leq \sum_v \frac{d_v(d_v + 1)}{2} = B. \quad (6)$$

Clearly that for  $\alpha = 4$

$$B \leq \alpha A$$

and this proves that any  $\pi$ -order gives 4-approximation for the Minimum Linear Arrangement problem on general interval graphs.

Easy to see that given some interval representation of interval graph, the  $\pi$ -order is equivalent to the order by start (or finish) time of the intervals.

## References

- [1] M. Golumbic, *Algorithmic graph theory and perfect graphs*, Academic Press, 1980.
- [2] Celina M. Herrera de Figueredo, Joao Meidanis and Celia Picinin de Mello. *A lexBFS Algorithm for Proper Interval Graph Recognition*, Instituto de Computago, UNICAMP, Technical Report DCC-04/93, 1993.
- [3] I. Safro, *M.Sc. Thesis. The Minimum Linear Arrangement Problem*, The Weizmann Institute of Science, 2002. <http://www.wisdom.weizmann.ac.il/~safro/thesis.ps>