

Precoloring extension on unit interval graphs*

Dániel Marx[†]

23rd February 2004

Abstract

In the precoloring extension problem we are given a graph with some of the vertices having a preassigned color and it has to be decided whether this coloring can be extended to a proper k -coloring of the graph. Answering an open question of Hujter and Tuza [6], we show that the precoloring extension problem is **NP**-complete on unit interval graphs.

1 Introduction

In graph vertex coloring we have to assign colors to the vertices such that neighboring vertices receive different colors. In the *precoloring extension* problem a subset W of the vertices have a preassigned color and we have to extend this to a proper k -coloring of the whole graph. Formally, we will investigate the following problem:

Precoloring Extension

Input: A graph $G(V, E)$, a subset $W \subseteq V$, a coloring c' of W and an integer k .

Question: Is there a proper k -coloring c of G extending the coloring c' (that is, $c(v) = c'(v)$ for every $v \in W$)?

Since vertex coloring is the special case $W = \emptyset$, the precoloring extension problem is **NP**-complete in every class of graphs where vertex coloring is **NP**-complete. Therefore we can hope to solve precoloring extension efficiently only on graphs that are easy to color, for example on perfect graphs. Biró, Hujter and Tuza [1, 5, 6] started a systematic study of precoloring extension in perfect graphs. It turns out that for some classes of perfect graphs (e.g., split graphs, complements of bipartite graphs, cographs) not only coloring is easy, but even the more general precoloring extension problem can be solved in polynomial time. On the other hand, for some other classes (bipartite graphs, line graphs of bipartite graphs) precoloring extension is **NP**-complete.

A graph is an *interval graph* if it can be represented as the intersection graph of a set of intervals. It is a *unit interval graph* if it can be represented by intervals of unit length and it is a *proper interval graph* if it can be represented in such a way that no interval is properly contained in another. It can be shown that

*Research is supported by grants OTKA 44733, 42559 and 42706 of the Hungarian National Science Fund.

[†]Department of Computer Science and Information Theory, Budapest University of Technology and Economics, H-1521 Budapest, Hungary. dmarx@cs.bme.hu

these two latter classes of graphs are the same [3], in fact they are exactly the interval graphs that are claw-free [11] (contain no induced $K_{1,3}$). These interval graphs are also called *indifference graphs*.

Interval graph coloring arises in various applications including scheduling [2] and single row VLSI routing [10]. There is a simple greedy algorithm that colors an interval graph with minimum number of colors. However, Biró, Hujter and Tuza [1] proved that the precoloring extension problem is **NP**-complete on interval graphs, even if every color is used at most twice in the precoloring (they also gave a polynomial time algorithm for the case where every color is used only once). In [6] they asked what is the complexity of the precoloring extension problem in the more restricted case of unit interval graphs. In Section 3, we prove that this problem is also **NP**-complete. The proof is by reduction from a disjoint paths problem whose **NP**-completeness was proved in [7]. In Section 2, we briefly overview the relevant definitions and results concerning the disjoint paths problem.

2 The disjoint paths problem

In the disjoint paths problem we are given a graph G and a set of source–destination pairs $(s_1, t_1), (s_2, t_2), \dots, (s_k, t_k)$ (called the *terminals*), our task is to find k disjoint paths P_1, \dots, P_k such that path P_i connects vertex s_i to vertex t_i . There are four basic variants of the problem: the graph can be directed or undirected, and we can require edge disjoint or vertex disjoint paths. Here only the directed, edge disjoint problem is considered, 'disjoint' will mean edge disjoint throughout this paper. The problem is often described in terms of a supply graph and a demand graph, as follows:

Disjoint Paths

Input: The directed *supply graph* G and the directed *demand graph* H on the same set of vertices.

Question: Find a path P_e in G for each $e \in E(H)$ such that these paths are edge disjoint and path P_e together with edge e form a directed circuit.

With a slight abuse of terminology, we say that a demand $\vec{uv} \in H$ starts in v and ends in u (since the directed path satisfying this demand starts in v and ends in u). An undirected graph is called *Eulerian* if every vertex has even degree, and a directed graph is Eulerian if the indegree equals the outdegree at every vertex. The disjoint paths problem is motivated by practical applications and the deep theory behind it, see [13, 4] for a survey of the results in this area.

A graph is a *grid graph* if it is a finite subgraph of the rectangular grid. A directed grid graph is a grid graph with the horizontal edges directed to the right and the vertical edges directed to the bottom. Clearly, every directed grid graph is acyclic. A *rectangle* is a grid graph with $n \times m$ nodes such that $v_{i,j}$ ($1 \leq i \leq n$, $1 \leq j \leq m$) is connected to $v_{i',j'}$ if and only if $|i - i'| = 1$ and $j = j'$, or $i = i'$ and $|j - j'| = 1$. The study of grid and rectangle graphs is motivated by applications in VLSI-layout.

The undirected edge disjoint paths problem is **NP**-complete even in the special case when G is planar (or even if $G + H$ is planar [8]). Vygen proved that the directed edge disjoint paths problem is **NP**-complete even if the supply graph G is planar and acyclic [14] or even if G is a directed grid graph [13]. In

[7] it is shown that the problem remains **NP**-complete if G is a directed grid graph and $G + H$ is Eulerian.

It is noted in [13] that the disjoint paths problem is not easier in rectangle graphs than in general grid graphs: if we add a new edge \overrightarrow{uv} to G and a new demand from u to v , then the new demand can reach v in the grid only using the new edge. Thus, without changing the solvability of the problem, we can add new edges and demands until G becomes a full rectangle. Notice that $G + H$ remains Eulerian after adding these new edges and new demands.

Theorem 2.1. *The disjoint paths problem is **NP**-complete on directed rectangles even if $G + H$ is Eulerian.*

The following observation will be useful:

Lemma 2.2. *In the directed case, if $G + H$ is Eulerian, and G is acyclic, then every solution of the disjoint paths problem uses all the edges of G .*

Proof. Assume that a solution is given. Take a demand edge of H and delete from $G + H$ the directed circuit formed by the demand edge and its path in the solution. Continue this until the remaining graph contains no demand edges, then it is a subgraph of G . Since we deleted only directed circuits, it remains Eulerian, but the only Eulerian subgraph of G is the empty graph, thus the solution used all the edges. ■

For purely technical reasons, we introduce the following variant of the disjoint paths problem. For every demand, not only the terminals are given, but here also the first and last edge of the path is also prescribed:

Directed Edge Disjoint Paths with Terminal Edges

Input: The supply graph G and the demand graph H on the same set of vertices (both of them directed), and for every edge $e \in H$, a pair of edges (s_e, t_e) of G .

Question: Find a path P_e in G for each $e \in E(H)$ such that these paths are edge disjoint, P_e and e form a directed circuit and the first/last edge of P_e is s_e, t_e , respectively.

As shown in the following theorem, this variant of the problem is **NP**-complete as well. It will be the basis of the reduction in Section 3.

Theorem 2.3. *The Directed Edge Disjoint Paths with Terminal Edges problem is **NP**-complete on rectangle graphs, even when $G + H$ is Eulerian.*

Proof. It is shown in [7] that the disjoint paths problem is **NP**-complete on directed grid graphs with $G + H$ Eulerian. The reduction in [7] constructs grid graphs with the following additional properties:

- at most one demand ends in each vertex v ,
- if a demand ends in v , then exactly one edge of G enters v ,
- at most two demands start in each vertex u ,
- if a demand starts in u , then no edge of G enters u .

If two demands α and β start at a vertex u , then we slightly modify G and H . Two new supply edges \overrightarrow{xu} and \overrightarrow{yu} are attached to u , there is place for these edges since no edge enters u in G . Demand graph H is modified such that the start vertex of demand α is set to x , the start vertex of β is set to y . Clearly, these modifications do not change the solvability of the instance, and G remains a grid graph. Moreover, $G + H$ remains Eulerian. Therefore we can assume that the instance has the following two properties as well:

- At most one demand starts from each vertex u ,
- If a demand starts in u , then exactly one edge of G leaves u .

If these properties hold, then in every solution of the disjoint paths problem a demand going from u to v has to leave u on the unique edge leaving u , and has to enter v on the unique edge entering v . Therefore prescribing the first and the last edge of every demand does not change the problem. Thus we can conclude that the disjoint paths with terminal edges problem is **NP**-complete in grid graphs. Furthermore, when we add new edges to G and H to make G a rectangle (as described in the remark before Theorem 2.1), then obviously it can be prescribed that the first and the last edge of the new demand is the new edge, hence it follows that the problem is **NP**-complete on directed rectangles as well. ■

3 Precoloring extension

The aim of this section is to prove the **NP**-completeness of precoloring extension on proper interval graphs (recall that proper interval graphs are the same as unit interval graphs). In [1] the **NP**-completeness of precoloring extension on interval graphs is proved by a reduction from circular arc graph coloring. A similar reduction is possible from *proper* circular arc coloring to the precoloring extension of *proper* interval graphs, but the analogy doesn't help here, because proper circular arc coloring can be done in polynomial time [9, 12]. In this section, we follow a different path: the **NP**-completeness of precoloring extension on proper interval graphs is proved by reduction from a planar disjoint paths problem investigated in Section 2.

An important idea of the proof is demonstrated on Figure 1. In any k -coloring of the intervals in (a), for all i , interval $I_{1,i}$ has the same color as $I_{0,i}$: interval $I_{1,0}$ must receive the only color not used by $I_{0,1}, I_{0,2}, \dots, I_{0,k-1}$; interval $I_{1,1}$ must receive the color not used by $I_{1,0}, I_{0,2}, I_{0,3}, \dots, I_{0,k-1}$, and so on. In case (b), the intervals are slightly modified. If all the $I_{0,i}$ intervals are colored, then there are two possibilities: either the color of $I_{1,i}$ is the same as the color of $I_{0,i}$ for $i = 0, \dots, k-1$, or we swap the colors of $I_{1,1}$ and $I_{1,2}$. Blocks of type (a) and (b) will be the building blocks of our reduction.

Theorem 3.1. *The precoloring extension problem is **NP**-complete on proper interval graphs.*

Proof. The reduction is from the Eulerian directed edge disjoint paths with terminal edges problem on rectangle graphs, whose **NP**-completeness was shown in Theorem 2.3. First we modify the given rectangle graph G . As in the remark before Theorem 2.1, new edges are added to the rectangle $KLMN$ to obtain the shape shown on Figure 2, without changing the solvability of the instance. Hereinafter it is assumed that G has such a form. The entire G is contained

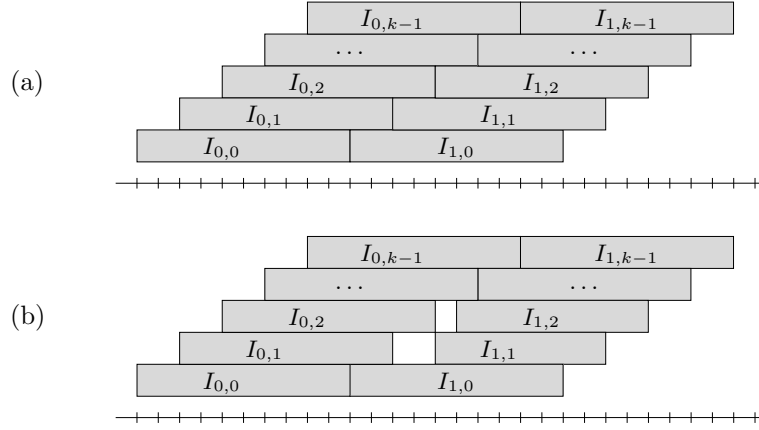


Figure 1: (a) In every k -coloring c of the (open) intervals, $I_{0,i}$ and $I_{1,i}$ receive the same color for $0 \leq i \leq k-1$ (b) In every k -coloring c of the intervals, $c(I_{0,i}) = c(I_{1,i})$ for $i \neq 1, 2$ and either $c(I_{0,1}) = c(I_{1,1})$, $c(I_{0,2}) = c(I_{1,2})$ or $c(I_{0,1}) = c(I_{1,2})$, $c(I_{0,2}) = c(I_{1,1})$ holds.

between the two diagonal lines X and Y , the vertices on X have outdegree 1, the vertices on Y have indegree 1 and the vertices of G between X and Y are Eulerian. If the rectangle $KLMN$ contains $r \times s$ vertices, then there are $m = r + s$ vertices on both X and Y , and every directed path from a vertex of X to a vertex of Y has length m . Now consider the parallel diagonal lines A_0, A_1, \dots, A_{m-1} as shown on the figure, and denote by E_i the set of edges intersected by A_i . Clearly this forms a partition of the edges, and every set E_i has size m . Let $E_i = \{e_{i,0}, \dots, e_{i,m-1}\}$, ordered in such a way that $e_{i,0}$ is the lower left edge.

We can assume that H is a DAG, otherwise there would be no solution, since G is acyclic. Exactly one demand starts from each vertex on line X , exactly one demand terminates at each vertex on Y , and the indegree equals the outdegree in every other vertex of H , this follows from $G + H$ Eulerian. From these facts, it is easy to see that H can be decomposed into m disjoint paths D_1, \dots, D_m such that every path connects a vertex on X with a vertex on Y . We assign a color to each demand: if demand α is in D_i , then give the color i to α .

Based on the disjoint paths problem, we define a set of intervals and a precoloring. Every interval $I_{i,j}$ corresponds to an edge $e_{i,j} \in E_i$ of the supply graph G . Let $v_{i,j}$ be the tail vertex of $e_{i,j}$, and denote by $\delta_G(v_{i,j})$ the outdegree of $v_{i,j}$ in G . The intervals $I_{i,j}$ ($0 \leq i \leq m-1$, $0 \leq j \leq m-1$) are defined as follows (see Figure 3):

$$I_{i,j} = \begin{cases} (2(im+j), 2(im+j)+2m) & \text{if } \delta_G(v_{i,j}) = 1, \\ (2(im+j)+2, 2(im+j)+2m) & \text{if } \delta_G(v_{i,j}) = 2 \text{ and } e_{i,j} \text{ is vertical,} \\ (2(im+j)+1, 2(im+j)+2m) & \text{if } \delta_G(v_{i,j}) = 2 \text{ and } e_{i,j} \text{ is horizontal.} \end{cases}$$

Notice that the intervals are open, hence two intervals that share only an endpoint do not intersect.

If the prescribed start edge and end edge of a demand with color c is e' and e'' , then precolor the intervals corresponding to e' and e'' with color c . This assignment is well defined, since it can be assumed that the start and end edges of the demands are different, otherwise it is trivial that the problem has no

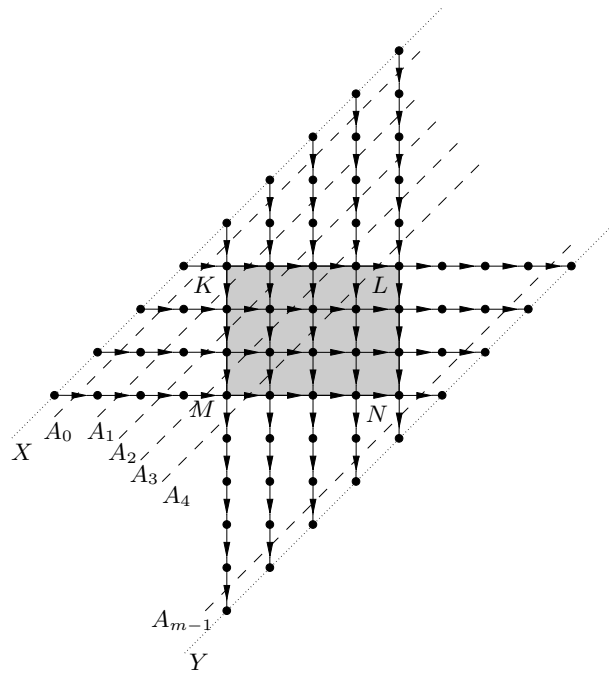


Figure 2: Partitioning the edges of the extended grid ($r = 4, s = 5$).

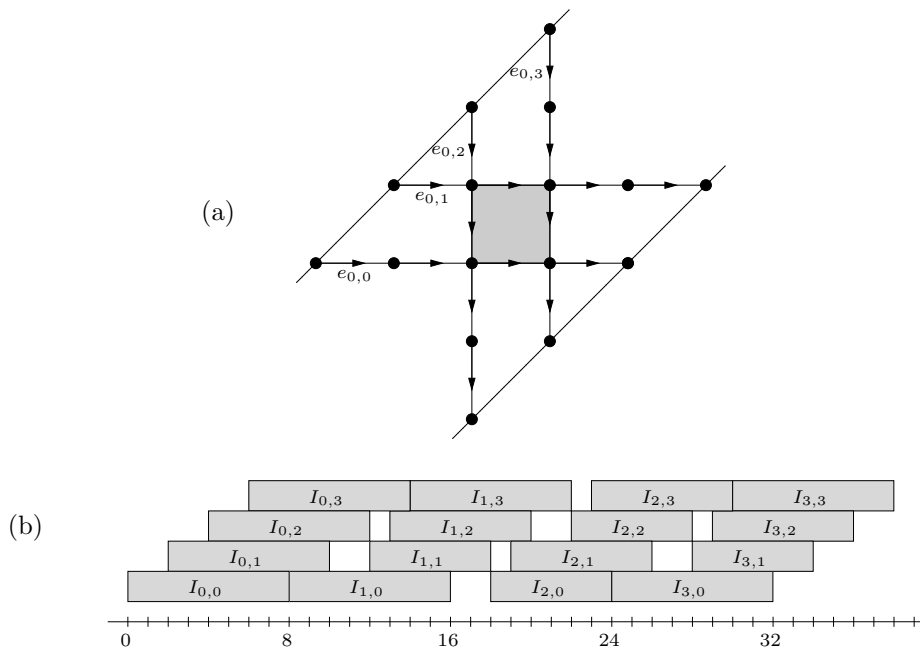


Figure 3: An example of the reduction with $r = s = 2$: (a) the grid graph, (b) the corresponding proper interval graph.

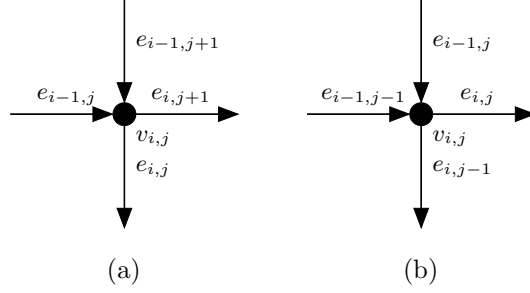


Figure 4: The edges incident to the tail of $e_{i,j}$. (a) $e_{i,j}$ is vertical (b) $e_{i,j}$ is horizontal

solution. This completes the description of the reduction, we claim that the precoloring of the constructed interval graph can be extended to a coloring with m colors if and only if the disjoint paths problem has a solution.

First we observe certain properties of the intervals. Let $I_i = \{I_{i,0}, \dots, I_{i,m-1}\}$, that is, the set of intervals corresponding to E_i . The set I_i forms a clique in the graph, and the elements of I_i and $I_{i'}$ are not intersecting if $i' \geq i + 2$. The interval $I_{i,j}$ does not intersect $I_{i-1,j'}$ for $j' \leq j$, and it does intersect $I_{i-1,j'}$ for $j' > j + 1$. It may or may not intersect $I_{i-1,j+1}$.

Assume that P_1, \dots, P_n is a solution of the disjoint paths problem. If an edge $e_{i,j}$ is used by a demand with color c , then color the edge $e_{i,j}$ and the corresponding interval $I_{i,j}$ with color c . Since by Lemma 2.2 every edge of the graph is used by a demand, every interval receives a color. Furthermore, the demands use the prescribed start and end edges, and so this coloring is compatible with the precoloring given above. Notice that the set of edges in the grid graph that receive the color c forms a directed path from a vertex on X to a vertex on Y . Thus all m colors appear on the intervals in I_i , every interval has different color in this set.

It has to be shown that this coloring is proper. By the observations made above, it is sufficient to verify that two intersecting intervals $I_{i,j}$ and $I_{i-1,j'}$ do not have the same color. Since the edges having color c form a path, if $e_{i-1,j'}$ and $e_{i,j}$ have the same color, then the head of $e_{i-1,j'}$ and the tail of $e_{i,j}$ must be the same vertex $v_{i,j}$. Assume first that $\delta_G(v_{i,j}) = 1$, then $j = j'$, which implies that $I_{i,j}$ and $I_{i-1,j'}$ are not intersecting. For the case $\delta_G(v_{i,j}) = 2$, it will be useful to refer to Figure 4. If $\delta_G(v_{i,j}) = 2$ and $e_{i,j}$ is vertical, then $e_{i,j+1}$ is horizontal and its tail is also $v_{i,j}$ (see Figure 4a). Moreover, in this case $e_{i-1,j}$ is horizontal, $e_{i-1,j+1}$ is vertical, and $v_{i,j}$ is the head of both edges. Therefore if $\delta_G(v_{i,j}) = 2$ and $e_{i,j}$ is vertical, then $j' = j$ or $j' = j + 1$, which implies that the right endpoint of $I_{i-1,j'}$ is not greater than $2((i-1)m + j + 1) + 2m = 2(im + j) + 2$, the left endpoint of $I_{i,j}$. If $e_{i,j}$ is horizontal, then $j' = j$ or $j' = j - 1$ (see Figure 4b), hence intervals $I_{i-1,j'}$ and $I_{i,j}$ are clearly not intersecting.

On the other hand, assume that there is a proper extension of the precoloring with m colors. Color every edge $e_{i,j}$ of the grid graph with the color assigned to the corresponding interval $I_{i,j}$. First we prove that the set of edges having color c forms a directed path R_c in the graph. Since the intervals in I_i have different colors, every one of the m colors appears exactly once on the edges in E_i . Thus it is sufficient to prove that the tail $v_{i,j}$ of the unique edge $e_{i,j} \in E_i$ having color c is the same as the head of the unique edge $e_{i-1,j'} \in E_{i-1}$ having color c .

Assume first that $\delta_G(v_{i,j}) = 1$, then we have to show that $j = j'$. Denote by $x = 2(im + j)$, the left endpoint of $I_{i,j}$, which is also the right endpoint of $I_{i-1,j}$ (as an example, consider interval $I_{1,3}$ on Figure 3b). If $j' > j$, then $I_{i-1,j'}$ and $I_{i,j}$ intersect (both of them contain $x + \epsilon$), which contradicts the assumption that $I_{i,j}$ and $I_{i-1,j'}$ have the the same color. Assume therefore that $j' < j$. It is clear from the construction that the left endpoint of every interval $I_{i,1}, \dots, I_{i,j-1}$ is strictly smaller than x (it is not possible that $\delta_G(v_{i,j-1}) = 2$ and $e_{i,j-1}$ is vertical, since that would imply $v_{i,j-1} = v_{i,j}$ and $\delta_G(v_{i,j}) = 2$). The right endpoint of every interval $I_{i-1,j}, \dots, I_{i-1,m-1}$ is not smaller than x , thus $\{I_{i,0}, \dots, I_{i,j-1}, I_{i-1,j}, I_{i-1,j+1}, \dots, I_{i-1,m-1}\}$ is a clique of size m in the interval graph, since they all contain $x - \epsilon$. Now $I_{i,0}, \dots, I_{i,j-1}$ intersect $I_{i,j}$, and $I_{i-1,j}, \dots, I_{i-1,m-1}$ intersect $I_{i,j'}$, thus color c cannot appear in this clique, a contradiction.

Now assume that $\delta_G(e_{i,j}) = 2$ and $e_{i,j}$ is vertical, we have to show that $j' = j$ or $j' = j + 1$ holds (see for example $I_{1,1}$ on Figure 3b). If $j' > j + 1$, then $I_{i-1,j'}$ intersects $I_{i,j}$, a contradiction. Assume therefore that $j' < j$ and let $y = 2(im + j)$, the right endpoint of $I_{i-1,j}$. It can be verified that $\{I_{i,0}, \dots, I_{i,j-1}, I_{i-1,j}, \dots, I_{i-1,m-1}\}$ is a clique of size m , since all of them contain $y - \epsilon$. Color c cannot appear on $I_{i,0}, \dots, I_{i,j-1}$ because of $I_{i,j}$, and it cannot appear on $I_{i-1,j}, \dots, I_{i-1,m-1}$ because of $I_{i-1,j'}$. Thus there is a clique of size m without color c , a contradiction.

If $\delta_G(e_{i,j}) = 2$ and $e_{i,j}$ is horizontal, then we have to show that j' is either j or $j - 1$ (see for example $I_{3,2}$ on Figure 3b). If $j' \geq j + 1$, then $I_{i,j}$ intersects $I_{i-1,j'}$, therefore it can be assumed that $j' < j - 1$. Let $z = 2((i - 1)m + j - 1) + 2m$, the right endpoint of $I_{i-1,j-1}$. Point $z - \epsilon$ is contained in each of $I_{i,0}, \dots, I_{i,j-2}, I_{i-1,j-1}, \dots, I_{i-1,m-1}$, hence they form a clique of size m . However, intervals $I_{i-1,j'}$ and $I_{i,j}$ forbid the use of color c on this clique, a contradiction.

We have shown that the set of edges with color c are contained in a path R_c . Because of the precoloring, the path R_c goes through the prescribed start and end edges of every demand with color c . Furthermore, since the demands with color c correspond to a directed path D_c in H , all these demands can be satisfied using only the edges of R_c , without two demands using the same edge. Thus there is solution to the disjoint path problem, proving the correctness of the reduction.

Since the precoloring extension problem is obviously in **NP** and the reduction above can be done in polynomial time, we have proved that it is **NP**-complete on unit interval graphs. ■

References

- [1] M. Biró, M. Hujter, and Zs. Tuza. Precoloring extension. I. Interval graphs. *Discrete Math.*, 100(1-3):267–279, 1992.
- [2] M. Biró, M. Hujter, and Zs. Tuza. Cross fertilisation of graph theory and aircraft maintenance scheduling. In G. Davidson, editor, *Thirty/Second Annual Symposium*, pages 307–317. AGIFORS (Airline Group of the International Federation of Operation Research Societies, 1993.
- [3] K. P. Bogart and D. B. West. A short proof that “proper = unit”. *Discrete Math.*, 201(1-3):21–23, 1999.
- [4] A. Frank. Packing paths, circuits, and cuts—a survey. In *Paths, flows, and VLSI-layout (Bonn, 1988)*, pages 47–100. Springer, Berlin, 1990.

- [5] M. Hujter and Zs. Tuza. Precoloring extension. II. Graph classes related to bipartite graphs. *Acta Mathematica Universitatis Comenianae*, 62(1):1–11, 1993.
- [6] M. Hujter and Zs. Tuza. Precoloring extension. III. Classes of perfect graphs. *Combin. Probab. Comput.*, 5(1):35–56, 1996.
- [7] D. Marx. Eulerian disjoint paths problem in grid graphs is NP-complete, 2004. To appear in *Discrete Applied Mathematics*. <http://dx.doi.org/10.1016/j.dam.2003.12.003> .
- [8] M. Middendorf and F. Pfeiffer. On the complexity of the disjoint paths problem. *Combinatorica*, 13(1):97–107, 1993.
- [9] J. B. Orlin, M. A. Bonuccelli, and D. P. Bovet. An $O(n^2)$ algorithm for coloring proper circular arc graphs. *SIAM J. Algebraic Discrete Methods*, 2(2):88–93, 1981.
- [10] A. Recki. Minimax results and polynomial algorithms in VLSI routing. In *Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity (Prachatice, 1990)*, pages 261–273. North-Holland, Amsterdam, 1992.
- [11] F. S. Roberts. Indifference graphs. In *Proof Techniques in Graph Theory (Proc. Second Ann Arbor Graph Theory Conf., Ann Arbor, Mich., 1968)*, pages 139–146. Academic Press, New York, 1969.
- [12] A. Teng and A. Tucker. An $O(qn)$ algorithm to q -color a proper family of circular arcs. *Discrete Math.*, 55(2):233–243, 1985.
- [13] J. Vygen. Disjoint paths. Technical Report 94816, Research Institute for Discrete Mathematics, University of Bonn, 1994.
- [14] J. Vygen. NP-completeness of some edge-disjoint paths problems. *Discrete Appl. Math.*, 61(1):83–90, 1995.