

# Minimum Fill-in for Chordal Bipartite Graphs

T. Kloks

RUU-CS-93-11 March 1993



# **Utrecht University**

Department of Computer Science

Padualaan 14, P.O. Box 80.089, 3508 TB Utrecht, The Netherlands, Tel.: ... + 31 - 30 - 531454

			•	
			•	
			e.	
			•	

# Minimum Fill-in for Chordal Bipartite Graphs

T. Kloks

 $\frac{\text{Technical Report RUU-CS-93-11}}{\text{March 1993}}$ 

Department of Computer Science
Utrecht University
P.O.Box 80.089
3508 TB Utrecht
The Netherlands

ISSN: 0924-3275

## Minimum Fill-in for Chordal Bipartite Graphs

T. Kloks \*
Department of Computer Science
Utrecht University
P.O.Box 80.089
3508 TB Utrecht, The Netherlands

#### Abstract

Chordal bipartite graph are exactly those bipartite graph in which every cycle of length at least six has a chord. The MINIMUM FILL-IN problem is the problem of finding a chordal embedding of the graph with a minimum number of edges. We present a polynomial time algorithm for the exact computation of the minimum fill-in for all chordal bipartite graphs.

#### 1 Introduction

Chordal bipartite graphs (or weakly chordal bipartite graphs) are bipartite graphs such that every induced cycle of length at least six has a chord. The chordal bipartite graphs form a large class of perfect graphs containing for example the convex and biconvex bipartite graphs, the bipartite permutation graphs and the bipartite distance hereditary graphs (or (6, 2)-chordal bipartite graphs). Many NP-complete problems remain NP-complete when restricted to the class of chordal bipartite graphs. For example Hamiltonian cycle, Hamiltonian path, dominating set, connected dominating set, independent dominating set and steiner tree [12, 3]. Since chordal bipartite graphs are bipartite and hence perfect, the weighted version of stable set and some other problems like the clique covering problem are solvable in polynomial time [6]. Recognizing chordal bipartite graphs can be done in  $O(\min(m \log n, n^2))$  [11].

The MINIMUM FILL-IN problem is the problem of finding a chordal embedding of a graph with a minimum number of edges. The problem is sometimes also called the CHORDAL GRAPH COMPLETION problem and is of importance in relation with the performance of Gaussian elimination on sparse matrices. The minimum fill-in

<sup>\*</sup>This author is supported by the Foundation for Computer Science (S.I.O.N.) of the Netherlands Organization for Scientific Research (N.W.O.). Email: ton@cs.ruu.nl.

problem is NP-complete in general [14]. In [13] an  $O(n^5)$  time algorithm is given for the class of bipartite permutation graphs. Notice that this class of graphs is properly contained in the class of chordal bipartite graphs. In this paper we show that the minimum fill-in can be computed in  $O(n^5)$  time for all chordal bipartite graphs.

We do not claim that our algorithm is a very practical one, however we feel that it is one of the first non-trivial polynomial time algorithms for computing the minimum fill-in for a relatively large class of graphs.

#### 2 Preliminaries

In this section we start with some necessary definitions and lemmas. For more information the reader is referred to [5] or [2].

**Definition 2.1** A graph is called chordal bipartite (or weakly chordal bipartite) if it is bipartite and each cycle of length at least six has a chord.

If G = (V, E) is a graph then we use the notation N(x) for the set of neighbors of a vertex x. Chordal bipartite graphs can be characterized by the existence of a certain ordering of the edges.

**Definition 2.2** Let G = (X, Y, E) be a bipartite graph. Then  $(u, v) \in E$  is called a bisimplicial edge if  $N(x) \cup N(y)$  induces a complete bipartite subgraph of G.

**Definition 2.3** Let G = (X,Y,E) be a bipartite graph. Let  $(e_1,\ldots,e_k)$  be an ordering of the edges of G. For  $i=0,\ldots,k$  define the subgraph  $G_i = (X_i,Y_i,E_i)$  as follows.  $G_0 = G$  and for  $i \geq 1$ ,  $G_i$  is the subgraph of  $G_{i-1}$  with  $X_i = X_{i-1}$ ,  $Y_i = Y_{i-1}$  and  $E_i = E_{i-1} \setminus \{e_i\}$  (i.e. the edge  $e_i$  is removed but not the end vertices). The ordering  $(e_1,\ldots,e_k)$  is a perfect edge without vertex elimination ordering for G if each edge  $e_i$  is bisimplicial in  $G_i$ , and  $G_k$  has no edge.

The following characterization appears for example in [2].

**Lemma 2.1** G is chordal bipartite if and only if there is a perfect edge without vertex elimination ordering.

The following lemma shows that we can start a perfect edge without vertex elimination ordering with *any* bisimplicial edge.

**Lemma 2.2** Let G be chordal bipartite. Let e be a bisimplicial edge in G. Let G' be the graph obtained from G by deleting the edge e but not the endvertices of e. Then G' is chordal bipartite.

*Proof.* Assume G' has a chordless cycle C of length  $\geq$  6. Then, clearly, x and y must be elements of C. The neighbors of x and y in the cycle form a square. This shows that C cannot be chordless in G'.

In [4] it is shown that a bisimplicial edge in a chordal bipartite graph with n vertices can be found in  $O(n^2)$  time.

Corollary 2.1 We can find a perfect edge without vertex elimination scheme in time  $O(n^2m)$ , where n is the number of vertices and m is the number of edges.

This shows that chordal bipartite graphs can be recognized efficiently. If G = (X, Y, E) is a bipartite graph then we call the sets of vertices X and Y the color classes of G. A maximal complete bipartite subgraph M of a bipartite graph G is an ordered pair M = (A, B) with  $A \subseteq X$  and  $B \subseteq Y$ . By definition there is an edge between every vertex of A and every vertex of B. In this paper we call a graph M = (A, B) complete bipartite only if  $|A| \ge 2$  and  $|B| \ge 2$ . As a first important result we obtain the following (see also [10]).

**Lemma 2.3** If G = (X, Y, E) is chordal bipartite then it contains at most |E| maximal complete bipartite subgraphs.

*Proof.* Consider a perfect edge without vertex elimination ordering  $(e_1, \ldots, e_k)$ . Consider a maximal complete bipartite subgraph (A, B) and let  $e_i = (x, y)$  be the first edge in the ordering with  $x \in A$  and  $y \in B$ . Then, since  $e_i$  is bisimplicial and (A, B) is maximal, A = N(y) and B = N(x). Hence (A, B) is completely and uniquely determined by  $e_i$ . This proves the lemma.

The problem we solve in this paper is finding a triangulation of a chordal bipartite graph such that the maximum clique size is minimum.

**Definition 2.4** A graph is chordal if it has no induced chordless cycle of length at least four.

**Definition 2.5** A triangulation of a graph G is a graph H with the same vertex set as G such that G is a subgraph of H and H is chordal.

Here are two problems that have drawn much attention because of the large number of applications. The first is to find a triangulation such that the maximum clique size is as small as possible. This problem is called the TREEWIDTH problem. For more information on treewidth we refer to [9]. The second important problem related to triangulations of graphs is the so called MINIMUM FILL-IN problem. In this case one tries to find a triangulation of a graph with a minimum number of edges. Both problems are NP-complete [1, 14].

We start with the following easy but important observation.

**Lemma 2.4** If G = (A, B) is a complete bipartite graph and H is a triangulation of G then either H[A] or H[B] is a complete subgraph of H.

*Proof.* Assume there are two vertices x and y in A which are not adjacent in H and two vertices p and q in B which are also not adjacent in H. Since G is complete bipartite the four vertices  $\{x,y,p,q\}$  form a chordless cycle of length four in H which is a contradiction.

Corollary 2.2 If G is a bipartite graph and H is a triangulation of G then for every complete bipartite subgraph in G at least one of the color classes induces a complete subgraph in H.

We say that the color class is completed in H. Suppose we make a list of all maximal complete bipartite subgraphs of a bipartite graph G and for each we choose one color class to be completed. In [10] it is shown that in general the result need not be a triangulation, hence we restrict the choices of the color classes to be completed.

**Definition 2.6** Let  $M_1 = (A_1, B_1)$  and  $M_2 = (A_2, B_2)$  be two maximal complete bipartite subgraphs of G. We say that  $M_1$  and  $M_2$  cross if either  $A_2 \subset A_1$  and  $B_1 \subset B_2$ , or  $A_1 \subset A_2$  and  $B_2 \subset B_1$ .

**Definition 2.7** Let G be a bipartite graph and let  $\mathcal{M}$  be the set of all maximal complete bipartite subgraphs of G. For each  $M \in \mathcal{M}$  let C(M) be one of the color classes. The set  $\mathcal{C} = \{C(M) \mid M \in \mathcal{M}\}$  is called a feasible set if for each pair  $M_1 = (A_1, B_1), M_2 = (A_2, B_2) \in \mathcal{M}$  that cross with  $A_1 \supset A_2$  and  $B_1 \subset B_2$  not both  $A_1$  and  $B_2$  are in  $\mathcal{C}$ .

Notice that a feasible set always exists: Let  $\mathcal{C}$  be the set of color classes containing for each maximal complete bipartite subgraph the color class with the least number of vertices. This is a feasible set. In [10] the following theorem was proved.

**Theorem 2.1** Let G be chordal bipartite and C be a feasible set of G. Let H be the graph obtained from G by making each  $C \in C$  complete. Then H is a triangulation of G.

Let G be chordal bipartite. If  $\mathcal{C}$  is a feasible set, we denote by  $H_{\mathcal{C}}$  the triangulation of G obtained from G by making every color class in  $\mathcal{C}$  complete. We show that there exists a feasible set such that the number of edges in  $H_{\mathcal{C}}$  is minimum over all triangulations of G. In [10] the following result was shown.

**Theorem 2.2** Let G be chordal bipartite and let H be a triangulation of G. There is a feasible set C such that  $H_C$  is a subgraph of H.

Corollary 2.3 Let G be chordal bipartite. There is a feasible set  $\mathcal C$  such that, over all triangulations of G,  $H_{\mathcal C}$  has the minimum number of edges.

### 3 Couples and umbrellas

In this section let G = (X, Y, E) be chordal bipartite and let  $\mathcal{M}$  be the set of all maximal complete bipartite subgraphs.

Construct a digraph W with vertex set  $\mathcal{M}$  as follows. If  $M_1 = (A_1, B_1)$  and  $M_2 = (A_2, B_2)$  are elements of  $\mathcal{M}$  that cross with  $A_1 \subset A_2$  then we direct an arc from  $M_1$  to  $M_2$ .

#### Lemma 3.1 W is transitively oriented.

*Proof.* Clearly, the orientation is acyclic. Consider three elements  $M_i = (A_i, B_i)$ , i = 1, 2, 3 and assume there is an arc from  $M_1$  to  $M_2$  and from  $M_2$  to  $M_3$ . Then  $A_1 \subset A_2 \subset A_3$  and hence  $B_1 \supset B_2 \supset B_3$ . This clearly implies that there is also an arc from  $M_1$  to  $M_3$ .

A source in a directed graph is a vertex without incoming arcs. A sink is a vertex without outgoing arcs. Notice that a vertex can be a source and a sink at the same time only if it is isolated.

**Definition 3.1** An umbrella is either a source or a sink of W.

Notice that the subgraph of W induced by the umbrellas is bipartite (with possibly isolated vertices).

We define a feasible set restricted to the umbrellas.

**Definition 3.2** Let  $\mathcal{U}$  be the set of umbrellas. For each  $U \in \mathcal{U}$  choose one color class C(U). The set  $\mathcal{C}^0 = \{C(U) \mid U \in \mathcal{U}\}$  is called a spine if for each pair  $U_1 = (A_1, B_1)$  and  $U_2 = (A_2, B_2)$  that cross with  $A_2 \subset A_1$  not both  $A_1$  and  $B_2$  are in  $\mathcal{C}^0$ .

One of the main observations is that a spine can be extended to a feasible set.

**Theorem 3.1** Let  $C^0$  be a spine. There is a feasible set C with  $C^0 \subseteq C$ .

*Proof.* We can construct  $\mathcal{C}$  in a greedy manner as follows. Start with  $\mathcal{C} = \mathcal{C}^0$ . Consider the maximal complete bipartite subgraphs of  $\mathcal{M} \setminus \mathcal{C}^0$  one by one. Assume that at a certain moment there are *no* maximal complete bipartite subgraphs  $M_1 = (A_1, B_1)$  and  $M_2 = (A_2, B_2)$  that cross with  $A_2 \subset A_1$  and both  $A_1$  and  $B_2$  in  $\mathcal{C}$ . Notice that by definition, initially, when  $\mathcal{C} = \mathcal{C}^0$  this is the case. Now consider a new maximal complete bipartite subgraph  $(A, B) \in \mathcal{M} \setminus \mathcal{C}$ .

If there is a maximal complete bipartite subgraph (C, D) that crosses with (A, B) such that  $C \subset A$  and  $D \in C$ , then put  $B \in C$ . Notice that in this case there is no maximal complete bipartite subgraph (K, L) that crosses with (A, B) with  $A \subset K$  and  $K \in C$ , otherwise (K, L) and (C, D) would also cross with  $C \subset K$  and  $K, D \in C$ .

If there is no maximal complete bipartite subgraph (C, D) that crosses with (A, B) with  $C \subset A$  and  $D \in C$  then put  $A \in C$ .

This proves that we can extend the spine  $C^0$  to a feasible set C.

**Definition 3.3** A couple is a pair of vertices p and q which are both in the same color class of G and which have at least two common neighbors.

Notice that each couple is contained in at least one maximal complete bipartite subgraph.

We assign each couple to a color class of an umbrella as follows. Let  $p, q \in X$ . Consider the maximal complete bipartite subgraph M = (C, D) with one color class  $D = N(p) \cap N(q)$  and the other color class  $C = \bigcap_{d \in D} N(d)$ . Choose any umbrella U = (K, L) such that in the digraph W there is an arc from U to M, or, if M is an umbrella itself, choose U = M. The couple p, q is under the color class K of the umbrella U.

Now consider a couple r, s in Y and let M = (C, D) be the maximal complete bipartite subgraph with  $C = N(r) \cap N(s)$  and  $D = \bigcap_{c \in C} N(c)$ . In this case choose an umbrella U = (K, L) such that in W there is an arc from M to U or, in case M is an umbrella itself, choose U = M. In this case we assign the couple to the color class L and we say that the couple is under the color class L of the umbrella U. In this way every couple is under exactly one color class of an umbrella.

Lemma 3.2 If the umbrella is not an isolated vertex in W then at most one of the color classes has couples under it. If the umbrella is an isolated vertex in W then the only couples under each color class are the pairs of vertices in this color class.

*Proof.* Obvious.

**Definition 3.4** Let U = (K, L) be an umbrella. The weight of each color class is the number of couples that are under it.

We define the weight of a spine.

**Definition 3.5** Let  $C^0$  be a spine. The weight of  $C^0$ ,  $w(C^0)$ , is the sum of the weights of the color classes in  $C^0$ .

We can now state our main theorem.

**Theorem 3.2** Let G = (X, Y, E) be a chordal bipartite graph and let  $C^0$  be a spine. Let C be any feasible set with  $C^0 \subseteq C$ . Then the number of edges in  $H_C$  is  $|E| + w(C^0)$ .

*Proof.* Consider an edge (p,q) of  $H_c$  that is not in E. Hence p and q are in the same color class of G. Without loss of generality assume that p and q are both in X. There must be a maximal complete bipartite subgraph M = (A, B) such that  $p, q \in A$  and  $A \in C$ . This shows that p, q is a couple.

Consider M' = (C, D) with  $D = N(p) \cap N(q)$  and  $C = \bigcap_{d \in D} N(d)$ . Notice that M and M' cross or are the same with  $C \subseteq A$ . Since C is a feasible set with  $A \in C$  we also have  $C \in C$ .

Since p, q is a couple in X there is an umbrella  $U = (U_1, U_2)$  such that p, q is under  $U_1$  of U. In the digraph W there is an arc from the umbrella U to M' or U = M'. It follows that  $U_1 \subseteq C$  and hence  $U_1 \in C$ . Hence the edge (p, q) is counted in  $w(C^0)$ .

Since each couple is under exactly *one* color class of an umbrella each edge of  $H_{\mathcal{C}}$  is counted exactly once.

### 4 The algorithm

In this section we give an algorithm to compute the minimum fill-in for chordal bipartite graphs. Let G = (X, Y, E) be chordal bipartite. In fact we give an algorithm that computes a spine which has minimum weight. It is then easy, using Theorem 3.1, to extend this spine to a feasible set. Completing each color class then gives the triangulation of G with a minimum number of edges by Theorem 3.2.

Consider the subgraph of W induced by the umbrellas. First consider the isolated vertices of the digraph. Notice that for each isolated umbrella we can take the smallest color class (the weight of the color class is in this case simply the number of unordered pairs of vertices) and put this in the spine. Since the umbrella is isolated, it cannot cross with any other umbrella. Hence there is spine with minimum weight containing this set of color classes as a subset.

Let  $T = (U_1, U_2)$  be the subgraph of W induced by the umbrellas without the isolated umbrellas. Let  $U_1$  be the set of umbrellas which are sources and let  $U_2$  be the set of umbrellas which are sinks. Of the vertices of  $U_1$  only the color class that is contained in X has a positive weight and of the vertices that are in  $U_2$ , only the color class that is contained in Y has a positive weight. Hence we can safely define the weight of an umbrella as the weight of that color class that has non zero weight.

We can formulate the problem as a coloring problem on a weighted bipartite graph as follows. Try to find a coloring of the vertices with two colors red and black. If a vertex is colored red, this means that the color class that is in Y is in the spine. If a vertex is colored black, the color class that is contained in X is in the spine. It follows from the non crossing condition that we must have the following restriction on the coloring.

**Definition 4.1** A vertex coloring of T is called correct if the following condition is satisfied. If an umbrella of  $U_1$  is colored red, then all its neighbors in  $U_2$  are also colored red. Or, equivalently, if a vertex of  $U_2$  is colored black then all its neighbors in X are also black.

**Definition 4.2** The weight of a correct coloring of  $T = (U_1, U_2)$  is defined as the sum of the weights of the black vertices of  $U_1$  and the weights of the red vertices of  $U_2$ .

**Lemma 4.1** Consider a correct coloring of the vertices of T. Let  $C^0$  be the set of color classes defined as follows. If a vertex is colored red then the color class of the umbrella contained in Y is in  $C^0$  and if a vertex is colored black then the color class contained in X is in  $C^0$ . Then  $C^0$  is a spine and the weight of the spine is the same as the weight of the coloring. Conversely, if  $C^0$  is a spine, then color an umbrella red if the color class that is contained in Y is in the spine and otherwise color the vertex black. This is a correct vertex coloring and the weight is equal to the weight of the spine.

*Proof.* Obvious.  $\Box$ 

Consider an independent set S in the bipartite graph. We can color the vertices of  $S \cap U_1$  red and the vertices of  $S \cap U_2$  black. If the rest of the vertices in  $U_1$  are colored black, and the rest of the vertices of  $U_2$  are colored red, then the weight of this correct coloring is  $w(U_1 \cup U_2) - w(S)$ . (The weight of a subset of vertices is defined as the sum of the weights of the vertices in the subset). On the other hand, consider a correct coloring. If S is the set of red vertices of  $U_1$  and black vertices of  $U_2$  then this set S must be an independent set. This proves the following.

**Lemma 4.2** Consider a correct coloring. The set of red vertices of  $U_1$  together with the set of black vertices of  $U_2$  is an independent set. On the other hand, let S be an independent set. Color the vertices of  $U_1$  that are in S red and the other vertices of  $U_1$  black. Color the vertices of  $U_2$  that are in S black and the other vertices of  $U_2$  red. This is a correct coloring with weight  $w(U_1 \cup U_2) - w(S)$ .

Hence the coloring problem is equivalent with the problem of finding a maximum weight independent set in the bipartite graph.

Change the weighted bipartite graph into a bipartite graph without weights as follows. Consider the vertices one by one. If the weight of a vertex is  $\boldsymbol{w}$  we replace the vertex by  $\boldsymbol{w}$  copies of the vertex, and each copy is made adjacent to all neighbors of the original vertex. Notice that if a vertex of the new graph is in a maximum independent set, then all copies are in the independent set. It follows that the problem of finding a maximum weight independent set in the old graph is equivalent with the maximum independent set of the new graph.

Since the sum of all weights is at most the number of pairs of vertices of G, the number of vertices in the unweighted bipartite graph is at most  $n^2$ , where n is the number of vertices in G.

Finally, notice that we can find a maximum independent set in a bipartite graph as follows. The number of vertices of a maximum independent set is equal to the total number of vertices minus the point covering number (i.e., the smallest number of vertices that cover all edges [7]). Since the graph is bipartite, the point covering number is equal to the number of edges in a maximum matching (see, e.g., [7]). A maximum matching in a bipartite graph can be found using a maximum flow algorithm [8]. We obtain the following result.

**Theorem 4.1** Let G = (X, Y, E) be a chordal bipartite graph. There exists a polynomial time algorithm to find the minimum fill-in of G. The algorithm can be implemented to run in  $O(n^5)$  time.

*Proof.* The constructed bipartite graph has at most  $n^2$  vertices. Hence a maximum matching can be found in  $O(n^5)$  time (see, e.g., [8]).

#### 5 Conclusions

In this paper we showed that computing the minimum fill-in of a chordal bipartite graph can be done in polynomial time. In this section we mention some open problems.

It would be of interest to know if there are other large classes of graphs for which the minimum fill-in can be computed in polynomial time.

Chordal bipartite graphs are properly contained in perfect elimination bipartite graphs. It would be interesting to know if MINIMUM FILL-IN is also solvable in polynomial time for this somewhat larger class of graphs.

It is very likely that there are faster algorithms for subclasses of chordal bipartite graphs, like convex bipartite graphs.

Another interesting question is of course whether the time bound of this algorithm can be improved. It could also be very useful to have a fast algorithm that approximates minimum fill-in within a constant factor for chordal bipartite graphs.

## 6 Acknowledgements

I like to thank H. Bodlaender, A. Brandstädt, A. Jacobs, D. Kratsch and H. Müller for valuable discussions.

### References

- [1] S. Arnborg, D. G. Corneil and A. Proskurowski, Complexity of finding embeddings in a k-tree, SIAM J. Alg. Disc. Meth. 8, (1987), pp. 277–284.
- [2] A. Brandstädt, Special Graph Classes—A Survey, Schriftenreihe des Fachbereichs Mathematik, SM-DU-199 (1991), Universität Duisburg Gesamthochschule.
- [3] P. Damaschke and H. Müller, Hamiltonian circuits in convex and chordal bipartite graphs, To appear.
- [4] L. Goh and D. Rotem, Recognition of perfect elimination bipartite graphs, Information Processing Letters 15, (4), (1982), pp. 179-182.

- [5] M. C. Golumbic, Algorithmic Graph Theory and Perfect Graphs, Academic Press, New York, 1980.
- [6] M. Grötschel, L. Lovász and A. Schrijver, Polynomial algorithms for perfect graphs, *Annals of Discrete Mathematics* **21**, (1984), pp. 325–356.
- [7] F. Harary, *Graph Theory*, Addison-Wesley Publ. Comp., Reading, Massuchusetts, (1969).
- [8] J. E. Hopcroft and R. M. Karp, A n<sup>5/2</sup> algorithm for maximum matching in bipartite graphs, SIAM J. Comput. 2, (1973), pp. 225–231.
- [9] T. Kloks, *Treewidth*, Ph. D. Thesis, Utrecht University, The Netherlands. To appear.
- [10] T. Kloks and D. Kratsch, Treewidth of chordal bipartite graphs, 10th Annual Symposium on Theoretical Aspects of Computer Science, Springer-Verlag, Lecture Notes in Computer Science 665, (1993), pp. 80-89.
- [11] A. Lubiw, Doubly lexical orderings of matrices, SIAM J. Comput. 16, (5), (1987), pp. 854-879.
- [12] H. Müller and A. Brandstädt, The NP-completeness of Steiner tree and Dominating set for chordal bipartite graphs, Theoretical Computer Science 53, (1987), pp. 257-265.
- [13] J. Spinrad, A. Brandstädt and L. Stewart, Bipartite permutation graphs, Discrete Applied Mathematics 18, (1987), pp. 279–292.
- [14] M. Yannakakis, Computing the minimum fill-in is NP-complete, SIAM J. Alg. Disc. Meth. 2, (1981), pp. 77-79.