Determining possible sets of leaves for spanning trees of dually chordal graphs

Pablo De Caria and Marisa Gutierrez

CONICET, Departamento de Matemática, Universidad Nacional de La Plata, C.C. 172, (1900), La Plata, Argentina*

Abstract

It will be proved that the problem of determining whether a set of vertices of a dually chordal graphs is the set of leaves of a tree compatible with it can be solved in polynomial time by establishing a connection with finding clique trees of chordal graphs with minimum number of leaves.

Keywords: Leaf, leafage, compatible tree, clique tree, chordal, dually chordal.

1 Introduction

Chordal and dually chordal graphs were found to have many applications, especially in biology. Both classes are endowed with characteristic tree structures, clique trees in chordal graphs and compatible trees in dually chordal graphs, which in several cases are connected with the solution of problems associated to the applications. A good example of this are *phylogenetic trees* [4, 5], used to model the evolutionary history of species, proteins, etc. In them, it is necessary that leaves represent the present individuals (or objects) and inner vertices should indicate possible ancestors. This makes desirable, also in a more general context, the ability to determine what vertices can be the leaves of a compatible tree or a clique tree.

The *leafage* of a chordal graph is the minimum number of leaves of a clique tree of the graph. A polynomial algorithm, running in time $O(n^3)$, to find the leafage of a chordal graph has been proposed recently [3]. The goal of this paper is to show that this enables an answer to the following problem: given a dually

*Complete address: Calle 50 y 115, C.C. 172, C.P. 1900, La Plata, Buenos Aires, Argentina. Tel/FAX: (540221)4245875-4229850. E-mails: pdecaria@mate.unlp.edu.ar/marisa@mate.unlp.edu.ar

chordal graph G and $A \subset V(G)$, determine if there is a tree compatible with G whose set of leaves is A. For that purpose, every dually chordal graph is found to be the clique graph of a chordal graph in such a way that there is a correspondence between the compatible trees of the former and the clique trees of the latter. This is stronger than the well known fact that dually chordal graphs are the clique graphs of chordal graphs. Then the problem is transformed into that of finding the leafage of a chordal graph.

2 Some graph terminology

This paper deals just with graphs without loops or multiple edges. For a graph G, $V(G)$ is the set of its vertices and $E(G)$ that of its edges. A set $A \subseteq V(G)$ is *complete* if all its elements are pairwise adjacent vertices of G. A maximal complete set is a *clique* and $\mathcal{C}(G)$ will be used to denote the family of cliques of G. A *clique edge cover* of G is defined as any subset F of $\mathcal{C}(G)$ such that any edge of G is contained in at least one element of F.

Given two vertices v and w of a graph G, the *distance* between v and w, or $d(v, w)$, is the length of a shortest path connecting v and w in G. For a vertex $v \in V(G)$, the *closed neighborhood* of v, $N[v]$, is the set composed of v and all the vertices adjacent to it. We say that w *dominates* v when $N[v] \subseteq N[w]$. The *disk* centered at vertex v with radius k is the set $N^k[v] := \{w \in V(G),\ d(v, w) \leq k\}$.

Let T be a tree. For all $v, w \in V(T)$, $T[v, w]$ will denote the path in T from v to w. And $\mathcal{L}(T)$ will denote the set of leaves of T.

Let $\mathcal{F}$ be a family of nonempty sets. The *intersection graph* of $\mathcal{F}$ has the members of $\mathcal{F}$ as vertices, two of them being adjacent if and only if their intersection is nonempty. The *clique graph* $K(G)$ of a graph G is the intersection graph of $\mathcal{C}(G)$.

A graph such that $\mathcal{C}(G)$ is a Helly family, i.e., any subfamily of pairwise intersecting cliques has a nonempty intersection, is called a *clique-Helly graph.*

3 Basic notions and properties

A *chord* of a cycle is an edge joining two nonconsecutive vertices of the cycle. *Chordal* graphs are those without chordless cycles of length at least four. A *clique tree* T of G is a spanning tree of $K(G)$ such that, for any $v \in V(G)$, the set $\{C \in \mathcal{C}(G),\ v \in C\}$ induces a subtree of T. One of the many characterizations for chordal graphs says that a graph is chordal if and only if it has a clique tree [6].

A vertex w is a *maximum neighbor* of v if $N^2[v] \subseteq N[w]$. A linear ordering $v_1 ... v_n$ of the vertices of G is a *maximum neighborhood ordering* of G if, for

$i = 1, ..., n$, v_i has a maximum neighbor in $G[\{v_i, ..., v_n\}]$. *Dually chordal graphs* can be defined as those possessing a maximum neighborhood ordering.

Moreover, more characterizations of dually chordal graphs have been given. In fact, given a connected graph G, it is dually chordal if and only if [1]:

1. *There is a spanning tree T of G such that any clique of G induces a subtree in T.*

2. *There is a spanning tree T of G such that any closed neighborhood of G induces a subtree in T.*

3. *G is clique-Helly and $K(G)$ is chordal.*

4. *G is the clique graph of a chordal graph.*

It is even true that any spanning tree fulfilling 1. also fulfills 2. and vice versa. Such a tree will be said to be *compatible* with G. We also have the following equivalence:

Theorem 1. *[2] Let T be a spanning tree of a graph G. Then T is compatible with G if and only if, for all $x, y, z \in V(G)$, $xy \in E(G)$ and $z \in T[x, y] \setminus \{x, y\}$ implies that $xz \in E(G)$ and $yz \in E(G)$.*

4 Leaves and dominated vertices

Before the goal of this paper is achieved, some properties about domination will be necessary to find some conditions that the leaves of a compatible tree should satisfy. The graphs considered are always connected.

Lemma 1. *Let G be a dually chordal graph and T a tree compatible with G. If v is a leaf of T and w is the vertex such that $vw \in E(T)$ then v is dominated by w.*

Proof. For any vertex u in $N[v] \setminus \{v, w\}$ it holds that $w \in T[u, v]$. From Theorem 1 we infer that w is adjacent to u and thus $N[v] \setminus \{v, w\} \subseteq N[w]$. As $\{v, w\}$ is also a subset of $N[w]$, the inclusion $N[v] \subseteq N[w]$ follows. □

Corollary 1. *Let G be a dually chordal graph, $|V(G)| \geq 3$, and T a tree compatible with G. Then each vertex in $\mathcal{L}(T)$ is dominated by at least one vertex of $V(T) \setminus \mathcal{L}(T)$.*

Proof. Trees with more than two vertices do not have adjacent leaves. Then, for any vertex $v \in \mathcal{L}(T)$, the only vertex adjacent to it in T is not in $\mathcal{L}(T)$ and dominates it according to Lemma 1. □

Lemma 2. *Let G be a dually chordal graph and T be a tree compatible with G. Then, given $v \in V(G)$, the set $D = \{w \in V(G) : N[v] \subseteq N[w]\}$, i.e., v itself and the vertices dominating it, induces a subtree of T.*

Proof. Let $w \in V(G)$. Then $w \in D$ if and only if, for all $u \in N[v]$, $u \in N[w]$, that is, $w \in N[u]$ for all $u \in N[v]$. Thus $w \in D$ if and only if $w \in \bigcap_{u \in N[v]} N[u]$ and so $D = \bigcap_{u \in N[v]} N[u]$.

Since T is compatible with G, any closed neighborhood induces a subtree in T. And if some subsets induce subtrees so does their intersection. Therefore D induces a subtree. □

As it was said before, clique trees of chordal graphs will be essential for the solution of the problem. In the next theorem we are going to see that not only every dually chordal graph G is the clique graph of a chordal graph, as the characterization of dually chordal graphs indicates, but also that some of the chordal graphs whose clique graphs equal G have the especial property that their clique trees are exactly the trees compatible with G. Thus, any problem about the compatible trees of G can be viewed as a problem about clique trees of any of those chordal graphs. This allows to take advantage of the fact that many problems about clique trees of chordal graphs have been comprehensively studied. Among them, we can find those regarding leaves.

Theorem 2. *Let G be a dually chordal graph and F be a clique edge cover of G. Let H be the intersection graph of $F \cup \{\{v\} : v \in V(G)\}$. Then*

(1) *H is chordal.*

(2) *$K(H) \simeq G$.*

(3) *Any clique tree of H is isomorphic to a tree compatible with G and vice versa.*

Proof. Let T be a tree compatible with G. Then any member of $F \cup \{\{v\} : v \in V(G)\}$ induces a subtree in T. As intersection graphs of subtrees of a tree are chordal [7], (1) follows.

Given any vertex $v \in V(G)$, the set $D_v = \{\{v\}\} \cup \{C \in F : v \in C\}$ is a clique of H because D_v is complete and the equality $N_H[\{v\}] = D_v$ implies maximality (and also that $\{v\}$ is simplicial in H). In fact, every clique of H is equal to D_v, for some $v \in V(G)$. A proof of this is given below.

Let $D \in C(H)$. Then the elements of $F \cap D$ are pairwise intersecting. As dually chordal graphs are clique-Helly, there is a vertex w which is an element of each clique of G in $F \cap D$ and therefore $F \cap D \subsetneq N_H[\{w\}] = D_w$. This implies that $F \cap D$ is not a maximal complete set of H and thus there exists $v \in V(G)$ such that $\{v\} \in D$. Hence, v is an element of each clique of G in D, so $D \subseteq D_v$. Since D and D_v are in $C(H)$, it follows that $D = D_v$. Therefore, $C(H) = \{D_v : v \in V(G)\}$.

Now we need to demonstrate that $D_u D_v \in E(K(H))$ if and only if $uv \in E(G)$, which implies that $K(H) \simeq G$. And the reasoning is as follows:

$$D_uD_v \in E(K(H)) \Leftrightarrow D_u \cap D_v \neq \emptyset \Leftrightarrow \exists C \in F,\ C \in D_u \wedge C \in D_v \Leftrightarrow \exists C \in F,\ u \in C \wedge v \in C \Leftrightarrow uv \in E(G)$$

where for the last equivalence we use the fact that F covers all the edges of G. This proves (2).

Let T be a clique tree for H and T' the spanning tree of G such that $uv \in E(T')$ if and only if $D_uD_v \in E(T)$. Let x, y be vertices adjacent in G and $z \in T'[x,y]\setminus\{x,y\}$. Then D_x and D_y are adjacent in $K(H)$ and let $C \in D_x \cap D_y$. As T is a clique tree, the subset $\{D \in \mathcal{C}(H) : C \in D\}$ induces a subtree of T, implying that D_z also belongs to it because $D_z \in T[D_x, D_y]$. Consequently $D_x \cap D_z \neq \emptyset$ and $D_y \cap D_z \neq \emptyset$ and hence $xz, yz \in E(G)$, making T' compatible with G.

Conversely, let T be a tree compatible with G and T' the spanning tree of $K(H)$ such that $D_uD_v \in E(T')$ if and only if $uv \in E(T)$. For any $v \in V(G)$, the set $\{D \in \mathcal{C}(H) : \{v\} \in D\} = \{D_v\}$ so it obviously induces a subtree. Let $C \in F$, D_x and D_y such that $C \in D_x \cap D_y$ and D_z be any vertex of $T'[D_x, D_y] \setminus \{D_x, D_y\}$. Then $x \in C$, $y \in C$ and $z \in T[x,y] \setminus \{x,y\}$. Since T is compatible with G and C induces a subtree in T, $z \in C$, that is, $C \in D_z$. This implies that the set $\{D \in \mathcal{C}(H) : C \in D\}$ induces a subtree in T' and therefore T' is a clique tree of H. □

Having narrowed down before what the elements of $\mathcal{L}(T)$ can be for a tree T compatible with a dually chordal graph G, it remains to introduce an auxiliary graph G' which will contain information about the problem. The results required now are the following:

Lemma 3. *Let G be a dually chordal graph, T a tree compatible with G and u, v, w vertices such that $uv \in E(T)$, $v \in T[u,w]$ and $N[u] \cap N[v] \subseteq N[w]$. Then $T' = T - uv + uw$ is also compatible with G.*

Proof. Let x be any vertex of G. We need to prove that $N[x]$ induces a subtree in T'. Call $T[A]$ and $T[B]$ the connected components of $T - uv$, with $u \in A$ and $v \in B$. The proof is divided into three cases.

If $N[x] \subseteq A$ then $N[x]$ induces the same subtree in T and T'. If $N[x] \subseteq B$ the reasoning is similar. Otherwise we have two vertices $y, z \in N[x]$ such that $y \in A$ and $z \in B$. As $N[x]$ induces a subtree in T and $u, v \in T[y,z]$ we conclude that $u, v \in N[x]$, that is, $x \in N[u] \cap N[v]$ and therefore $x \in N[w]$. Now, u and v are connected in T' by the path formed by merging uw and $T[w,v]$ (contained in $N[x]$ because $w, v \in N[x]$ and T is compatible with G); and any other two vertices of $N[x]$ adjacent in T are still adjacent in T'. Therefore, vertices of $N[x]$ adjacent in T are connected in T' by paths within $N[x]$ and this is enough to claim that $N[x]$ induces a subtree in T', making T' compatible with G. □

Theorem 3. *Let G be a dually chordal graph and $A \subseteq V(G)$ be a set of vertices, each being dominated by a vertex in $V(G) \setminus A$. Let G' be a graph constructed from G by adding, for each $v \in A$, a vertex v^* and the edge vv^*. Then G' is dually chordal. Moreover, there is a tree T compatible with G such that $\mathcal{L}(T) = A$ if and only if there is a tree T' compatible with G' such that $\mathcal{L}(T') = A^* := \{v^*,\ v \in A\}$.*

Proof. Let T be a tree compatible with G. Then the tree T' such that $V(T') = V(G) \cup A^*$ and $E(T') = E(T) \cup \{vv^*,\ v \in A\}$ is compatible with G', so this graph is dually chordal. Furthermore, if $\mathcal{L}(T) = A$ then $\mathcal{L}(T') = A^*$.

Conversely, suppose that there exists a tree T' compatible with G' such that $\mathcal{L}(T') = A^*$ and set $T_0 = T' - A^*$. Choose T' so that $|\mathcal{L}(T_0)|$ is maximized. Since $\mathcal{C}(G) \subseteq \mathcal{C}(G')$, any clique of G induces a subtree of T' and thus in T_0 as well, so we conclude that T_0 is compatible with G. Now we show that $\mathcal{L}(T_0) = A$, which will prove the claim.

It is straightforward that $\mathcal{L}(T_0) \subseteq A$, otherwise any vertex in $\mathcal{L}(T_0) \setminus A$ would also be a leave of T', which is a contradiction.

If $\mathcal{L}(T_0) \neq A$, take a vertex $u \in A \setminus \mathcal{L}(T_0)$ and let w be a vertex in $V(G) \setminus A$ dominating u and w' the vertex adjacent to u in $T_0[u, w]$.

By Lemma 3, if for any vertex x different from w' and adjacent to u in T_0 we add the edge wx to T_0 and remove ux, we get a new tree T_1 compatible with G such that the degree of w is bigger in T_1 than in T_0, u is a leaf of T_1 and the remaining vertices have the same degree in T_0 and T_1. Then $\mathcal{L}(T_1) = \mathcal{L}(T_0) \cup \{u\}$, contradicting the way T_0 was chosen. Therefore, $\mathcal{L}(T_0) = A$. □

Now it is possible to prove the main theorem:

Theorem 4. *Let G be a dually chordal graph and A be a subset of $V(G)$ such that for each vertex of A there is a vertex in $V(G) \setminus A$ dominating it. Determining if there exists a tree compatible with G and whose set of leaves equals A can be reduced, in polynomial time, to the problem of finding the leafage of a chordal graph and hence it is itself polynomial.*

Proof. Let G' be the same graph as in Theorem 3 and H' be a chordal graph such that $K(H') \simeq G'$ and constructed as in Theorem 2. Denote by T' a clique tree for H' with minimum number of leaves and let T^* be a tree compatible with G' and isomorphic to T'. By part (3) of Theorem 2, T^* is a compatible tree for G' with minimum number of leaves.

If $\mathcal{L}(T^*) = A^*$, Theorem 3 implies that there is a tree T compatible with G and $\mathcal{L}(T) = A$. Otherwise, since the degree in G' of the vertices in A^* equals 1, $A^* \subsetneq \mathcal{L}(T^*)$. As the number of leaves of T^* is minimum, no tree compatible with G' has A^* as set of leaves and this time Theorem 3 implies that there is no tree T compatible with G and such that $\mathcal{L}(T) = A$.

We finish the proof by showing that H' can be constructed in polynomial time. For each edge of G' take $C \in \mathcal{C}(G')$ containing it. The resulting collection of cliques is a clique edge cover of G', call it F, whose cardinality is bounded by $|E(G')|$, and so by $|E(G)| + |V(G)|$. Thus, the intersection graph of $F \cup \{\{v\} : v \in V(G')\}$ can be obtained in polynomial time. Set H' equal to this graph. □

Consequently, given G dually chordal graph and $A \subset V(G)$ such that each vertex of A is dominated by a vertex in $V(G) \setminus A$, determining if there exists a tree T compatible with G and such that $\mathcal{L}(T) = A$ depends on whether the leafage of H' equals $|A|$ or not.

It is suitable to stress that, in case that the answer to the problem is affirmative, a tree compatible with G with set of leaves equal to A can be efficiently constructed. The algorithm presented in [3] produces a tree with $|A|$ leaves isomorphic to a tree compatible with G' and then it is easy to obtain a tree T_0 compatible with G and $\mathcal{L}(T_0) \subseteq A$. In case that $\mathcal{L}(T_0) \neq A$, we can apply the ensuing mechanism of increasing the number of leaves described in Theorem 3 until the set of leaves equals A.

If it is not true that any vertex of A is dominated by one of $V(G) \setminus A$, deciding whether the tree exists is trivial because, by Corollary 1, in this case only graphs with at most 2 vertices have to be considered.

References

[1] A. Brandstädt, F. Dragan, V. Chepoi and V. Voloshin, Dually chordal graphs, Proceedings of the 19th International Workshop on Graph Theoretic Concepts in Computer Science, Lecture Notes in Comp. Sci., 790 (1994), pp. 237-251.

[2] M. Gutierrez, Tree-Clique Graphs, Estudos e comunicações do Instituto de Matemática, Universidade Federal do Rio de Janeiro, 63 (1996), pp. 7-26.

[3] M. Habib, J. Stacho, Polynomial-time Algorithm for the Leafage of Chordal Graphs, Algorithms - ESA 2009, Lecture Notes in Comp. Sci., 5757 (2009), pp. 290-300.

[4] F.R. McMorris, T.J. Warnow, T. Wimer, Triangulating vertex-colored graphs, SIAM J. Discrete Math. 7 (2) (1994), pp. 296-306.

[5] J. Meidanis and J. Setubal, Introduction to Computational Molecular Biology, PWS Publishing, Boston MA, 1997.

[6] Y. Shibata, On the tree representation of chordal graphs, J. Graph Theory, 12 (1988), pp. 421-428.