



**HAL**  
open science

# Characterizing path graphs by forbidden induced subgraphs

Benjamin Lévêque, Frédéric Maffray, Myriam Preissmann

► **To cite this version:**

Benjamin Lévêque, Frédéric Maffray, Myriam Preissmann. Characterizing path graphs by forbidden induced subgraphs. 2008. hal-00261413v2

**HAL Id: hal-00261413**

**<https://hal.science/hal-00261413v2>**

Preprint submitted on 12 Mar 2008 (v2), last revised 17 Apr 2008 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Characterizing path graphs by forbidden induced subgraphs

Benjamin Lévêque, Frédéric Maffray, Myriam Preissmann

March 12, 2008

## Abstract

A graph is a path graph if it is the intersection graph of a family of subpaths of a tree. In 1970, Renz asked for a characterization of path graphs by forbidden induced subgraphs. Here we answer this question by listing all graphs that are not path graphs and are minimal with this property.

## 1 Introduction

A hole is a chordless cycle of length at least four. A graph is chordal (or triangulated) if it contains no holes as an induced subgraph. Gavril [3] proved that a graph is chordal if and only if it is the intersection graph of a family of subtrees of a tree.

An *interval graph* is the intersection graph of a family of intervals; equivalently, it is the intersection graph of a family of subpaths of a path. An *asteroidal triple* in a graph  $G$  is a set of three non adjacent vertices such that for any two of them, there exists a path between them in  $G$  that does not intersect the neighborhood of the third. Lekkerkerker and Boland [5] proved that a graph is an interval graph if and only if it is chordal and contains no asteroidal triple. They derive from this result the list of minimal forbidden subgraphs for interval graphs.

Another interesting class is the class of path graphs. A graph is a *path graph* if it is the intersection graph of a family of subpaths of a tree. Clearly, the class of path graphs is included in the class of chordal graphs and contains the class of interval graphs. Several characterizations of path graphs have been given [4, 6, 7] but no characterization by forbidden subgraphs is known, whereas such results exists for intersection graphs of subpaths of a paths (interval graphs [3]) and subtrees of a tree (chordal graphs [5]).

In 1970, Renz [7] gave two examples of graphs that are chordal and not path graphs and are minimal with this property. He also asked for a complete list of such graphs. In this paper we answer this question and obtain a characterizat on of path graphs by forbidden induced subgraphs.

## 2 Special simplicial vertices in chordal graphs

A vertex is *simplicial* if its neighborhood is a clique.

Let  $G$  be a chordal graph, and let  $\mathcal{Q}(G)$  be the set of all maximal cliques of  $G$ .

Given two vertices  $u, v$  in a graph  $G$ , a  $u, v$ -separator is a set  $S$  of vertices of  $G$  such that  $u$  and  $v$  lie in two different components of  $G \setminus S$  and  $S$  is minimal with this property. A set is a *separator* if it is a  $u, v$ -separator for some  $u, v$  in  $G$ . Let  $\mathcal{S}(G)$  be the set of separators of  $G$ . Given a simplicial vertex  $v$ , let  $Q_v = N(v) \cup \{v\}$  and  $S_v = Q_v \cap N(V \setminus Q_v)$ . Since  $v$  is simplicial,  $Q_v \in \mathcal{Q}$  (remark that  $S_v$  is not necessarily in  $\mathcal{S}$ ).

A *clique tree*  $T$  of  $G$  is a tree whose vertices are in  $\mathcal{Q}$  and such that for each vertex  $v$  of  $G$ , the set of cliques that contain  $v$  induces a subtree of  $T$ . A graph is chordal if and only if it has a clique tree representation.

For a clique tree  $T$ , the label of an edge  $QQ'$  of  $T$  is  $S_{QQ'} = Q \cap Q'$ . For all edges  $QQ'$ , we have  $S_{QQ'} \in \mathcal{S}$ . The number of times an element  $S$  of  $\mathcal{S}$  appears as a label of an edge is constant for every clique-tree.

Given  $X \subseteq \mathcal{Q}$ , let  $G(X)$  be the subgraph of  $G$  induced by all the vertices that appears in members of  $X$ . If  $T$  is a clique tree of  $G$ , then  $T[X]$  is the subtree of  $T$  of minimum size whose vertices contains  $X$ . Note that if  $|X| = 2$ , then  $T[X]$  is a path. For every vertex  $a$ , set  $T^a = T[\{Q \in \mathcal{Q} \mid a \in Q\}]$ .

Given a subtree  $T'$  of a clique-tree  $T$  of  $G$ . Let  $\mathcal{Q}(T')$  be the set of vertices of  $T'$  and  $\mathcal{S}(T')$  be the set of separators of  $G(\mathcal{Q}(T'))$ .

The following lemma is clear.

**Lemma 1** *A vertex is simplicial if and only if it does not belong to any separator.*

Dirac [2] proved that a chordal graph that is not a clique contains two non adjacent simplicial vertices. We need to generalize this theorem to the following. Let us say that a simplicial vertex  $v$  is *special* if  $S_v$  is (inclusionwise) maximal in  $\mathcal{S}$ .

**Theorem 1** *In a chordal graph that is not a clique, there exist two non adjacent special simplicial vertices.*

*Proof.* We prove the lemma by induction on  $|\mathcal{Q}|$ . By the hypothesis,  $G$  is not a clique, so  $|\mathcal{Q}| \geq 2$ .

*Case 1:  $\mathcal{S}$  has only one maximal element  $S$ .* Let  $Q, Q'$  be two maximal cliques such that  $Q \cap Q' = S$ . Let  $v \in Q \setminus Q'$  and  $v' \in Q' \setminus Q$ . The set  $S$  is the only maximal separator and it does not contain  $v$  or  $v'$ . So  $v$  and  $v'$  do not belong to any element of  $\mathcal{S}$ , and they are simplicial by Lemma 1 and  $S_v = S_{v'} = S$ .

*Case 2:  $\mathcal{S}$  has two distinct maximal elements  $S, S'$ .* Let  $T$  be a clique tree of  $G$ . Let  $Q_1, Q_2, Q'_1, Q'_2 \in \mathcal{Q}$  be such that  $S = S_{Q_1 Q_2}$ ,  $S' = S_{Q'_1 Q'_2}$ , and  $Q_2, Q_1, Q'_1, Q'_2$  appear in this order along the path  $T[\{Q_2, Q_1, Q'_1, Q'_2\}]$  (maybe  $Q_1 = Q'_1$ ).

Let  $Y$  be the subtree of  $T \setminus Q_1$  that contains  $Q_2$ , and let  $Z$  be the tree that consists of  $Y$  plus the vertex  $Q_1$  and the edge  $Q_1 Q_2$ . The subtree  $Z$  does not contain  $Q'_2$ , so  $G(\mathcal{Q}(Z))$  has strictly fewer maximal cliques than  $G$ . By the induction hypothesis, there exist two non adjacent simplicial vertices  $v, w$  of  $G(\mathcal{Q}(Z))$  such that  $S_v, S_w$  are maximal elements of  $\mathcal{S}(Z)$ . At most one of  $v, w$  is in  $Q_1$  since they are not adjacent. Suppose  $v$  is not in  $Q_1$ . We claim that  $v$  is a simplicial vertex of  $G$  and that  $S_v$  is a maximal element

of  $\mathcal{S}$ . Vertex  $v$  does not belong to any element of  $\mathcal{S}(Z)$ . If it belongs to an element of  $\mathcal{S} \setminus \mathcal{S}(Z)$ , then it also belongs to  $S \in \mathcal{S}(Z)$ , a contradiction. So  $v$  does not belong to any element of  $\mathcal{S}$  and it is a simplicial vertex of  $G$  by Lemma 1. The set  $S_v$  is a maximal element of  $\mathcal{S}(Z)$ . If it is not a maximal element of  $\mathcal{S}$ , then it is included in  $S \in \mathcal{S}(Z)$ , a contradiction. So  $v$  is a special simplicial vertex of  $G$ .

Likewise, let  $Y'$  be the subtree of  $T \setminus Q'_1$  that contains  $Q'_2$ , and let  $Z'$  be the tree that consists of  $Y'$  plus the vertex  $Q'_1$  and the edge  $Q'_1Q'_2$ . Just like with  $v$ , we can find a simplicial vertex  $v'$  of  $G(\mathcal{Q}(Z'))$  not in  $Q'_1$  that is a simplicial vertex of  $G$  with  $S_{v'}$  being a maximal element of  $\mathcal{S}$ .

Vertices  $v, v'$  are not adjacent as there are separated by  $Q_1$ .  $\square$

Algorithms LexBFS [8] and MCS [10] are linear time algorithms that were developed to find a simplicial vertex in a chordal graph. But the simplicial vertex that is found by these algorithms is not necessarily special. For example, on the graph with vertices  $a, b, c, d, e, f, g, h$  and edges  $ab, bc, cd, de, ef, fg, gh, bd, be, gd, ge$ , both algorithms LexBFS and MCS will always end on one of simplicial vertices  $a, h$ , which are not special.

The proof of Theorem 1 can be turned into a polynomial time algorithm to find a special simplicial vertex in a chordal graph. We do not know how to find such a vertex in linear time.

### 3 Forbidden induced subgraphs

Figures 1, 2, 3, 4 and 5 give a list of minimal forbidden subgraphs for path graphs. Let  $\mathcal{P}$  be the class of graphs that do not contain any  $F_i$  ( $0 \leq i \leq 16$ ) as an induced subgraph. In the next sections, will prove that graphs in  $\mathcal{P}$  are exactly path graphs.

Each graph in Figure 2 is obtained by adding a universal vertex to some minimal forbidden subgraph for interval graphs. Graphs  $F_{10}(n)_{n \geq 8}$  are also forbidden in interval graphs. Graphs  $F_6$  and  $F_{10}(8)$  are from Renz [7, Figures 1 and 5].

The following lemma is clear.

**Lemma 2** *In a path graph the neighborhood of every vertex is an interval graph.*

Lekkerkerker and Boland [5] proved that a graph is an interval graph if and only if it is chordal and contains no asteroidal triple. So we can deduce the following corollaries.

**Corollary 1** *In a path graph the neighborhood of every vertex contains no asteroidal triple.*

**Corollary 2**  $F_1, \dots, F_5$  are not path graphs.

One could have hoped (as we did initially) that a chordal graph in which the neighborhood of every vertex contains no asteroidal triple is a path graph; but this is not true, as observed already by Renz [7] and shown by graphs  $F_6, \dots, F_{16}$ . (Renz's graphs is our  $F_6$  and  $F_{10}(8)$ .)

It is a routine matter (and we omit the details) to check that each  $F_i$  ( $0 \leq i \leq 16$ ) is not a path graph and that for every vertex  $x$  of  $F_i$  the graph  $F_i \setminus x$  is a path graph. So we have:

**Theorem 2**  $F_0, \dots, F_{16}$  are minimally not path graphs.

## 4 Co-special simplicial vertices

A *clique path tree*  $T$  of  $G$  is a clique tree of  $G$  such that, for each vertex  $v$  of  $G$ , the subtree induced by the set of cliques that contain  $v$  is a path. Clearly, a graph is a path graph if and only if it has a clique path tree.

A simplicial vertex  $v$  is *co-special* if  $S_v$  is a minimal element of  $\mathcal{S}$  and  $G \setminus S_v$  has exactly two components (note that in that case  $S_v$  appears on exactly one label of any path tree of  $G$ ).

**Lemma 3** Let  $G$  be a minimally not path graph. Then either  $G$  is one of  $F_{11}, \dots, F_{15}$  or every simplicial vertex of  $G$  is co-special.

*Proof.* Suppose on the contrary that  $G$  is minimally not path graph, different from  $F_{11}, \dots, F_{15}$ , and there is a simplicial vertex  $v$  of  $G$  that is not co-special. All simplicial vertices of  $F_0, \dots, F_{10}$  are co-special, so  $G$  is not any of these graphs; moreover it does not contain any of them strictly (for otherwise  $G$  would not be minimally not path graph). So  $G$  is in  $\mathcal{P}$ .

Let  $T_0$  be a clique path tree of  $G \setminus v$ . Let  $Q'_v \in \mathcal{Q}(G \setminus v)$  be such that  $S_v \subseteq Q'_v$ . If  $Q'_v = S_v$ , then we can add  $v$  to  $Q'_v$  to obtain a clique path tree of  $G$ , a contradiction. So  $Q'_v \neq S_v$  and  $S_v \in \mathcal{S}$ .

Let  $T'$  be the maximal subtree of  $T_0$  that contains  $Q'_v$  and such that no label of the edges of  $T_0$  is included in  $S_v$ . Remark that  $T'$  plus vertex  $Q_v$  and edge  $Q'_v Q_v$  is a clique tree of  $G(\mathcal{Q}(T') \cup \{Q_v\})$  (but not necessarily a clique path tree), and in that tree only one label is included in  $S_v$ . Vertex  $v$  is not co-special, so in  $T_0$  there is an edge whose label is included in  $S_v$ , and so  $T'$  has strictly fewer vertices than  $T_0$ . So  $G(\mathcal{Q}(T') \cup \{Q_v\})$  is a path graph. Let  $T$  be a clique path tree of this graph.

We claim that  $Q_v$  is a leaf of  $T$ . If not, then there are at least two labels of  $T$  that are included in  $S_v$ , which contradicts the definition of  $T'$  (the number of times a label appears in a clique tree is constant).

Let  $T_1, \dots, T_\ell$  be the subtrees of  $T_0 \setminus T'$  ( $\ell \geq 1$ ). For  $1 \leq i \leq \ell$ , let  $Q_i Q'_i$  be the edge between  $T_i$  and  $T'$  with  $Q_i \in T_i$  and  $Q'_i \in T'$ . Note that  $Q_1, \dots, Q_\ell$  are pairwise disjoint; on the other hand,  $Q'_v, Q'_1, \dots, Q'_\ell$  are not necessarily pairwise disjoint). Let  $S_i = Q_i \cap Q'_i$  and  $v_i \in Q_i \setminus Q'_i$ . Let  $\mathcal{H} = (V_{\mathcal{H}}, E_{\mathcal{H}})$  be the intersection graph of  $S_1, \dots, S_\ell$ , that is,  $V_{\mathcal{H}} = \{S_1, \dots, S_\ell\}$  and  $E_{\mathcal{H}} = \{S_i S_j \mid S_i \cap S_j \neq \emptyset\}$ . For  $1 \leq i \leq \ell$ , let  $\mathcal{R}_i = \{S \in \mathcal{S}(T') \mid S_i \cap S \neq \emptyset \text{ and } S_i \setminus S \neq \emptyset\}$ . Let  $X = \{S_i \mid \mathcal{R}_i \neq \emptyset\}$ .

**Claim 1**  $\mathcal{H}$  contains no odd cycle.

*Proof.* Suppose on the contrary that there is an odd cycle  $S_1 \cdots S_p S_1$  in  $\mathcal{H}$ , with length  $p = 2r + 1$  ( $r \geq 1$ ). Let  $I_j = S_j \cap S_{j+1}$ , with  $S_{p+1} = S_1$ . Suppose that for some  $j \neq k$  we have  $I_j \cap I_k \neq \emptyset$ ; then there is a common vertex in the cliques  $Q_j, Q_{j+1}, Q_k, Q_{k+1}$ , and the number of different cliques among these is at least three, which contradicts the fact that  $T_0$  is a clique path tree as these three cliques do not lie on a common path of  $T_0$ . For  $1 \leq j \leq p$ , let  $s_j \in I_j$ . By the preceding remark, the  $s_j$ 's are pairwise distinct. By the definition of  $T'$ , we have  $S_j \subseteq S_v$  for each  $1 \leq j \leq p$ , so the  $s_j$ 's are all in  $Q_v$  and  $Q'_v$ . Let  $v' \in Q'_v \setminus Q_v$ . Let us consider the subgraph induced by  $v, v', v_1, \dots, v_p, s_1, \dots, s_p$ . Each of the non-adjacent vertices  $v$  and  $v'$  is adjacent to all of the clique formed by the  $s_j$ 's. Each vertex  $v_j$  is adjacent to  $s_{j-1}$  and  $s_j$  (with  $s_0 = s_p$ ) and not to any other  $s_i$  or to  $v$ . Vertex  $v'$  can have at most two neighbors among the  $v_j$ 's. If  $v'$  has no neighbor among them, then  $v, v', v_1, \dots, v_p, s_1, \dots, s_p$  induce  $F_{11}(4r + 4)_{r \geq 1}$ . If  $v'$  has exactly one neighbor among the  $v_j$ 's, then  $v, v', v_1, \dots, v_p, s_1, \dots, s_p$  induce  $F_{12}(4r + 4)_{r \geq 1}$ . If  $v$  has two consecutive neighbors  $v_j, v_{j+1}$ , then  $v, v', v_j, v_{j+1}, s_{j-1}, s_j, s_{j+1}$  induce  $F_2$ . If  $v$  has two non-consecutive neighbors  $v_j, v_k$ , then we can assume that  $1 \leq j < j + 1 < k \leq p$  and  $k - j$  is odd,  $k - j = 2s + 1$  with  $s \geq 1$ , and then  $v, v', v_j, \dots, v_k, s_j, \dots, s_{k-1}$  induce  $F_{14}(4s + 5)_{s \geq 1}$ , in all cases we obtain a contradiction.  $\square$

**Claim 2**  $\mathcal{H}$  contains no odd path between two vertices in  $X$ .

*Proof.* Suppose on the contrary that there is an odd path  $S_1 \cdots S_p$  (with  $p = 2k, k \geq 1$ ) in  $\mathcal{H}$  between two vertices  $S_1, S_p$  of  $X$ , and assume that  $p$  is minimum with this property. By the minimality, all interior vertices  $S_j$  ( $1 < j < p$ ) are not in  $X$ . For  $1 \leq j < p$ , let  $s_j$  be a vertex in  $S_j \cap S_{j+1}$ . As in the preceding claim, the  $s_j$ 's are pairwise distinct and lie in  $Q_v$  and  $Q'_v$ . Let  $P$  be the subpath  $T(\{Q'_1, Q'_2\})$ . If  $p \neq 2$ , then  $S_2$  is not in  $X$ , so  $Q'_3 = Q'_1$ ; then  $S_3$  is not in  $X$ , so  $Q'_4 = Q'_2$ , and so on. So, the two extremities of  $P$  are  $Q'_1 = Q'_3 = \dots = Q'_{p-1}$  and  $Q'_2 = Q'_4 = \dots = Q'_p$ . Since  $S_1, S_p$  are in  $X$ , the sets  $\mathcal{R}_1, \mathcal{R}_p$  are non empty. We distinguish between several cases.

*Case 1: there exists  $R_1 \in \mathcal{R}_1$  and  $R_p \in \mathcal{R}_p$  that are labels of edges not in  $P$ .* We can assume that  $R_1$  is the label of an edge  $K_0 L_0$  of  $T'$ , with  $L_0 \in P, K_0 \notin P$ . Let  $s_0 \in S_1 \cap R_1$  and  $v_0 \in K_0 \setminus L_0$ . Vertex  $s_0$  belongs to  $Q_v$  and  $Q'_v$  by the definition of  $T'$  and is distinct from  $s_1, \dots, s_{p-1}$  because  $T'$  is a clique path tree. Vertex  $v_0$  is not adjacent to any of  $s_1, \dots, s_{p-1}$  because  $T'$  is a clique path tree and  $v_0 \notin L_0$ . Similarly, we can assume that  $R_p$  is the label of an edge  $K_{p+1} L_{p+1}$  of  $T'$ , with  $L_{p+1} \in P$ . Let  $s_p \in S_p \cap R_p$  and  $v_{p+1} \in K_{p+1} \setminus Q'_{p+1}$ . Vertex  $s_p$  belongs to  $Q_v$  and  $Q'_v$  and is distinct from  $s_0, \dots, s_{p-1}$ . Vertex  $v_{p+1}$  is not adjacent to any of  $s_1, \dots, s_{p-1}$ . Since  $T'$  is a clique path tree,  $Q'_v$  lies between  $Q'_1$  and  $L_0$  and between  $L_{p+1}$  and  $Q'_p$  along  $P$ . So  $Q'_1, L_{p+1}, Q'_v, L_0, Q'_p$  lie in this order on  $P$  (these cliques are not necessarily all distinct). We distinguish between two cases.

*Case 1.1:  $K_0$  and  $K_{p+1}$  are equal.* Then  $L_0 = L_{p+1} = Q'_v$  and we can assume that  $v_0 = v_{p+1}$ . By the definition of  $T'$ , there exists  $y \in R_1 \setminus S_v$ . Vertex  $y$  is distinct from  $s_0, s_p$  as it is not in  $S_v$ . Vertex  $y$  is adjacent to  $v_0, s_0, \dots, s_p$  and to no vertex among  $v, v_1, \dots, v_p$ . So  $v, y, v_0, \dots, v_p, s_0, \dots, s_p$  induce  $F_{12}(4k + 4)_{k \geq 1}$ , a contradiction.

*Case 1.2:  $K_0$  and  $K_{p+1}$  are different.* Then  $v_0$  and  $v_{p+1}$  are distinct. We can choose vertices  $x_1, \dots, x_r$  ( $r \geq 1$ ) not in  $S_v$  and on the labels of  $T'[\{K_0, K_{p+1}\}]$  such that they form a chordless path  $v_0-x_1-\dots-x_r-v_{p+1}$ . These vertices are distinct from and adjacent to  $s_0, \dots, s_p$  and not to  $v, v_1, \dots, v_p$ . If  $r = 1$ , then  $v, v_0, \dots, v_{p+1}, s_0, \dots, s_p, x_1$  induce  $F_{14}(4k+5)_{k \geq 1}$ . If  $r = 2$ , then  $v, v_0, \dots, v_{p+1}, s_0, \dots, s_p, x_1, x_2$  induce  $F_{15}(4k+6)_{k \geq 1}$ . If  $r \geq 3$ , then  $v, v_0, v_{p+1}, s_0, s_p, x_1, \dots, x_r$  induce  $F_{10}(r+5)_{r \geq 3}$ . A contradiction.

*Case 2: there exists  $R_1 \in \mathcal{R}_1$  and  $R_p \in \mathcal{R}_p$  that are labels of edges in  $P$ .* We can assume that  $R_1$  is the label of an edge  $K_0L_0$  of  $T'$ , with  $L_0 \in P$ ,  $K_0 \in P$ ,  $S_1 \subseteq L_0$ ,  $S_1 \not\subseteq K_0$ . Let  $s_0 \in S_1 \setminus R_1$  and  $v_0 \in K_0 \setminus L_0$ . Vertex  $s_0$  belongs to  $Q_v$  and  $Q'_v$  by the definition of  $T'$  and is distinct from  $s_1$  because it is not in  $K_0$ . Similarly, we can assume that  $R_p$  is the label of an edge  $K_{p+1}L_{p+1}$  of  $T'$ , with  $L_{p+1} \in P$ ,  $K_{p+1} \in P$ ,  $S_p \subseteq L_{p+1}$ ,  $S_p \not\subseteq K_{p+1}$ . Let  $s_p \in S_p \setminus R_p$  and  $v_{p+1} \in K_{p+1} \setminus Q'_{p+1}$ . Vertex  $s_p$  belongs to  $Q_v$  and  $Q'_v$  and is distinct from  $s_0, s_1$  because it is not in  $K_{p+1}$ . Since  $T'$  is a clique path tree,  $Q'_v$  lies between  $Q'_1$  and  $L_0$  and between  $L_{p+1}$  and  $Q'_p$  along  $P$ . So  $Q'_1, K_{p+1}, L_{p+1}, Q'_v, L_0, K_0, Q'_p$  lie in this order on  $P$ . We can choose vertices  $x_1, \dots, x_r$  ( $r \geq 1$ ) not in  $S_v$  and on the labels of  $T'[\{K_0, K_{p+1}\}]$  such that they form a chordless path  $v_0-x_1-\dots-x_r-v_{p+1}$ . These vertices are distinct from and adjacent to  $s_0, \dots, s_p$ . If  $r = 1$ , then  $v, v_0, v_{p+1}, s_0, s_1, s_p, x_1$  induce  $F_2$ . If  $r = 2$ , then  $v, v_0, v_{p+1}, s_0, s_1, s_p, x_1, x_2$  induce  $F_3$ . If  $r \geq 3$ , then  $v, v_0, v_{p+1}, x_1, \dots, x_r, s_0, s_p$  induce  $F_{10}(r+5)_{r \geq 3}$ . A contradiction.

*Case 3: there is  $R_1 \in \mathcal{R}_1$  that is the label of an edge in  $P$  and  $R_p \in \mathcal{R}_p$  that is the label of an edge not in  $P$  (or the symmetric case).* We can assume that  $R_1$  is the label of an edge  $K_0L_0$  of  $T'$ , with  $L_0 \in P$ ,  $K_0 \in P$ ,  $S_1 \subseteq L_0$ ,  $S_1 \not\subseteq K_0$ . Since  $T'$  is a clique path tree,  $Q'_v$  lies between  $Q'_1$  and  $L_0$  along  $P$ . So  $Q'_1, Q'_v, L_0, K_0, Q'_p$  lie in this order on  $P$ . Let  $s_0 \in S_1 \setminus R_1$  and  $v_0 \in K_0 \setminus L_0$ . Vertex  $s_0$  belongs to  $Q_v$  and  $Q'_v$  by the definition of  $T'$  and is distinct from  $s_1, \dots, s_{p-1}$  because it is not in  $K_0$ . Vertex  $v_0$  is adjacent to  $s_1, \dots, s_{p-1}$ . Vertex  $v_0$  is not in  $L_0$ , so not in  $S_v$ , so not in any  $S_i$ . So  $v_0$  is not adjacent to  $v, v_1, \dots, v_p$ . We can assume that  $R_p$  is the label of an edge  $K_{p+1}L_{p+1}$  of  $T'$ , with  $L_{p+1} \in P$ . Let  $s_p \in S_p \cap R_p$  and  $v_{p+1} \in K_{p+1} \setminus Q'_{p+1}$ . Vertex  $s_p$  belongs to  $Q_v$  and  $Q'_v$  and is distinct from  $s_0, \dots, s_{p-1}$ . Vertex  $v_{p+1}$  is not adjacent to any of  $s_1, \dots, s_{p-1}$ . If vertex  $s_0$  sees  $v_{p+1}$ , then  $R_p \in \mathcal{R}_1$  and we can apply case 1. So we can assume that  $v_{p+1}$  is not adjacent to  $s_0$ . Since  $T'$  is a clique path tree,  $Q'_v$  lies between  $L_{p+1}$  and  $Q'_p$  along  $P$ . We can choose vertices  $x_1, \dots, x_r$  ( $r \geq 1$ ) not in  $S_v$  and on the labels of  $T'[\{K_0, K_{p+1}\}]$  such that they form a chordless path  $v_0-x_1-\dots-x_r-v_{p+1}$ . These vertices are distinct from and adjacent to  $s_0, \dots, s_p$  and not to  $v, v_1, \dots, v_p$ . If  $r = 1$ , then  $v, v_0, \dots, v_{p+1}, s_0, \dots, s_p, x_1$  induce  $F_{13}(4k+5)_{k \geq 1}$ . If  $r \geq 2$ , then  $v, v_0, v_{p+1}, x_1, \dots, x_r, s_0, s_p$  induce  $F_5(r+5)_{r \geq 2}$ . A contradiction. □

By the preceding claims,  $\mathcal{H}$  is a bipartite graph  $(A, B, E_{\mathcal{H}})$  such that  $X \subseteq A$ .

We claim that all the subtrees  $T_i$  can be linked to  $T$  to get a clique path tree of  $G$ . For each  $S_i \in A$ , we add an edge  $Q_v Q_i$  between  $T$  and  $T_i$ . This creates no illegal branching because  $A$  is a stable set of  $\mathcal{H}$  and  $Q_v$  is a leaf of  $T$ . For each  $S_i \in B$ , let  $Q''_i \in \mathcal{Q}(T)$  be such that  $Q''_i \cap S_i \neq \emptyset$  and the length of  $T[\{Q_v, Q''_i\}]$  is maximal. For

each  $S_i \in B$ , we have  $\mathcal{R}_i = \emptyset$  so  $S_i \subseteq Q_i''$  and we can add an edge  $Q_i''Q_i$  between  $T$  and  $T_i$ . This creates no illegal branching because  $B$  is a stable set of  $\mathcal{H}$  and by the definition of  $Q_i''$ . Thus we obtain a clique path tree of  $G$ , a contradiction.  $\square$

## 5 Characterization of path graphs

In this section we prove the main theorem, that is, the class of path graphs is exactly  $\mathcal{P}$ . We could not find a characterization similar to the one found by Lekkerkerker and Boland [5] for interval graphs (“an interval graph is a chordal graph with no asteroidal triple”). We know that in a path graph, the neighborhood of every vertex contains no asteroidal triple but the converse is not true. So we prove directly that a graph that does not contain any of the excluded subgraphs is a path graph.

**Lemma 4** *In a graph in  $\mathcal{P}$ , the neighborhood of every vertex does not contain an asteroidal triple.*

*Proof.* It suffices to check that when a universal vertex is added to a minimal forbidden induced subgraph for interval graphs ([5]), then one obtains a graph that contains one of  $F_0, \dots, F_5, F_{10}$ . The easy details are left to the reader.  $\square$

Given three non adjacent vertices  $a, b, c$ , we say that  $a$  is the *middle* of  $b, c$  if every path between  $b$  and  $c$  contains a vertex from  $N(a)$ . If  $a, b, c$  is not an asteroidal triple, then at least one of them is the middle of the others.

Let us say that a vertex  $x$  is *complete* to a set  $S$  of vertices if  $x$  is adjacent to every vertex in  $S$ .

**Lemma 5** *In a chordal graph  $G$  with clique tree  $T$ , a vertex  $a$  is the middle of  $b, c$  if and only if for all cliques  $Q_b$  and  $Q_c$  such that  $b \in Q_b$  and  $c \in Q_c$ , there is an edge of the path  $T[\{Q_b, Q_c\}]$  such that  $a$  is complete to its label.*

*Proof.* Suppose that  $a$  is the middle of  $b, c$ . Let  $Q_b$  and  $Q_c$  be cliques such that  $b \in Q_b$  and  $c \in Q_c$ , and suppose there is no edge of  $T[\{Q_b, Q_c\}]$  such that  $a$  is complete to its label. For each edge on  $T[\{Q_b, Q_c\}]$ , one can select a vertex that is not adjacent to  $a$ . Then the set of selected vertices forms a path from  $b$  to  $c$  that uses no vertex from  $N(a)$ , a contradiction.

Suppose now that for all cliques  $Q_b$  and  $Q_c$  such that  $b \in Q_b$  and  $c \in Q_c$ , there is an edge of the path  $T[\{Q_b, Q_c\}]$  such that  $a$  is complete to its label. Suppose that there exists a path  $x_0 \dots x_r$ , with  $b = x_0$  and  $c = x_r$  and none of the  $x_i$ 's is in  $N(a)$ . We can assume that this path is chordless. Then, for  $1 \leq i \leq r$ , let  $Q_i$  be a maximal clique containing  $x_{i-1}, x_i$ . Then  $Q_1, \dots, Q_r$  appear in this order along a subpath of  $T$ . On each  $T[\{Q_i, Q_{i+1}\}]$  ( $1 \leq i \leq r-1$ ), vertex  $a$  is not adjacent to  $x_i$ , so  $a$  is not complete to any label of  $T[\{Q_1, \dots, Q_r\}]$ , but  $Q_1$  contains  $b$  and  $Q_r$  contains  $c$ , a contradiction.  $\square$

Now we are ready to prove the main theorem. Part of the proof has been done in the previous section. Lemma 3 deals with the case where there exists a simplicial vertex that



is the middle of two other vertices; now we have to look at the case where all simplicial vertices are not the middle of any pair of vertices.

**Theorem 3** *A chordal graph is a path graph if and only if it does not contain any of  $F_0, \dots, F_{16}$  as an induced subgraph.*

*Proof.* By Theorem 2, a path graph is in  $\mathcal{P}$ . Suppose now that there exists a minimally not path graph  $G$  in  $\mathcal{P}$ . Graph  $G$  is chordal. By Theorem 1, there is a special simplicial vertex  $q$  of  $G$ . By Lemma 3,  $q$  is co-special. Let  $Q = Q_q$  and  $S_Q = S_q \in \mathcal{S}$ . Let  $T_0$  be a clique path tree of  $G(\mathcal{Q} \setminus Q)$ . Let  $Q' \in \mathcal{Q} \setminus Q$  be such that  $S_Q \subseteq Q'$ . We add the edge  $QQ'$  to  $T_0$  to obtain a clique tree  $T'_0$  of  $G$ .

For each clique  $L \in \mathcal{Q} \setminus \{Q, Q'\}$ , let  $L'$  be the neighbor of  $L$  along  $T_0[\{L, Q'\}]$ . Let  $S_L = L \cap L'$ . Let  $\mathcal{S}_L$  be the set of labels of edges incident to  $L$  in  $T_0$ . Let  $\bar{L}$  be the clique such that  $L, L', \bar{L}, \bar{L}'$  appear in this order along  $T_0[\{L, Q'\}]$ , no label of  $T_0[\{\bar{L}', Q'\}]$  is included in  $S_L$ , and  $S_{\bar{L}} \subseteq S_L$ .

Let  $\mathcal{L}$  be the set of cliques  $L$  of  $\mathcal{Q} \setminus \{Q, Q'\}$  such that no element of  $S_L \setminus S_{\bar{L}}$  contains  $S_{\bar{L}}$ .

For each clique  $L \in \mathcal{L}$ , we define a subtree  $T_L$  of  $T'_0$ , where  $T_L$  is the biggest subtree of  $T'_0$  that contains  $Q'$  and for which no label is included in  $S_L$ . Note that  $\bar{L}'$  is in  $T_L$  and  $\bar{L}$  is not in  $T_L$ . Subtree  $T_L$  contains  $Q$  as  $q$  is special and co-special, and so  $S_Q \not\subseteq S_L$ .

**Claim 1** *For each clique  $L \in \mathcal{L}$  we have  $L' \in T_L$ .*

*Proof.* Suppose on the contrary that  $L' \notin T_L$ . Then  $\bar{L} \neq L$ . When we remove the edges  $LL'$  and  $\bar{L}\bar{L}'$  from  $T'_0$ , there remain three connected subtrees. Let  $T_1$  be the subtree containing  $L$ ,  $T_2$  be the subtree containing  $L', \bar{L}$ , and  $T_3$  be the subtree containing  $\bar{L}', Q', Q$ . Let  $T_4$  be the tree formed by  $T_1, T_3$  plus the edge  $L\bar{L}'$ . Then  $T_4$  is a clique tree of  $G(\mathcal{Q}(T_4))$ . The set  $\mathcal{Q}(T_4)$  contains strictly fewer maximal cliques than  $\mathcal{Q}$ , so let  $T_5$  be a clique path tree of  $G(\mathcal{Q}(T_4))$ .

We claim that there is an edge of  $T_5$  that is incident to  $L$  and that has  $S_{\bar{L}}$  as a label. On the clique tree  $T_4$ , the label  $S_{\bar{L}}$  is on the edge  $L\bar{L}'$ , so it is also a label of  $T_5$ . So there is an edge with label  $S'_{\bar{L}}$ , incident to  $L$  such that  $S_{\bar{L}} \subseteq S'_{\bar{L}} \subseteq L$ . Suppose that  $\bar{S}_L \subsetneq S'_{\bar{L}}$ . Then there is an edge of  $T_1$  or  $T_3$  with label  $S'_{\bar{L}}$ . No label of  $T_1$  can be  $S'_{\bar{L}}$  by the definition of  $\mathcal{L}$ . All the labels of  $T_3$  that are included in  $L$  are also included in  $S_{\bar{L}}$ , so no label of  $T_3$  can be  $S'_{\bar{L}}$ . So  $S_{\bar{L}} = S'_{\bar{L}}$ .

Let  $L\bar{L}''$  be an edge of  $T_5$  incident to  $L$  with label  $S_{\bar{L}}$  (maybe  $\bar{L}'' = \bar{L}'$ ). We can remove this edge from  $T_5$  and replace it by the subtree  $T_2$  and edges  $LL', \bar{L}\bar{L}''$ . Thus we obtain a clique path tree of  $G$ , a contradiction.  $\square$

Let  $\mathcal{L}^*$  be the subset of  $\mathcal{L}$  such that  $T_L$  is a strict subtree of  $T'_0 \setminus L$ . Let  $A$  be the set of vertices  $a$  of  $Q$  such that  $Q'$  is a vertex of  $T_0^a$  that is not a leaf. Then  $A$  is not empty, for otherwise  $T'_0$  would be a clique path tree of  $G$ . For each  $a \in A$ , the leaves of  $T_0^a$  are in  $\mathcal{L}$  and we claim that at least one of them is in  $\mathcal{L}^*$ . Let  $a \in A$  and let  $L_1, L_2$  be the leaves of  $T_0^a$ . For  $i = 1, 2$ , let  $l_i \in L_i \setminus S_{L_i}$ . The three vertices  $q, l_1, l_2$  are adjacent to  $a$  so

they do not form an asteroidal triple by Lemma 4, so one of them is the middle of the others. Vertex  $q$  cannot be the middle of  $l_1, l_2$ , for otherwise by Lemma 5 there would be an edge of  $T_0[\{L_1, L_2\}]$  with a label included in  $S_Q$ , contradicting that  $q$  is co-special. So one of  $l_1, l_2$  is the middle of the others (maybe both). By symmetry we can assume that  $l_1$  is the middle of  $q, l_2$ . So there is an edge of  $T'_0[\{Q, L_2\}]$  with a label included in  $S_{L_1}$ . So  $T_{L_1}$  is a strict subtree of  $T'_0 \setminus L_1$  and so  $L_1 \in \mathcal{L}^*$ . So  $\mathcal{L}^*$  is not empty.

We choose  $L \in \mathcal{L}^*$  such that the subtree  $T_L$  is maximal. Let  $S_{Q'}$  be the label of the edge of  $T_0[\{L, Q'\}]$  that is incident to  $Q'$ . Vertex  $q$  is special and co-special, so there exists  $s_Q$  in  $S_Q \setminus S_{Q'}$ , and we have  $s_Q \notin S_L$ . We add the edge  $LL'$  to  $T_L$  to obtain a clique tree  $T'_L$  of  $G(\mathcal{Q}(T_L) \cup \{L\})$ . The subtree  $T'_L$  is a strict subtree of  $T'_0$ , so we can consider a clique path tree  $T$  of  $G(\mathcal{Q}(T'_L))$ . We claim that  $L$  is a leaf of  $T$ . If not, then there are at least two labels of  $T$  that are included in  $S_L$ , which contradicts the definition of  $T_L$ .

We define  $\mathcal{U}, \mathcal{V}, \mathcal{W}$  as follows :

$$\mathcal{U} = \{U \in \mathcal{Q} \setminus \mathcal{Q}(T'_L) \mid UL \text{ is an edge of } T_0\}$$

$$\mathcal{V} = \{V \in \mathcal{Q} \setminus \mathcal{Q}(T'_L) \mid \exists V' \in \mathcal{Q}(T_L), \text{ s.t. } VV' \text{ is an edge of } T_0 \text{ and } S_{VV'} \cap S_Q = \emptyset\}$$

$$\mathcal{W} = \{W \in \mathcal{Q} \setminus \mathcal{Q}(T'_L) \mid \exists W' \in \mathcal{Q}(T_L), \text{ s.t. } WW' \text{ is an edge of } T_0 \text{ and } S_{WW'} \cap S_Q \neq \emptyset\}$$

For each  $U \in \mathcal{U} \cup \mathcal{V} \cup \mathcal{W}$ , let  $T_U$  be the connected component of  $T'_0 \setminus T'_L$  that contains  $U$ .

We claim that the  $S_V$ 's, with  $V \in \mathcal{V} \cup \mathcal{W}$ , are pairwise disjoint. For if they are not disjoint, then there is a vertex in  $S_V \cap S_{V'}$  with  $V, V'$  in  $\mathcal{V} \cup \mathcal{W}$ . But then  $S_V, S_{V'}$  are included in  $S_L$ , so this vertex is on three labels of  $T_0$  that are not on a path, contradicting that  $T_0$  is a clique path tree.

Let  $\mathcal{U}_1 = \{U \in \mathcal{U} \mid \exists W \in \mathcal{W} \text{ such that } U \cap W \neq \emptyset\}$ .

**Claim 2** *There exists  $U \in \mathcal{U}_1$ , such that  $S_U \setminus Q' \neq \emptyset$ .*

*Proof.* We define  $\mathcal{U}_{p \geq 1}, \mathcal{V}_{p \geq 0}$  as follows. Let  $V_0 = W$  and for  $p \geq 1$ :

$$\mathcal{U}_p = \{U \in \mathcal{U} \setminus (\mathcal{U}_1 \cup \dots \cup \mathcal{U}_{p-1}) \mid \exists V \in \mathcal{V}_{p-1} \text{ such that } U \cap V \neq \emptyset\}$$

$$\mathcal{V}_p = \{V \in \mathcal{V} \setminus (\mathcal{V}_1 \cup \dots \cup \mathcal{V}_{p-1}) \mid \exists U \in \mathcal{U}_{p-1} \text{ such that } V \cap U \neq \emptyset\}$$

Let  $k$  be the smallest  $k \geq 1$  such that there exists  $U \in \mathcal{U}_k$  with  $S_U \setminus Q' \neq \emptyset$ . Let  $k = \infty$  if it does not exist.

Suppose by contradiction that  $k > 1$ . For all  $1 \leq p \leq k-1$  and all  $U \in \mathcal{U}_p$ , we have  $S_U \subseteq Q'$  and we define  $U'' \in \mathcal{Q}(T)$  such that  $U'' \cap S_U \neq \emptyset$  and the length of  $T[\{L, U''\}]$  is maximal.

Suppose that there exists  $U_p \in \mathcal{U}_p$ ,  $1 \leq p \leq k-1$ , such that  $S_{U_p} \not\subseteq U''$ , and let  $p$  be minimal with this property. Let  $V_0, \dots, V_{p-1}, U_1, \dots, U_p$  be such that  $V_i \in \mathcal{V}_i$ ,  $U_i \in \mathcal{U}_i$ ,  $V_{i-1} \cap U_i \neq \emptyset$ , and  $U_i \cap V_i \neq \emptyset$ . Let  $u_i \in U_i \setminus S_{U_i}$ , let  $v_i \in V_i \setminus S_{V_i}$ . Let  $x_1, \dots, x_r$  be such that  $x_1 \in V_0 \cap U_1$ ,  $x_2 \in U_1 \cap V_1$ ,  $\dots$ ,  $x_r \in V_{p-1} \cap U_p$  with  $r = 2p-1$ . By the definition of  $\mathcal{V}$ , none of  $x_2, \dots, x_r$  is in  $Q$ . Let  $x_0 \in V_0 \cap Q$  (maybe  $x_0 = x_1$ ). As  $x_0 \in S_{V_0}$ , vertex

$x_0$  is also in  $L$ . None of  $U_2, \dots, U_p$  can contain  $x_0$  by the definition of  $\mathcal{U}_1$ . Let  $Z$  be a clique of  $T_L$  such that  $Z' \in T_0^{x_0}$ ,  $S_{U_p} \subseteq Z'$ ,  $S_{U_p} \cap Z \neq \emptyset$  and  $S_{U_p} \setminus Z \neq \emptyset$  (such a  $Z$  exists because of  $U_p''$ ). Let  $z \in Z \setminus Z'$ . Vertex  $Q'$  is on  $T[\{L, Z'\}]$  as  $S_{U_p} \subseteq Q'$ . We choose vertices  $y_1, \dots, y_s$  on the labels of  $T_0[Z, Q]$  such that none of them is in  $S_L$  and  $z-y_1-\dots-y_s-q$  is a chordless path.

If  $Z \in T_0^{x_0}$ , then let  $b \in S_{U_p} \setminus Z$ . As  $q$  is special and co-special, we have  $S_Q \not\subseteq S_Z$ , so let  $c \in S_Q \setminus S_Z$ . Then  $z, l, q$  form an asteroidal triple, with paths  $z-y_1-\dots-y_s-q$  and  $l-b-c-q$ , and they lie in the neighborhood of  $x_0$ , a contradiction. So  $Z \notin T_0^{x_0}$ . Let  $x_{r+1} \in Z \cap U_p$ . If  $x_{r+1} \in Q$ , then  $z, l, q$  form an asteroidal triple, with paths  $z-y_1-\dots-y_s-q$  and  $l-x_0-q$ , and they lie in the neighborhood of  $x_{r+1}$ , a contradiction; so  $x_{r+1} \notin Q$ . The  $S_{U_i}$ 's are all included in  $Q'$  and so in  $S_L$  too. They are pairwise disjoint, for otherwise  $T_0$  is not a clique path tree. Let  $l \in L \setminus S_L$ . Vertex  $l$  is not in any of the  $S_{U_i}$ 's, and  $l$  is adjacent to  $x_0, \dots, x_{r+1}$  but none of  $u_1, \dots, u_p, v_0, \dots, v_{p-1}, y_1, \dots, y_s, z, q$ .

Suppose that one of  $x_1, x_0$  is in  $V_0 \cap U_1 \cap Q$ . Then we can assume that  $x_0 = x_1$ . Let  $L_1, L_2 \in \mathcal{L}$  be the leaves of  $T_0^{x_0}$  with  $U_1 \in T_0[\{L_1, Q'\}]$  and  $V_0 \in T_0[\{L_2, Q'\}]$ . Every edge of  $T_L$  is not included in  $S_L$  and so is not included in  $S_{L_1}$ . So  $T_{L_1}$  contains  $T_L$ . If  $L_1 \in \mathcal{L}^*$ , then  $T_{L_1} = T_L$  by the maximality of  $T_L$ . But then  $L_1'$  is not in  $T_{L_1}$ , which contradicts Claim 1. So  $L_1 \notin \mathcal{L}^*$  and  $L_2 \in \mathcal{L}^*$ . Every edge of  $T_L$  is not included in  $S_L$  and so is not included in  $S_{V_0}$  and also not in  $S_{L_2}$ . So  $T_{L_2}$  contains  $T_L$ . Vertex  $x_{r+1} \notin S_{V_0}$ , so  $x_{r+1} \notin S_{L_2}$ , so  $S_L \not\subseteq S_{L_2}$ , so  $T_{L_2}$  contains  $L$ , which contradicts the maximality of  $T_L$ . So  $x_0 \neq x_1$ ,  $x_0 \notin U_1$ ,  $x_1 \notin Q$ .

If  $s = 1$ , then  $u_1, \dots, u_p, v_0, \dots, v_{p-1}, x_0, \dots, x_{r+1}, y_1, q, z, l$  induce  $F_{14}(4p+5)_{p \geq 1}$ . If  $s = 2$ , then  $u_1, \dots, u_p, v_0, \dots, v_{p-1}, x_0, \dots, x_{r+1}, y_1, y_2, q, z, l$  induce  $F_{15}(4p+6)_{p \geq 1}$ . If  $s \geq 3$ , then  $x_0, x_{r+1}, y_1, \dots, y_s, q, z, l$  induce  $F_{10}(s+5)_{s \geq 3}$ . A contradiction.

Therefore we have  $S_U \subseteq U''$  for every  $U \in \mathcal{U}_p$ ,  $1 \leq p \leq k-1$ .

Suppose that  $k$  is infinite. Then, the  $U_i$ 's are pairwise disjoint, for otherwise  $T_0$  is not a clique path tree as  $S_{U_i} \subseteq Q'$ . For each  $V \in \bigcup_{p \geq 0} \mathcal{V}_p$ , we add the edge  $VL$  between  $T_V$  and  $T$ . For each  $U \in \bigcup_{p \geq 1} \mathcal{U}_p$ , we add the edge  $UU''$  between  $T_U$  and  $T$ . For each  $U \in \mathcal{U} \setminus (\bigcup_{p \geq 1} \mathcal{U}_p)$ , we add the edge  $UL$  between  $T_U$  and  $T$ . For each  $V \in \mathcal{V} \setminus (\bigcup_{p \geq 1} \mathcal{V}_p)$ , we define  $V'' \in \mathcal{Q}(T)$  such that  $V'' \cap S_V \neq \emptyset$  and the length of  $T[\{L, V''\}]$  is maximal. By the definition of  $\mathcal{V}$ , we have  $S_V \cap Q = \emptyset$ , so  $V'' \neq Q$ , so  $V''$  is a vertex of  $T_L$  on  $T_0[L, V]$  so it contains all  $S_V$  as  $S_V \subseteq S_L$ . Then we can add the edge  $VV''$  between  $T_V$  and  $T$  to obtain a clique path tree of  $G$ , a contradiction.

So  $k$  is finite and  $\geq 2$ .

As before, let  $V_0, \dots, V_{k-1}, U_1, \dots, U_k$  be such that  $V_i \in \mathcal{V}_i$ ,  $U_i \in \mathcal{U}_i$ ,  $V_{i-1} \cap U_i \neq \emptyset$ , and  $U_i \cap V_i \neq \emptyset$ . Let  $u_i \in U_i \setminus S_{U_i}$ , let  $v_i \in V_i \setminus S_{V_i}$ . Let  $x_1, \dots, x_r$  be such that  $x_1 \in V_0 \cap U_1$ ,  $x_2 \in U_1 \cap V_1$ ,  $\dots$ ,  $x_r \in V_{k-1} \cap U_k$  with  $r = 2k-1$ . None of  $x_2, \dots, x_r$  is in  $Q$ . Let  $x_0 \in V_0 \cap Q$ .

Suppose that one of  $x_1, x_0$  is in  $V_0 \cap U_1 \cap Q$ . Then we can assume that  $x_0 = x_1$ . Let  $L_1, L_2 \in \mathcal{L}$  be the leaves of  $T_0^{x_0}$  with  $U_1 \in T_0[\{L_1, Q'\}]$  and  $V_0 \in T_0[\{L_2, Q'\}]$ . As before, we have  $L_2 \in \mathcal{L}^*$ . Every edge of  $T_L$  is not included in  $S_L$  so it is also not in  $S_{V_0}$  and not in  $S_{L_2}$ . So  $T_{L_2}$  contains  $T_L$ . We have  $k \geq 2$ , so  $U_2$  exists and there is a vertex  $x_2 \in U_1 \cap V_1$ . Vertex  $x_2$  is not in  $S_{V_0}$ , so it is not in  $S_{L_2}$ , so  $S_L \not\subseteq S_{L_2}$ , so  $T_{L_2}$  contains

$L$ , which contradicts the maximality of  $T_L$ . So  $x_0 \neq x_1$ ,  $x_0 \notin U_1$ ,  $x_1 \notin Q$ .

Let  $s_{U_k} \in S_{U_k} \setminus Q'$ . Vertex  $s_{U_k}$  is not adjacent to any of  $q, s_Q, v_0, \dots, v_{k-1}$  because  $s_{U_k} \notin Q'$ , and by the minimality of  $k$ , vertex  $s_{U_k}$  is not adjacent to  $u_1, \dots, u_{k-1}$ . Then  $u_1, \dots, u_k, v_0, \dots, v_{k-1}, x_0, \dots, x_r, s_{U_k}, s_Q, q$  induce  $F_{16}(4k+3)_{k \geq 2}$ . A contradiction.  $\square$

Let  $U \in \mathcal{U}_1$  be such that  $S_U \setminus Q' \neq \emptyset$ . Let  $s_U \in S_U \setminus Q'$ . Vertex  $s_U$  is not adjacent to  $s_Q$ . Let  $u \in U \setminus S_U$ . Let  $W \in \mathcal{W}$  be such that  $U \cap W \neq \emptyset$ . Let  $w \in W \setminus S_W$ .

**Claim 3**  $S_W = S_L$ .

*Proof.* Suppose  $S_W \neq S_L$ . Then  $S_W \subsetneq S_L$ . By the definition of  $\mathcal{W}$ , there exists  $a \in W \cap Q$  and so  $a \in L$ .

Suppose  $S_W \subseteq S_U$ . Then  $a$  is in  $U$ . Let  $L_1, L_2 \in \mathcal{L}$  be the leaves of  $T_0^a$  with  $U \in T_0[\{L_1, Q'\}]$  and  $W \in T_0[\{L_2, Q'\}]$ . Every edge of  $T_L$  is not included in  $S_L$  so it is also not in  $S_{L_1}$ . By the definition of  $\mathcal{L}$ , the set  $S_L$  is not included in  $S_{L_1}$ . So  $T_{L_1}$  is strictly greater than  $T_L$ . So  $L_1 \notin \mathcal{L}^*$ , so  $L_2 \in \mathcal{L}^*$ . Every edge of  $T_L$  is not included in  $S_L$ , so it is also not in  $S_W$  and not in  $S_{L_2}$ . The same goes for  $S_L$  by the hypothesis. So  $T_{L_2}$  is strictly greater than  $T_L$ , which contradicts the definition of  $L$ . So  $S_W \not\subseteq S_U$ . Let  $b$  be a vertex of  $S_W \setminus S_U$ .

Vertex  $s_U$  is in  $S_U \setminus Q'$ , so  $s_U \notin S_W$ . Every labels of the edges of  $T_L$  is not included in  $S_L$ , so it is also not in  $S_W$ . So we can choose vertices  $x_1, \dots, x_r$  on the labels of  $T_0'[\{U, Q\}]$  such that none of the  $x_i$ 's is in  $S_W$ ,  $x_1 \in U$ ,  $x_r \in Q$  and  $x_1 \cdots x_r$  is a chordless path.

Then  $w-b-s_q-q$  is a path from  $w$  to  $q$  that avoids  $N(u)$ ; and  $u-x_1 \cdots x_r-q$  is a path from  $u$  to  $q$  that avoids  $N(w)$ . So  $q, u, w$  form an asteroidal triple. By Lemma 4, we have  $Q \cap U \cap V = \emptyset$ . So  $a \notin U$ .

Let  $c \in U \cap W$  (by the definition of  $\mathcal{U}_1$ ). Vertex  $c$  is not in  $Q$ . Let  $r = 1$ . Then  $x_1$  is different from  $s_U$  and  $s_Q$ , and  $q, u, w, a, c, s_Q, s_U, x_1$  induce  $F_8$ . Let  $r = 2$ . If  $x_1$  is adjacent to  $s_Q$ , then  $q, u, w, a, c, s_Q, s_U, x_1$  induce  $F_9$ ; and if  $x_1$  is not adjacent to  $s_Q$ , then  $q, u, w, a, c, s_Q, x_1, x_2$  induce  $F_9$ . If  $r \geq 3$ , then  $q, u, w, a, c, x_1, \dots, x_r$  induce  $F_{10}(r+5)_{r \geq 3}$ . In all cases we obtain a contradiction.  $\square$

**Claim 4**  $W \in \mathcal{L}^*$

*Proof.* Suppose that  $W \notin \mathcal{L}^*$ . Let  $a \in W \cap Q$ , we have  $a \in L$ . Let  $L_1, L_2 \in \mathcal{L}$  be the leaves of  $T_0^a$ , with  $L \in T_0[\{L_1, Q'\}]$  and  $W \in T_0[\{L_2, Q'\}]$ . Let  $K \in T_0[\{L_2, W\}] \cap \mathcal{L}$  be such that the length of  $T_0[\{K, W\}]$  is minimal. If  $W \in \mathcal{L}$ , then  $T_W = T_L$  and  $W \in \mathcal{L}^*$ , a contradiction. So  $W \notin \mathcal{L}$ , so  $W \neq K$ . The edges of  $T_L$  are not included in  $S_L$ , so they are also not in  $S_W$  and not in  $S_K$ . So  $T_K$  contains  $T_L$ . If  $K \in \mathcal{L}^*$ , then  $T_K = T_L$  by the maximality of  $T_L$ , which contradicts Claim 1; so  $K \notin \mathcal{L}^*$ . So  $T_K = T_0' \setminus K$  and the labels of  $T_0' \setminus K$  are not included in  $S_K$ , so  $S_W \not\subseteq S_K$ . Let  $X$  be the edge of  $T_0[\{W, K\}]$  such that  $X'$  contains  $S_W$  and  $X$  does not (maybe  $X' = W$ ,  $X = K$ ). The set  $S_X$  contains  $a$  but not all of  $S_X'$ . So no element of  $S_X' \setminus S_X$  contains  $S_X'$ . So  $X' \in \mathcal{L}$ , a contradiction to the definition of  $K$ .  $\square$

By claim 4, we have  $W \in \mathcal{L}^*$ . Then  $T_W = T_L$  is also maximal and what we have proved for  $L$  can be done for  $W$ . By Claim 2 we know that there exists  $X \notin T_L$  such that  $XW$  is an edge of  $T_0$  with  $S_X \cap S_W \neq \emptyset$  and  $S_X \setminus Q' \neq \emptyset$ . Let  $x \in X \setminus W$ . Let  $s_X \in S_X \setminus Q'$ . Vertex  $s_X$  is not in  $S_W$ , for otherwise it would also be in  $S_L$  and in  $Q'$ . Vertex  $s_Q$  is not in  $S_L$ , so not in  $S_W$ . So  $s_Q, s_X$  are not adjacent. We distinguish between two cases.

*Case 1:*  $U \cap X = \emptyset$ . Let  $a \in U \cap W$ , so  $a \notin X$ . Suppose  $a \notin Q$ . If there exists  $b \in X \cap Q$  then  $b$  is also in  $L$  and  $q, u, x, s_Q, s_U, s_X, a, b$  induces  $F_6$ , a contradiction. So there is no vertex in  $X \cap Q$ . Let  $c \in W \cap Q$ , we have  $c \in L$  and  $c \notin X$ . Let  $d \in X \cap S_W$ , we have  $d \notin Q$ . If  $c$  is adjacent to  $U$ , then  $q, u, x, s_Q, s_U, s_X, c, d$  induces  $F_6$ , if  $c$  misses  $u$ , then  $q, u, x, s_Q, s_U, s_X, a, c, d$  induce  $F_7$ , a contradiction. So  $a \in Q$ . Let  $e \in X \cap S_W$ . If  $e \notin Q$  then  $q, u, x, s_Q, s_U, s_X, a, e$  induce  $F_6$ , a contradiction, so  $e \in Q$ . Let  $f \in S_W \setminus S_Q$  (as  $q$  is special and co-special); maybe  $f$  is in  $U$  or  $X$ , but not in both. Then  $q, u, x, s_U, s_X, a, e, f$  induce  $F_9$  or  $F_{10}(8)$ , according to whether  $f$  is adjacent to none or exactly one of  $u, x$ , a contradiction.

*Case 2:*  $U \cap X \neq \emptyset$ . Suppose  $U \cap X \cap Q \neq \emptyset$ . Let  $z \in U \cap X \cap Q$ . Let  $L_1, L_2 \in \mathcal{L}$  be the leaves of  $T_0^z$ . Let  $i \in \{1, 2\}$  be such that  $L_i \in \mathcal{L}^*$ . The edges of  $T_L$  are not included in  $S_L = S_W$ , thus also not in  $S_{L_i}$ . So  $T_{L_i}$  contains  $T_L$ , so  $T_{L_i} = T_L$  by maximality of  $T_L$ . But this contradicts Claim 1. So  $U \cap X \cap Q = \emptyset$ .

Let  $a \in U \cap X$ . Vertex  $a$  is not in  $Q$ . Let  $b \in W \cap Q$ . If  $b \notin X \cup U$ , then  $q, u, x, s_Q, s_U, s_X, a, b$  induce  $F_6$ , a contradiction. If  $b \in X$ , then  $b \notin U$ . Let  $c \in S_W \setminus S_X$ . If  $c \in U \setminus Q$ , then  $q, u, x, s_Q, s_U, s_X, b, c$  induce  $F_6$ . If  $c \in Q \setminus U$ , then  $q, u, x, s_Q, s_U, s_X, a, c$  induce  $F_6$ . If  $c \in U \cap Q$ , then  $q, u, x, s_Q, s_U, s_X, a, b, c$  induce  $F_8$ . If  $c \notin U \cup Q$ , then  $q, u, x, s_Q, s_U, a, b, c$  induce  $F_{10}(8)$ . A contradiction. If  $b \in U$ , then  $b \notin X$ . Let  $d \in S_L \setminus S_U$ . If  $d \in X \setminus Q$ , then  $q, u, x, s_Q, s_U, s_X, b, d$  induce  $F_6$ . If  $d \in Q \setminus X$ , then  $q, u, x, s_Q, s_U, s_X, a, d$  induce  $F_6$ . If  $d \in U \cap Q$ , then  $q, u, x, s_Q, s_U, s_X, a, b, d$  induce  $F_8$ . If  $d \notin U \cup Q$ , then  $q, u, x, s_Q, s_U, a, b, d$  induce  $F_{10}(8)$ . A contradiction.

This ends the proof of Theorem 3. □

## 6 Recognition algorithm

Gavril [4] and Schaffer [9] gave polynomial time algorithms to recognize path graphs. The characterization that we give in this paper suggests a new recognition algorithm, which takes any graph  $G$  as input and either builds a clique path tree for  $G$  or finds one of  $F_0, \dots, F_{16}$ . We have not analyzed the exact complexity of such a method but it will give a new polynomial algorithm to recognize path graphs.

There are linear time recognition algorithms for interval graphs [1] and triangulated graphs [8] but surprisingly not for path graphs. One can hope that the work presented here will be helpful in the search for a linear time recognition algorithm for path graphs.

## References

- [1] K.S. Booth, G.S. Lueker, Testing for the consecutive ones property, interval graphs and graph planarity using PQ-tree algorithm, *J. Comput. Syst. Sci.* 13 (1976) 335–379.
- [2] G.A. Dirac, On rigid circuit graphs, *Abh. Math. Sem. Univ. Hamburg* 38 (1961) 71–76.
- [3] F. Gavril, The intersection graphs of subtrees in trees are exactly the chordal graphs, *J. Combin. Theory B* 16 (1974) 47–56.
- [4] F. Gavril, A recognition algorithm for the intersection graphs of paths in trees, *Discrete Math.* 23 (1978) 211–227.
- [5] C. Lekkerkerker, D. Boland, Representation of finite graphs by a set of intervals on the real line, *Fund. Math.* 51 (1962) 45–64.
- [6] C.L. Monma, V.K. Wei, Intersection graphs of paths in a tree, *J. Combin. Theory B* 41 (1986) 141–181.
- [7] P.L. Renz, Intersection representations of graphs by arcs, *Pacific J. Math.* 34 (1970) 501–510.
- [8] D.J. Rose, R.E. Tarjan, G.S. Lueker, Algorithmic aspects of vertex elimination of graphs, *SIAM J. Comput.* 5 (1976) 266–283.
- [9] A.A. Schäffer, A faster algorithm to recognize undirected path graphs, *Discrete Appl. Math.* 43 (1993) 261–295.
- [10] R.E. Tarjan, M. Yannakakis, Simple linear time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs, *SIAM J. Comput.* 13 (1984) 566–579.

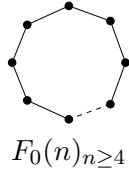


Figure 1: Forbidden subgraphs with no simplicial vertices

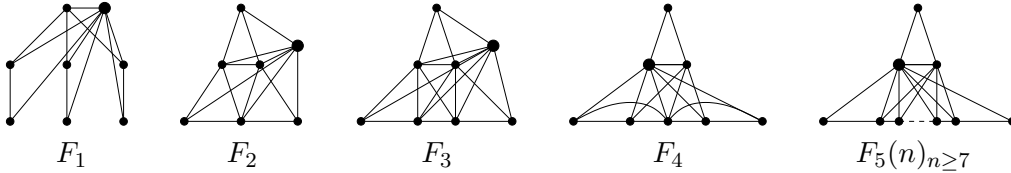


Figure 2: Forbidden subgraphs with a universal vertex

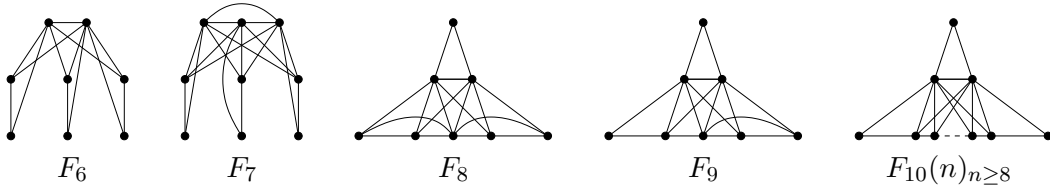


Figure 3: Forbidden subgraphs with no universal vertex and exactly three simplicial vertices

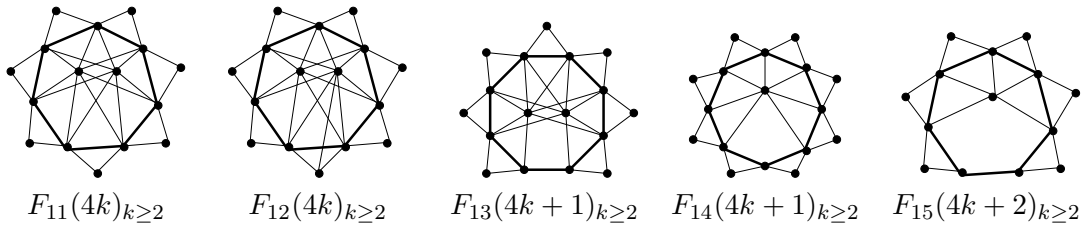


Figure 4: Forbidden subgraphs with at least one simplicial vertex that is not co-special. (bold edges form a clique)

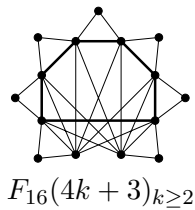


Figure 5: Forbidden subgraphs with  $\geq 4$  simplicial vertices that are all co-special. (bold edges form a clique)