Characterizing Path Graphs by Forbidden Induced Subgraphs

Benjamin Lévêque,¹ Frédéric Maffray,² and Myriam Preissmann²

¹ROSE, EPFL, LAUSANNE, SWITZERLAND
E-mail: benjamin.leveque@free.fr

²C.N.R.S., LABORATOIRE G-SCOP, GRENOBLE, FRANCE
E-mail: frederic.maffray@g-scop.inpg.fr, myriam.preissmann@g-scop.inpg.fr

Received May 6, 2008; Revised January 27, 2009
Published online 6 April 2009 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/jgt.20407

Abstract: A path graph is the intersection graph of subpaths of a tree. In 1970, Renz asked for a characterization of path graphs by forbidden induced subgraphs. We answer this question by determining the complete list of graphs that are not path graphs and are minimal with this property.

Keywords: intersection graphs; path graphs; forbidden induced subgraphs

1. INTRODUCTION

All graphs considered here are finite and have no parallel edges and no loop. A hole is a chordless cycle of length at least four. A graph is chordal (or triangulated) if it contains no hole as an induced subgraph. Gavril [7] proved that a graph is chordal if
and only if it is the intersection graph of a family of subtrees of a tree. In this paper, whenever we talk about the intersection of subgraphs of a graph we mean that the vertex sets of the subgraphs intersect.

An *interval graph* is the intersection graph of a family of intervals on the real line; equivalently, it is the intersection graph of a family of subpaths of a path. An *astereoidal 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 [13] proved that a graph is an interval graph if and only if it is chordal and contains no asteroidal triple. They derived from this result the list of minimal forbidden subgraphs for interval graphs.

An intermediate 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 [8, 15, 17] but no characterization by forbidden subgraphs was known, whereas such results exist for intersection graphs of subpaths of a path (interval graphs [13]), subtrees of a tree (chordal graphs [7]), and also for directed subpaths of a directed tree (directed path graphs [16]).

In 1970, Renz [17] asked for a complete list of graphs that are chordal, not path graphs, and are minimal with this property, and he gave two examples of such graphs. The list of minimal forbidden subgraphs for path graphs was extended in [21], but that list is incomplete. Here, we answer Renz’s question and obtain a characterization of path graphs by forbidden induced subgraphs. We will prove that the graphs presented in Figures 1–5 are all the minimal non-path graphs. In other words:

**Theorem 1.** A graph is a path graph if and only if it does not contain any members of the families of $F_0, \ldots, F_{16}$ as an induced subgraph.

We could not find a characterization similar to the one found by Lekkerkerker and Boland [13] 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
characterizing path graphs by forbidden induced subgraphs

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).

Asteroidal triple; but this condition is not sufficient. So we prove directly that a graph that does not contain any of the excluded subgraphs is a path graph. The initial proof of the characterization of interval graphs by Lekkerkerker and Boland [13] was fairly complicated. It was simplified by Halin [12] by using the concept of prime graph decomposition. Cameron et al. [3] translated Halin’s proof in terms of clique tree. Our proof is, in its principle, a generalization of the proof presented in [3].

2. Special Simplicial Vertices in Chordal Graphs

In a graph \( G \), a clique is a set of pairwise adjacent vertices. Let \( Q(G) \) be the set of all (inclusionwise) maximal cliques of \( G \). When there is no ambiguity we will write \( Q \) instead of \( Q(G) \).

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 \( S(G) \) be the set of separators of \( G \). When there is no ambiguity we will write \( S \) instead of \( S(G) \).

The neighborhood of a vertex \( v \) is the set \( N(v) \) of vertices adjacent to \( v \). For a set \( X \) of vertices, let \( N(X) = (\bigcup_{v \in X} N(v)) \setminus X \). Let us say that a vertex \( u \) is complete to a set \( X \)
of vertices if \( X \subseteq N(u) \). A vertex is simplicial if its neighborhood is a clique. It is easy to see that a vertex is simplicial if and only if it does not belong to any separator. 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, \( Q_v \) is the unique maximal clique containing \( v \). Remark that \( S_v \) is not necessarily in \( S \); for example, in the graph \( H \) with vertices \( a, b, c, d, e \) and edges \( ab, bc, cd, de, bd \), we have \( S_v = \{b, d\} \) and \( S(H) = \{\{b\}, \{d\}\} \).

A classical result \([1, 11]\) (see also \([9]\)) states that, in a chordal graph \( G \), every separator is a clique; moreover, if \( S \) is a separator, then there are at least two components of \( G \setminus S \) that contain a vertex that is complete to \( S \), and so \( S \) is the intersection of two maximal cliques.

A clique tree \( T \) of a graph \( G \) is a tree whose vertices are the members of \( Q \) and such that, for each vertex \( v \) of \( G \), those members of \( Q \) that contain \( v \) induce a subtree of \( T \), which we will denote by \( T_v \). Note that \( G \) is the intersection graph of these subtrees. Gavril \([7]\) proved the classical result that a graph is chordal if and only if it has a clique tree.

Clique trees are very useful when studying chordal graphs or subclasses of chordal graphs as they give the structure of graphs for which they are a clique tree. We recall the definitions and properties of clique trees that we need in the article, but the reader who is not familiar with this notion can refer to classical books of graph theory (like \([9, 14]\)). Our proofs are done in the clique tree. Occasionally, we will have to refer to the original graph (for example, to obtain the forbidden subgraphs explicitly) but most of the time everything can be understood just by studying the clique tree.

In a clique tree \( T \), the label of an edge \( QQ' \) of \( T \) is defined as \( S_{QQ'} = Q \cap Q' \). Note that every edge \( QQ' \) satisfies \( S_{QQ'} \in S \); indeed, there exist vertices \( v \in Q \setminus Q' \) and \( v' \in Q' \setminus Q \) and the set \( S_{QQ'} \) is a \( \{v, v'\} \)-separator. The number of times an element \( S \) of \( S \) appears as a label of an edge is equal to \( c - 1 \), where \( c \) is the number of components of \( G \setminus S \) that contains a vertex complete to \( S \) \([7, 14]\). As pointed out above, \( c \) is at least two; moreover, it depends only on \( S \) and not on \( T \); so, for a given \( S \in S \), the number \( c - 1 \) is the same in every clique tree.

Given a set \( X \subseteq Q \) of maximal cliques, let \( G(X) \) denote the subgraph of \( G \) induced by all the vertices that appear in members of \( X \). If \( T \) is a clique tree of \( G \), then \( T[X] \) denotes the subtree of \( T \) of minimum size such that its set of vertices contains \( X \). Note that if \( |X| = 2 \), then \( T[X] \) is a path.

Given a subtree \( T' \) of a clique-tree \( T \) of \( G \), let \( Q(T') \) be the set of vertices of \( T' \) and \( S(T') \) be the set of separators of \( G(Q(T')) \). It is easy to verify the following important property: \( T' \) is a clique tree of \( G(Q(T')) \). Moreover, if \( T' \neq T \) there exists a leaf \( L \) of \( T \) not in \( T' \) and a vertex in \( L \) that is not in any vertex of \( T' \), so \( G(Q(T')) \) is a strict induced subgraph of \( G \).

Dirac \([6]\) 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 a member of \( S \) and is (inclusionwise) maximal in \( S \).

**Theorem 2.** In a chordal graph that is not a clique, there exist two non-adjacent special simplicial vertices.
\textbf{Proof.} By the hypothesis \( G \) is not a clique, so \(|Q| \geq 2\) and \( S \neq \emptyset \). Let \( T \) be a clique tree of \( G \).

Let us choose, in the set of vertices of \( T \) incident to an edge with (inclusionwise) maximal label, two maximal cliques \( Q_1, Q_2 \) that are at a maximum distance in \( T \). Since \( S \neq \emptyset \) these maximal cliques are distinct.

For \( i = 1, 2 \), let \( Q'_i \) be the neighbor of \( Q_i \) on \( T[Q_1, Q_2] \) (possibly \( Q'_1 = Q'_2 \) or both \( Q'_1 = Q_2 \) and \( Q'_2 = Q_1 \)). By the choice of \( Q_1, Q_2 \), the label \( S_{Q_iQ'_i} \) of \( Q_iQ'_i \) is maximal and no edge of \( T_i \), the subtree of \( T \setminus Q'_i \) that contains \( Q_i \), has a maximal label. So the label of each edge of \( T_i \) is included in \( S_{Q_iQ'_i} \). Let \( v_i \in Q_i \setminus Q'_i \). As \( v_i \) is not in \( S_{Q_iQ'_i} \), it is not in any label of \( T_i \) and so not in any label of \( T \). Thus, \( v_i \) is simplicial and \( Q_{v_i} = Q_i \).

All the labels of the edges incident to \( Q_i \) are included in \( S_{Q_iQ'_i} \), so \( S_{v_i} = S_{Q_iQ'_i} \) and \( v_i \) is special. Since \( Q_{v_1} \) and \( Q_{v_2} \) are distinct cliques, \( v_1 \) and \( v_2 \) are not adjacent. \( \blacksquare \)

Algorithms LexBFS \cite{18} and MCS \cite{20} are linear time algorithms that were developed to find a simplicial elimination ordering in a chordal graph. (A simplicial elimination ordering is an ordering of the vertices \( v_1, \ldots, v_n \) such that, for \( 1 \leq i \leq n \), vertex \( v_i \) is simplicial in the graph induced by vertices \( v_1, \ldots, v_{i-1} \).) The last vertex found by these algorithms is simplicial in the whole graph. This vertex is not necessarily special simplicial. For example, on the graph with vertices \( a, b, c, d, e, f \) and edges \( ab, bc, cd, eb, ec, fb, fc \), every application of LexBFS or MCS will end on one of the simplicial vertices \( a, d \), which are not special. The proof of Theorem 2 can be turned into a polynomial time algorithm to find a special simplicial vertex in a chordal graph. We leave open the problem of finding a special simplicial elimination ordering in linear time. (A special simplicial elimination ordering is a simplicial elimination ordering where vertex \( v_i \) is a special simplicial vertex in the graph induced by vertices \( v_1, \ldots, v_{i-1} \).)

### 3. FORBIDDEN INDUCED SUBGRAPHS

A clique path tree \( T \) of \( G \) is a clique tree of \( G \) such that, for each vertex \( v \) of \( G \), the subtree \( T^v \) induced by the cliques that contain \( v \) is a path. Note that \( G \) is the intersection graph of these subpaths. Gavril \cite{8} proved that a graph is a path graph if and only if it has a clique path tree. A graph \( G \) is a minimal non-path graph if \( G \) is not a path graph but any induced subgraph of \( G \) distinct from \( G \) is a path graph. Note that any induced subgraph of a path graph is also a path graph, so it is enough to require that \( G \setminus v \) is a path graph for every vertex \( v \) of \( G \).

Consider graphs \( F_0, \ldots, F_{16} \) presented in Figures 1–5. Let us make a few remarks about them. Each graph in Figure 2 is obtained by adding a universal vertex to some minimal forbidden subgraph for interval graphs. Clearly, in a path graph the neighborhood of every vertex is an interval graph; so \( F_1, \ldots, F_5 \) are not path graphs. Graphs \( F_{10}(n)_{n \geq 8} \) are also forbidden in interval graphs. Graphs \( F_6 \) and \( F_{10}(8) \) are from Renz \cite[Figures 1 and 5]{17}. For \( i \in \{0, 1, 3, 4, 5, 6, 7, 9, 10, 13, 15, 16\} \), Panda \cite{16} proved that \( F_i \) is a minimal non-directed path graph, so \( F_i \setminus x \) is a directed path graph for every

\textit{Journal of Graph Theory} DOI 10.1002/jgt
vertex x (obviously every directed path graph is a path graph). In general, we have the following:

**Theorem 3.** \( F_0, \ldots, F_{16} \) are families of minimal non-path graphs.

**Proof.** Clearly, \( F_0 \) is a minimal non-path graph. As pointed out above, \( F_1, \ldots, F_5 \) are not path graphs and it is easy to verify that by deleting any vertex of these graphs one obtains a path graph.

We do not give a detailed proof for each graph, but we show a general statement that can be used to prove that \( F_6, \ldots, F_{16} \) are not path graphs. Let \( F = F_i \) for some \( i \in \{6, \ldots, 16\} \). A maximal clique of \( F \) will be called *peripheral* if it contains only one separator \( S \) and \( G \setminus S \) has only two connected components. Such a clique must be a leaf in any clique tree of \( F \). A maximal clique that is not peripheral will be called *central*. In any clique tree of \( F \) the central cliques will induce a subtree. It is easy to see that there is a set \( K \) of vertices of \( F \) such that every central clique of \( F \) contains \( K \) and every peripheral clique intersects \( K \). Therefore, if there is a clique path tree \( T \) of \( F \), then the central cliques must induce a path in \( T \), every peripheral clique must be adjacent to some extremity of this path, and if two peripheral cliques share a common vertex of \( K \) then they must be adjacent to distinct extremities. Another simple remark is that if a vertex of \( F \) is in exactly two maximal cliques, then these two cliques must be adjacent in any clique tree of \( G \).

Let \( P_1, \ldots, P_k \) and \( C_1, \ldots, C_\ell \) be, respectively, the peripheral and the central cliques of \( F \). The following conditions for a path tree of \( F \) are easy to check using our previous remarks:

- When \( F = F_6 \) or \( F_7 \), then \( k = \ell = 3 \) and \( P_i \) must be adjacent to \( C_i \).
- When \( F = F_8, F_9, F_{10}, F_{11}, F_{12}, F_{13}, \) or \( F_{16} \), then \( k \) is odd, for each \( i \in \{1, \ldots, k-1\} \), \( P_i \) and \( P_{i+1} \) must be adjacent to distinct extremities of the path induced by the central cliques, and this holds also for \( P_1 \) and \( P_k \).
- When \( F = F_{14} \) or \( F_{15} \), then \( k \) is even, for each \( i \in \{1, \ldots, k-1\} \), \( P_i \) and \( P_{i+1} \) must be adjacent to distinct extremities of the path induced by the central cliques, and \( P_1 \) and \( P_k \) must be adjacent to the same extremity.

In each of these cases, it is easy to see that no clique tree can satisfy all the conditions.

Furthermore, it is not difficult to check that when we delete any vertex of \( F \) then one of the above constraints is removed and this is sufficient to make it possible to construct a clique path tree. 

**4. CO-SPECIAL SIMPLICIAL VERTICES**

Let us say that a simplicial vertex \( v \) is co-special if \( S_v \) is a separator such that \( G \setminus S_v \) has exactly two components. Lekkerkerker and Boland [13] call this type of vertex strongly simplicial. Note that in that case \( S_v \) is a minimal element of \( S \) and it appears exactly once as a label of any path tree of \( G \). (The fact that \( S_v \) is a minimal element of \( S \) for some simplicial vertex \( v \) does not imply that \( v \) is co-special; for example, consider the graph with vertices \( a, b, c, d \) and edges \( ab, ac, ad \); in fact it has no co-special vertex.)
Also note that, in contrast with Theorem 2, a chordal graph does not necessarily have a simplicial vertex \( v \) where \( S_v \) is a minimal element of \( S \); for example, consider the graph with seven vertices \( a, b, c, d, e, f, g \) and edges \( bc, cd, ef, fg, ab, ac, ad, ae, af, ag \).

**Lemma 1.** Let \( G \) be a minimal non-path graph. Then either \( G \) is one of \( F_{11}, \ldots, F_{15} \) or every simplicial vertex of \( G \) is co-special.

**Proof.** Suppose on the contrary that \( G \) is a minimal non-path graph, different from \( F_{11}, \ldots, F_{15} \), and there is a simplicial vertex \( q \) of \( G \) that is not co-special. All simplicial vertices of \( F_9, \ldots, F_{10}, F_{16} \) 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 minimal). Therefore, \( G \) contains none of \( F_9, \ldots, F_{16} \).

The graph \( G \) is not a clique as it is not a path graph. Let us denote by \( Q \) the unique maximal clique that contains \( q \). Consider graph \( G(Q \setminus Q) \), which is equal to \( G(Q \setminus S_q) \). Since \( Q \setminus S_q \neq \emptyset \), and by the minimality of \( G \), it follows that \( G(Q \setminus Q) \) admits a clique path tree \( T_0 \). Let \( Q' \) be a vertex of \( T_0 \) such that \( S_q \subseteq Q' \). If \( Q' = S_q \), then by replacing \( Q' \) by \( Q \) in \( T_0 \) we obtain a clique path tree of \( G \), a contradiction. So \( Q' \neq S_q \). Let \( q' \in Q' \setminus S_q \), then \( S_{q'} = Q \cap Q' \) is a \( \{q, q'\} \)-separator. Let \( T_0' \) be obtained from \( T_0 \) by adding vertex \( Q \) and edge \( QQ' \). Remark that \( T_0' \) is a clique tree of \( G \) but not a clique path tree since \( G \) is not a path graph.

Let \( T' \) be the maximal subtree of \( T_0' \) that contains \( Q \) and \( Q' \) and such that the edge \( QQ' \) is the only edge whose label is included in \( S_q \). So \( T' \) is a clique tree of \( G(Q(T')) \). Since \( q \) is not co-special, there is an edge of \( T_0 \) whose label is included in \( S_q \), and so \( T' \) is a strict subgraph of \( T_0' \). So \( G(Q(T')) \) is a strict subgraph of \( G \) and by the minimality of \( G \) it is a path graph. Let \( T \) be a clique path tree of this graph. From now on, our goal will be to show that either \( G \) contains one of the forbidden subgraphs or \( T \) can be extended into a clique path tree of \( G \).

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

Let \( T_1, \ldots, T_\ell \) 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 \( Q_1, \ldots, Q_\ell \) are pairwise disjoint (but \( Q', Q_1, \ldots, Q_\ell \) are not necessarily pairwise disjoint). Let \( S_i = Q_i \cap Q'_i \) and \( v_i \in Q_i \setminus Q'_i \). Let \( \mathcal{H} \) be the intersection graph of \( S_1, \ldots, S_\ell \), that is, \( \mathcal{H} \) has vertex-set \( V_{\mathcal{H}} = \{S_1, \ldots, S_\ell \} \) and edge-set \( E_{\mathcal{H}} = \{S_iS_j \mid S_i \cap S_j \neq \emptyset \} \).

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

**Proof.** Suppose on the contrary, without loss of generality, that \( S_1 \cdots S_p - S_1 \) is an odd cycle in \( \mathcal{H} \), with length \( p = 2r + 1 \) \((r \geq 1)\). Let \( I_j = S_j \cap S_{j+1} \) \((j = 1, \ldots, p)\) 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 \( T_0 \). Therefore, we have \( I_j \cap I_k = \emptyset \) whenever \( j \neq k \). 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 \( S_j \subseteq S_q \) for each \( 1 \leq j \leq p \), so the \( s_j \)'s are all...
in \( Q \) and \( Q' \). Let us consider the subgraph induced by \( q, q', v_1, \ldots, v_p, s_1, \ldots, s_p \). Both \( q \) and \( q' \) are 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 \( q \). Vertex \( q' \) has no neighbor among the \( v_j \)'s, for otherwise \( q' \) is in some \( S_j \) and then also in \( S_q \subseteq Q \), a contradiction to its definition. Now \( \{q, q', v_1, \ldots, v_p, s_1, \ldots, s_p\} \) induces \( F_{11}(4r+4)_{r\geq1} \), a contradiction. Thus the claim holds.

By the preceding claim, \( \mathcal{H} \) is a bipartite graph.

For \( 1 \leq i \leq \ell \), let \( R_i = \{S \in S(T')| S_i \cap S \neq \emptyset \text{ and } S_i \setminus S \neq \emptyset \} \). Let \( X = \{S_i| R_i \neq \emptyset \} \). We remark that \( S_Q \notin R_i \).

**Claim 2.** There is no odd path between two vertices of \( X \) in \( \mathcal{H} \).

**Proof.** Suppose on the contrary, without loss of generality, that \( S_1 \cdots S_p \) is an odd path in \( \mathcal{H} \) between two vertices \( S_1, S_p \) of \( X \) (with \( p = 2k \), \( k \geq 1 \)), 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 \( S_j \setminus S_{j+1} \). As in the preceding claim, the \( s_j \)'s are pairwise distinct and lie in \( Q \) and \( Q' \). Let \( P \) be the path \( T'[Q_1', Q_p'] \). We note that when \( p > 2 \), then \( S_2 \) is not in \( X \), so \( Q_3' = Q_1' \), for otherwise \( T_0^{2} \) would not be a path; then \( S_3 \) is not in \( X \), so \( Q_4' = Q_2' \), and so on. Thus, the two extremities of \( P \) are \( Q_1' = Q_3' = \cdots = Q_{p-1}' \) and \( Q_2' = Q_4' = \cdots = Q_p' \). Since \( S_1 \) and \( S_p \) are in \( X \), the sets \( R_1, R_p \) are non-empty.

Let \( L_1 \) be the closest vertex to \( Q_1' \) in \( P \) such that there exists an edge incident to \( L_1 \) with label in \( R_1 \), and let \( L_1K_1 \) be such an edge and \( R_1 \) be its label (such an edge exists because \( R_1 \neq \emptyset \)). Similarly, let \( L_p \) be the closest vertex to \( Q_p' \) in \( P \) such that there exists an edge incident to \( L_p \) with label in \( R_p \), and let \( L_pK_p \) be such an edge and \( R_p \) be its label. So \( S_1 \subseteq L_1, S_1 \not\subseteq K_1 \) and \( S_p \subseteq L_p, S_p \not\subseteq K_p \). Each of \( K_1, K_p \) may be in \( P \) or not. Since \( T' \setminus Q = Q' \) is a clique path tree, \( Q' \) lies between \( Q_1' \) and \( L_1 \) and between \( L_p \) and \( Q_p' \) along \( P \). So \( Q_1', L_p, Q', L_1, Q_p' \) lie in this order on \( P \), and \( S_1 \) is included in all labels between \( Q_1' \) and \( L_1 \) in \( P \), and \( S_p \) is included in all labels between \( Q_p' \) and \( L_p \) in \( P \).

Let \( v_0 \in K_1 \setminus L_1 \) and \( v_{p+1} \in K_p \setminus L_p \). Since \( T_0' \) is a clique tree, \( v_0 \) and \( v_{p+1} \) are distinct from \( v_1, \ldots, v_p \) and not adjacent to \( q \).

Let \( s_0 \in S_1 \cap R_1 \) and \( s_p \in S_p \cap R_p \). Then \( v_0 \) and \( s_0 \) are adjacent, and \( v_{p+1} \) and \( s_p \) are adjacent. Since \( T_0 \) is a clique path tree, if \( K_1 \) or \( K_p \) is not in \( P \), then \( s_0 \) and \( s_p \) are different from each other, from \( s_1, \ldots, s_{p-1} \) and from \( v_0, \ldots, v_{p+1} \). Furthermore, if \( K_1 \) is not in \( P \), then \( v_0 \) is not adjacent to any of \( s_1, \ldots, s_p \); and if \( K_p \) is not in \( P \), then \( v_{p+1} \) is not adjacent to any of \( s_0, \ldots, s_{p-1} \).

Let \( s_0' \in S_1 \setminus R_1 \) and \( s_p' \in S_p \setminus R_p \). Then \( v_0 \) and \( s_0' \) are not adjacent, and \( v_{p+1} \) and \( s_p' \) are not adjacent. Since \( T_0 \) is a clique path tree, if \( K_1 \) or \( K_p \) is in \( P \), then \( s_0' \) and \( s_p' \) are different from each other, from \( s_1, \ldots, s_{p-1} \) and from \( v_0, \ldots, v_{p+1} \). Furthermore, if \( K_1 \) is in \( P \), then \( v_0 \) is adjacent to \( s_p' \) and to \( s_1, \ldots, s_p \); and if \( K_p \) is in \( P \), then \( v_{p+1} \) is adjacent to \( s_0' \) and to \( s_0, \ldots, s_{p-1} \).

Note that \( \{q, s_0', s_0, s_1, s_2, \ldots, s_p, s_p'\} \) induces a clique in \( G \). Moreover, \( v_1 \) is adjacent to \( s_0' \), \( v_p \) is adjacent to \( s_p' \), for \( i = 1, \ldots, p \), \( v_i \) is adjacent to \( s_{i-1} \) and \( s_i \), and there is no other edge between \( v_1, \ldots, v_p \) and that clique.
Suppose that $K_1 = K_p$. Then $L_1 = L_p = Q'$ and $K_1$ is not in $P$. By the definition of $T'$, there exists $y \in R_1 \setminus S_q$. Vertex $y$ is distinct from all $s_i$’s as it is not in $S_p$, and it is adjacent to all of $v_0, s_0, \ldots, s_p$ and to none of $q, v_1, \ldots, v_p$. Then $\{q, y, v_0, \ldots, v_p, s_0, \ldots, s_p\}$ induces $F_{12}(4k + 4)_{k \geq 1}$, a contradiction. So $K_1 \neq K_p$, and $v_0$ and $v_{p+1}$ are distinct non-adjacent vertices. We can choose vertices $x_1, \ldots, x_r \ (r \geq 1)$ not in $S_q$ and on the labels of $T'[K_1, K_p]$ such that $v_0 - x_1 - \cdots - x_r - v_{p+1}$ is a chordless path in $G$. Vertices $x_1, \ldots, x_r$ are distinct from and adjacent to $s_0' - s_p', s_0, \ldots, s_p$, and they are distinct from and not adjacent to any of $q, v_1, \ldots, v_p$.

Suppose that $L_1 = Q_p'$ and $L_p = Q_1'$. Then $K_1$ and $K_p$ are not in $P$. If $r = 1$, then $\{q, v_0, \ldots, v_{p+1}, s_0, \ldots, s_p, x_1\}$ induces $F_{15}(4k + 5)_{k \geq 1}$. If $r = 2$, then $\{q, v_0, \ldots, v_{p+1}, s_0, \ldots, s_p, x_1, x_2\}$ induces $F_{15}(4k + 6)_{k \geq 1}$. If $r \geq 3$, then $\{q, v_0, v_{p+1}, s_0, s_p, x_1, \ldots, x_r\}$ induces $F_{10}(r + 5)_{r \geq 3}$, a contradiction.

Suppose now that $L_1 \neq Q_p'$ and $L_p = Q_1'$. Then $K_p$ is not in $P$ and we may assume that $K_1$ is in $P$. If $r = 1$, then $\{q, v_0, \ldots, v_{p+1}, s_0', s_1, \ldots, s_p, x_1\}$ induces $F_{13}(4k + 5)_{k \geq 1}$. If $r \geq 2$, then $\{q, v_0, v_{p+1}, x_1, \ldots, x_r, s_0', s_p\}$ induces $F_{5}(r + 5)_{r \geq 2}$, a contradiction.

Suppose finally that $L_1 \neq Q_p'$ and $L_p \neq Q_1'$. Then we may assume that $K_1$ and $K_p$ are in $P$. If $r = 1$, then $\{q, v_0, v_{p+1}, s_0', s_1, s_p', x_1\}$ induces $F_2$. If $r = 2$, then $\{q, v_0, v_{p+1}, s_0', s_1, s_p', x_1, x_2\}$ induces $F_3$. If $r \geq 3$, then $\{q, v_0, v_{p+1}, x_1, \ldots, x_r, s_0', s_p'\}$ induces $F_{10}(r + 5)_{r \geq 3}$, a contradiction. Thus the claim holds.

By the preceding two claims, $\mathcal{H}$ is a bipartite graph, so its vertex-set can be partitioned into two stable sets $A_{\mathcal{H}}, B_{\mathcal{H}}$, and we may assume that $X \subseteq A_{\mathcal{H}}$. Now all the subtrees $T_i$ can be linked to $T$ to get a clique path tree of $G$ as follows. For each $S_i \in A_{\mathcal{H}}$, we add an edge $Q_i Q_j$ between $T$ and $T_i$. This creates a clique path tree on the corresponding subset of cliques because $A_{\mathcal{H}}$ is a stable set of $\mathcal{H}$ and $Q$ is a leaf of $T$. For each $S_i \in B_{\mathcal{H}}$, let $Q_i' \in Q(T)$ be such that $Q_i' \cap S_i \neq \emptyset$ and the length of $T(Q, Q_i')$ is maximal. Since $S_i \in B_{\mathcal{H}}$, we have $R_i = \emptyset$, so $S_i \subseteq Q_i'$ and we can add an edge $Q_i' Q_j$ between $T$ and $T_i$. This creates a clique path tree of $G$ because $B_{\mathcal{H}}$ is a stable set of $\mathcal{H}$ and by the definition of $Q_i'$, a contradiction.

5. CHARACTERIZATION OF PATH GRAPHS

In this section we prove the main theorem, that is, path graphs are exactly the graphs that do not contain any of $F_0, \ldots, F_{16}$.

**Lemma 2.** In a graph that does not contain any member of the families of $F_0, \ldots, F_5, F_{10}, F_{16}$, the neighborhood of every vertex does not contain an asteroidal triple.

**Proof.** Suppose that in a graph $G$ the neighborhood of some vertex $v$ contains an asteroidal triple. Then, by [13], the neighborhood contains a minimal forbidden induced subgraph $H$ for interval graphs. Then $H$ and $v$ induce one of $F_0, \ldots, F_5, F_{10}$ in $G$.

Given three non-adjacent vertices $a, b, c$, we say that $a$ is in 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 in the middle of the others.
Lemma 3. In a chordal graph $G$ with clique tree $T$, a vertex $a$ is in the middle of two vertices $b,c$ if and only if for all maximal cliques $Q_b$ and $Q_c$ with $b \in Q_b$ 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 in the middle of $b,c$. Let $Q_b$ and $Q_c$ be maximal cliques with $b \in Q_b$ 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 maximal cliques $Q_b$ and $Q_c$ with $b \in Q_b$ 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 $x_0 - \cdots - x_r$, with $b = x_0$ and $c = x_r$ and none of the $x_i$’s is in $N(a)$. We may assume that this path is chordless. For $1 \leq i \leq r$, let $Q_i$ be a maximal clique containing $x_{i-1}, x_i$. Then $Q_1, \ldots, Q_r$ 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 $T[Q_1, \ldots, Q_r]$, but $Q_1$ contains $b$ and $Q_r$ contains $c$, a contradiction. $\blacksquare$

Now we are ready to prove the main theorem. Part of the proof was done in the previous section. Lemma 1 deals with the case where there exists a simplicial vertex that is in the middle of two other vertices; now we have to look at the case where all simplicial vertices are not in the middle of any pair of vertices.

Proof of Theorem 1. By Theorem 3, a path graph does not contain any of $F_0, \ldots, F_{16}$. Suppose now that there exists a graph $G$ that does not contain any of $F_0, \ldots, F_{16}$ and is a minimal non-path graph. Since $G$ contains no $F_0$, it is chordal. By Theorem 2, there is a special simplicial vertex $q$ of $G$. By Lemma 1, $q$ is co-special. Let us denote by $Q$ the unique maximal clique containing $q$. It will be convenient to denote by $S_Q$ the separator $S_q$.

The graph $G$ is not a clique as it is not a path graph. Consider graph $G(Q \setminus Q)$, which is equal to $G \backslash (Q \setminus S_q)$. Since $Q \setminus S_q \neq \emptyset$, and by the minimality of $G$, it follows that $G(Q \setminus Q)$ admits a clique path tree $T_0$. Let $Q'$ be a vertex of $T_0$ such that $S_q \subset Q'$ (by the fact that $S_Q$ is a separator $Q'$ does exist). Let $T'_0$ be obtained from $T_0$ by adding vertex $Q'$ and edge $QQ'$. Remark that $T'_0$ is a clique tree of $G$ but not a clique path tree since $G$ is not a path graph.

Claim 1. For all non-adjacent vertices $u,w \notin Q$, there exists a path between $u$ and $w$ that avoids the neighborhood of $q$.

Proof. Suppose the contrary. Let $U,W \in Q$ be such that $u \in U$ and $w \in W$. We have $U \neq W$ since $u,w$ are not adjacent. By Lemma 3, there is an edge of $T_0[U,W]$ whose label is included in $S_Q$, contradicting that $q$ is co-special. Thus the claim holds. $\blacksquare$

For each clique $L \in Q \setminus \{Q, Q'\}$ we will use the following notation. Let $L'$ be the neighbor of $L$ along $T_0[L, Q']$ and $S_L$ be the label $L \cap L'$ of the edge $LL'$. Let $T_L$ be the largest subtree of $T'_0$ that contains $Q'$ and in which no label is included in $S_L$. Let $S'_L$ be the label of the edge of $T_0[L, Q']$ that has exactly one extremity in $T_L$.
Since \( q \) is special and co-special we have \( S_Q \not\subseteq S_L \), so \( T_L \) contains \( Q \). Note that \( S'_L \subseteq S_L \) by the definition of \( T_L \).

Let \( L \) be the set of cliques \( L \subseteq Q \setminus \{Q, Q'\} \) such that \( LL' \) is the only edge incident to \( L \) whose label contains \( S'_L \). In particular, for a vertex \( x \in Q' \), any leaf of \( T_0' \) which is not equal to \( Q' \) is in \( L \). Recall that \( T_0' \) is a path because \( T_0' \subseteq \) is a clique path tree. Let \( A \) be the set of vertices \( a \) of \( Q \) such that \( Q' \) is a vertex of \( T_0' \) that is not a leaf. Then \( A \) is not empty, for otherwise \( T_0' \) would be a clique path tree of \( G \).

**Claim 2.** For each clique \( L \in \mathcal{L} \) we have \( L' \in T_L \).

**Proof.** Suppose on the contrary that \( L' \not\in T_L \). Let \( \bar{L} \) be the clique in \( T_0[L, Q'] \) such that \( \bar{L} \not\in T_L \) and \( \bar{L} \not\in T_L \). Then \( \bar{L} \not\in L \) and the edge \( LL' \) has label \( S'_L \) (possibly \( \bar{L} = L' \)).

When we remove the edges \( LL' \) and \( LL' \) from \( T_0' \), there remain three subtrees \( T_1, T_2, T_3 \), where \( T_1 \) is the subtree that contains \( L \), \( T_2 \) is the subtree that contains \( L' \) and \( \bar{L} \), and \( T_3 \) is the subtree that contains \( \bar{L}, Q', Q \). Let \( T_4 \) be the tree formed by \( T_1 \) and \( T_3 \) plus the edge \( \bar{L} L' \).

Then, since \( S'_L \subseteq S_L \), \( T_4 \) is a clique tree of \( G(Q(T_4)) \). Let \( x \) be any vertex in \( L' \setminus L \). Vertex \( x \) does not belong to any vertex of \( T_1 \) as it is not in \( L \). Since \( S'_L \subseteq L \), vertex \( x \) does not belong to any vertex of \( T_3 \). So \( G(Q(T_4)) \) is a strict subgraph of \( G \) and there exists a clique path tree \( T_5 \) of \( G(Q(T_4)) \). Label \( S'_L \) is on the edge \( \bar{L} L' \) of \( T_4 \), so it is also a label of \( T_5 \). Consequently, there is an edge \( LL'' \) of \( T_5 \) with a label \( R \) such that \( S'_L \subseteq R \subseteq L \). (Possibly \( LL'' = \bar{L} \).) Suppose that \( R \neq S'_L \). Then there is an edge of \( T_1 \) or \( T_3 \) with label \( R \). But no label of \( T_1 \) can be \( R \) by the definition of \( L \); and all the labels of \( T_3 \) that are included in \( L \) are also included in \( S'_L \), so no label of \( T_3 \) can be \( R \), a contradiction.

By the preceding claim, every \( L \in \mathcal{L} \) satisfies \( S'_L = S_L \).

Let \( L^* \) be the set of all \( L \in \mathcal{L} \) such that \( T_L \) is a strict subtree of \( T_0' \setminus L \).

**Claim 3.** For any \( a \in A \), at least one leaf of \( T_0' \) is in \( L^* \).

**Proof.** Let \( L_1, L_2 \) be the leaves of \( T_0' \), as already noted, both are in \( \mathcal{L} \). For \( i = 1, 2 \), let \( \ell_i \in L_i \setminus S_{L_i} \). The three vertices \( q, \ell_1, \ell_2 \) are adjacent to \( a \), so they do not form an asteroidal triple by Lemma 2, and so one of them is in the middle of the other two.

Vertex \( q \) cannot be in the middle of \( \ell_1, \ell_2 \) by Claim 1. So we may assume up to symmetry that \( \ell_1 \) is in the middle of \( q, \ell_2 \). So, by Lemma 3, 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 \( L_1 \in L^* \). Thus the claim holds.

The preceding claim implies that \( L^* \) is not empty. We choose \( L \in L^* \) such that the subtree \( T_L \) is maximal. Let \( S_Q \) be the label of the edge of \( T_0[L, Q'] \) incident to \( Q' \). Vertex \( q \) is special and co-special, so there exists \( s_Q \in S_Q \setminus s_Q \), and we have \( s_Q \not\in S_L \).

Therefore, no clique of \( Q \setminus Q(T_L) \) contains \( s_Q \). We add the edge \( LL' \) to \( T_L \) to obtain a clique tree \( T_L' \) of \( G(Q(T_L) \cup \{L\}) \). Since \( L \in L^* \), we have \( T_L' \not\in T_0' \), and by the minimality of \( G \), there exists a clique path tree \( T \) of \( G(Q(T_L')) \). Note that \( L \) is a leaf of \( T \), for
otherwise there are at least two labels of $T$ that are included in $S_L$, which contradicts the definition of $T_L$. From now on, our goal will be to show that either $G$ contains one of the forbidden subgraphs, or $T$ can be extended into a clique path tree of $G$.

**Claim 4.** Let $a \in A$ be such that both leaves of $T_0^a$ are not in $T_L$. Let $L_a$ be a leaf of $T_0^a$ that belongs to $L^*$. Then $L_a'$ is in $T_L$, and every edge $KK'$ of $T_0$ with $K \notin T_L$ and $K' \in T_L$ satisfies $S_K \subseteq S_{L_a}$.

**Proof.** By Claim 3, $L_a$ exists. Since the labels of the edges of $T_L$ are not included in $S_L$, they are also not included in $S_{L_a}$. So $T_L$ is a subtree of $T_{L_a}$. By the maximality of $T_L$, we have $T_L = T_{L_a}$. By Claim 2, $L_a'$ is in $T_L$. By the definition of $T_{L_a}$, every edge $KK'$ of $T_0$ with $K \notin T_L$ and $K' \in T_L$ satisfies $S_K \subseteq S_{L_a}$. Thus the claim holds.

**Claim 5.** There exist $U, W \in Q \setminus Q(T_L')$ such that $UL$ is an edge of $T_0$, $S_U \setminus Q' \neq \emptyset$, $U \cap W \neq \emptyset$, $W' \in Q(T_L)$ and $W \cap Q' \neq \emptyset$.

**Proof.** We define sets $U, V$ as follows:

$$U = \{ U \in Q \setminus Q(T_L') | UL \text{ is an edge of } T_0 \}$$

$$V = \{ V \in Q \setminus Q(T_L') | V' \in Q(T_L) \}.$$

We observe that the members of $V$ are pairwise disjoint. For if there is a vertex $v$ in $V_1 \cap V_2$ for some $V_1, V_2 \in V$, then $v$ is on three labels (namely $S_{V_1}, S_{V_2},$ and $S_L$) of $T_0$ that do not lie on a common path, contradicting that $T_0$ is a clique path tree.

We define sets $U_p (p \geq 1)$ and $V_p (p \geq 0)$ as follows:

$$V_0 = \{ W \in V | W \cap Q \neq \emptyset \}$$

$$U_p = \{ U \in U \setminus (U_1 \cup \cdots \cup U_{p-1}) | \exists V \in V_{p-1} \text{ such that } U \cap V \neq \emptyset \} \quad (p \geq 1)$$

$$V_p = \{ V \in V \setminus (V_0 \cup \cdots \cup V_{p-1}) | \exists U \in U_p \text{ such that } V \cap U \neq \emptyset \} \quad (p \geq 1).$$

Consider the smallest $k \geq 1$ such that there exists $U \in U_k$ with $S_U \setminus Q' \neq \emptyset$. If no such $U$ exists, then let $k = \infty$. Claim 5 to be proved states that $k = 1$, so let us suppose on the contrary that $k \geq 2$. For all $1 \leq p \leq k-1$ and all $U \in U_p$, we have $S_U \subseteq Q'$; for each such $U$ we denote by $U''$ the vertex of $Q(T)$ such that $U'' \cap S_U \neq \emptyset$ and the length of $T[U,L,U'']$ is maximum. Remark that $S_U$ is included in $U''$ if and only if all vertices of $T$ that intersect $S_U$ contain $S_U$. Let us prove that:

$$S_U \subseteq U'' \text{ for every } U \in U_p, \quad 1 \leq p \leq k-1. \quad (1)$$

Suppose that there exists $U_p \in U_p, \ 1 \leq p \leq k-1$, such that $S_{U_p} \not\subseteq U''$, and let $p$ be minimum with this property. Let $V_0, \ldots, V_{p-1}, U_1, \ldots, U_p$ be such that $V_i \in V_i, U_i \in U_i, V_{i-1} \cap U_i \neq \emptyset,$ and $U_i \cap V_i \neq \emptyset$. We claim that $V_0 = V_1' = \cdots = V_{p-1}'$. For otherwise there exists $i \in \{1, \ldots, p-1\}$ such that $V_{i-1}' \neq V_i'$. Then $V_i'$ contains elements of $S_{U_i}$ but not all, and so $S_{U_i} \not\subseteq U''_{i'}$, which contradicts the minimality of $p$. Pick $u_i \in U_i \setminus S_{U_i}$ and $v_i \in V_i \setminus V_i$. Let $x_1, \ldots, x_r$ be such that $x_1 \in V_0 \cap U_1, x_2 \in U_1 \cap V_1, \ldots, x_r \in V_{p-1} \cap U_p$ with $r = 2p-1$. 

*Journal of Graph Theory* DOI 10.1002/jgt
By the definition of the $V_i$'s, none of $x_2, \ldots, x_r$ is in $Q$. Let $x_0 \in V_0 \cap Q$ (maybe $x_0 = x_1$). So $x_0 \in S_{V_0} \subseteq S_L \subseteq L$. None of $U_2, \ldots, U_p$ can contain $x_0$ by the definition of $U_i$. Note that $x_r$ is in $U_p$ and $V_{p-1}' = V_0'$; on the other hand we have $S_{U_p} \not\subseteq U_p'$. So there exists a clique $Z$ of $T_L$ such that $Z' \in T_0'$, $S_{U_p} \subseteq Z'$, $S_{U_p} \cap Z \neq \emptyset$ and $S_{U_p} \cap Z' \neq \emptyset$. Vertex $Q'$ is on $T_0'[L, Z']$ as $S_{U_p} \subseteq Q'$. Let $z \in Z \cap Z'$. We can find vertices $y_1, \ldots, y_r$ on the labels of $T_0'[Z, Q]$ such that none of them is in $S_L$ and $z - y_1 - \cdots - y_r - q$ is a chordless path in $G$. Let $\ell \in L \setminus S_L$. By Claim 1, there exists a path $P$ between $z$ and $\ell$ whose vertices are not neighbors of $q$.

If $Z \in T_0'$, 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 $z, \ell, q$ form an asteroidal triple (because of the three paths $P$, $z - y_1 - \cdots - y_r - q$, and $\ell - b - c - q$), and they lie in the neighborhood of $x_0$, a contradiction to Lemma 2. So $Z \not\in T_0'$. Let $x_{r+1} \in Z \cap U_p$. If $x_{r+1} \in Q$, then $z, \ell, q$ form an asteroidal triple (because of the three paths $P$, $z - y_1 - \cdots - y_r - q$, and $\ell - x_0 - q$), and they lie in the neighborhood of $x_{r+1}$, a contradiction again. So $x_{r+1} \not\in Q$. 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. Vertex $\ell$ is not in any of the $S_{U_i}$'s, and $\ell$ is adjacent to all of $x_0, \ldots, x_{r+1}$ and to none of $u_1, \ldots, u_p, v_0, \ldots, v_{p-1}, y_1, \ldots, y_r, z, q$.

Suppose that $V_0 \cap U_1 \cap Q \neq \emptyset$. Then we may assume that $x_0 = x_1$, so $x_0$ is in $A$ and the two leaves of $T_0'$ are not in $T_L$. By Claims 3 and 4, there exists a leaf $L_{x_0}$ of $T_0'$ that belongs to $L^*$ and $L_{x_0}'$ is in $T_L$, so $L_{x_0} = V_0$. But $x_{r+1}$ is in $Z \cap U_p$, so it is not in $S_{V_0}$; thus $S_{U_p} \not\subseteq S_{V_0}$, which contradicts the end of Claim 4. Therefore $V_0 \cap U_1 \cap Q = \emptyset$, so $x_0 \neq x_1, x_0 \not\in U_1, x_1 \notin Q$. Now, if $t = 1$, then $\{u_1, \ldots, u_p, v_0, \ldots, v_{p-1}, x_0, \ldots, x_{r+1}, y_1, q, z, \ell\}$ induces $F_{14}(4p+5)_{p\geq1}$. If $t = 2$, then $\{u_1, \ldots, u_p, v_0, \ldots, v_{p-1}, x_0, \ldots, x_{r+1}, y_1, y_2, q, z, \ell\}$ induces $F_{15}(4p+6)_{p\geq1}$. If $t \geq 3$, then $\{x_0, x_{r+1}, z, y_1, \ldots, y_t, q\}$ induces $F_{10}(s+5)_{s\geq1}$, a contradiction. Therefore (1) holds.

Suppose that $k$ is finite. Let $V_0, \ldots, V_{k-1}, U_1, \ldots, U_k$ be such that $V_i \in V_i$, $U_i \subseteq U_i$, $V_i \cup U_i \neq \emptyset$, and $U_j \cap V_i \neq \emptyset$. Let $u_i \in U_i \setminus S_{V_i}$ and $v_i \in V_i \setminus S_{V_i}$. Pick vertices $x_1 \in V_0 \cap U_1$, $x_2 \in U_1 \cap V_1, \ldots, x_r \in V_{k-1} \cap U_k$ with $r = 2k - 1$. By the definition of the $V_i$'s, none of $x_2, \ldots, x_r$ is in $Q$. Let $x_0 \in V_0 \cap Q$. Suppose that $V_0 \cap U_1 \cap Q \neq \emptyset$. Then we can assume that $x_0 = x_1$, so $x_0$ is in $A$ and the two leaves of $T_0'$ are not in $T_L$. By Claims 3 and 4, a leaf $L_{x_0}$ of $T_0'$ is in $L^*$ and $L_{x_0}'$ is in $T_L$, so $L_{x_0} = V_0$. But $x_2$ is in $S_{V_1}$ and not in $S_{V_0}$, so $S_{V_1} \not\subseteq S_{V_0}$, which contradicts the end of Claim 4. Therefore $V_0 \cap U_1 \cap Q = \emptyset$, and $x_0 \neq x_1, x_0 \not\in U_1, x_1 \notin Q$. Let $s_{U_k} \in S_{U_k} \setminus Q'$. Vertex $s_{U_k}$ is not adjacent to any of $q, s_Q, v_0, \ldots, v_{k-1}$ because $s_{U_k} \notin Q'$, and by the minimality of $k$, vertex $s_{U_k}$ is not adjacent to $u_1, \ldots, u_{k-1}$. Then $\{u_1, \ldots, u_k, v_0, \ldots, v_{k-1}, x_0, \ldots, x_r, s_{U_k}, q, q\}$ induces $F_{16}(4k+3)_{k\geq2}$, a contradiction.

Now $k$ is infinite. Then the members of $\bigcup_{p \geq 1} U_p$ are included in $Q'$ and pairwise disjoint, for otherwise $T_0$ is not a clique path tree. For each member $M$ of $U \cup V$, let $T_0'(M)$ be the component of $T_0' \setminus T_L$ that contains $M$. Starting from the clique path tree $T$ and the trees $T_0'(M)$ ($M \in U \cup V$), we build a new tree as follows. For each $V \in \bigcup_{p \geq 1} V_p$, we add the edge $V L$ between $T_0'(V)$ and $T$. For each $U \in \bigcup_{p \geq 1} U_p$, we add the edge $U U'$ between $T_0'(U)$ and $T$. For each $U \in U \setminus \bigcup_{p \geq 1} U_p$, we add the edge $U L$ between $T_0'(U)$ and $T$. For each $V \in V \setminus \bigcup_{p \geq 1} V_p$, we define $V'' \in Q(T)$ such that $V'' \cap S_V \neq \emptyset$ and the length of $T[L, V'']$ is maximum. By the definition of $V_0$, we have $S_V \cap Q = \emptyset$. 

Journal of Graph Theory DOI 10.1002/jgt
so $V'' \neq Q$, so $V''$ is a vertex of $T_L$ on $T_0[L, V]$ and it contains $S_V$ as $S_V \subseteq S_L$. Then we can add the edge $VV''$ between $T_0'(V)$ and $T$. Thus we obtain a clique path tree of $G$, a contradiction. So $k = 1$, and there exist $U \in \mathcal{U}_1$ and $W \in \mathcal{V}_0$ such that $S_U \setminus Q' \neq \emptyset$, $U \cap W \neq \emptyset$, and $W \cap Q \neq \emptyset$. Thus the claim holds.

Let $U, W$ be as in the preceding claim. Let $s_U \in S_U \setminus Q'$. Vertex $s_U$ is not adjacent to $s_Q$. Let $u \in U \setminus S_U$ and $w \in W \setminus S_W$. We have $W \in Q(T_L)$, so $S_W \subseteq S_L$. Moreover $W \cap Q \neq \emptyset$, so $W \cap Q' \cap L \neq \emptyset$, so $Q'$ is on $T_0[W, L]$ as $T_0$ is a clique path tree.

**Claim 6.** $S_W = S_L$.

**Proof.** Assume on the contrary that $S_W \neq S_L$. Then $S_W$ is a proper subset of $S_L$. Suppose that there exists $a \in U \cap W \cap Q \neq \emptyset$. Then $a$ is in $A$ and the two leaves of $T_0'$ are not in $T_L$. By Claims 3 and 4, a leaf $L_a$ of $T_0'$ in in $L^*$ and $L'_a$ is in $T_L$, so $L_a = W$. But $S_L \not\subseteq S_W$, so Claim 4 is contradicted. Therefore $U \cap W \cap Q = \emptyset$. By the definition of $U$ and $W$, there exists $b \in W \cap Q$ and $c \in U \cap W$. So $b \in U, c \in Q, b \neq c$. Since $s_U$ is in $S_U \setminus Q'$, we have $S_U \not\subseteq S_W$. The labels of the edges of $T_L$ are not included in $S_L$, so they are also not in $S_W$. Thus, we can choose vertices $x_1, \ldots, x_r$ 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 $u-x_1-\cdots-x_r-q$ is a path from $u$ to $q$ that avoids $N(w)$. Suppose $r = 1$. Then $x_1$ is different from $s_U$ and $s_Q$, and $\{w, b, c, u, s_U, x_1, s_Q, q\}$ induces $F_8$. Suppose $r = 2$. If $x_1$ is adjacent to $s_Q$, then $\{w, b, c, u, s_U, x_1, s_Q, q\}$ induces $F_9$, and if $x_1$ is not adjacent to $s_Q$, then $\{w, b, c, u, x_1, s_Q, q\}$ induces $F_9$. Finally, suppose $r \geq 3$. Then $\{w, b, c, u, x_1, \ldots, x_r, q\}$ induces $F_{10}(r+5)_{r \geq 3}$. In all cases we obtain a contradiction. Thus the claim holds.

**Claim 7.** $W \in \mathcal{L}^*$. 

**Proof.** In the connected component of $T_0'[W']$ that contains $W$, let $X \in Q$ be such that $S_W \subseteq X$ and the length of $T_0'[X, W]$ is maximum (possibly $X = W$). Then $S_W \subseteq S_X$ and $XX'$ is the only edge of $T_0'$ incident to $X$ that contains $S_W$, so $X \in \mathcal{L}$. Since $S_W \subseteq S_X$ we have that $W \not\in T_X$. Then, by Claim 2 we have $X = W$ and by Claim 6 we have $T_W = T_L$; so $W \in \mathcal{L}^*$. Thus the claim holds.

By Claim 7, we have $W \in \mathcal{L}^*$. By Claim 6, we have $T_W = T_L$, so $T_W$ is also maximal and what we have proved for $L$ can be done for $W$. Thus, by Claim 5, there exists $X \not\in T_W$ such that $XW$ is an edge of $T_0$ with $S_X \setminus Q' \neq \emptyset$ and $X \cap S_W \neq \emptyset$. Let $x \in X \setminus W$ and $s_X \in S_X \setminus Q'$. Vertex $s_X$ is not in $S_W$, for otherwise, it would also be in $S_L$, and in $Q'$. Vertex $s_U$ is not in $S_L$, for otherwise, it would also be in $S_W$ and in $Q'$. Vertex $s_Q$ is not in $S_W$ (=$S_L$). So $s_Q, s_X, s_U$ are pairwise non-adjacent.

Suppose that there exists a vertex $a \in \partial X \cap Q \neq \emptyset$. So $a \in A$, but none of the two leaves of $T_0'$ can satisfy Claim 4, a contradiction. Therefore $\partial U \cap \partial X \cap Q = \emptyset$.

Suppose that $U \cap X \neq \emptyset$, and let $a \in U \cap X$. So $a$ is not in $Q$. Let $b \in S_W \cap Q$ (= $S_L \cap Q$). So $b$ is not in $U \cap X$. If $b \not\in X \cup U$, then $\{q, u, x, s_Q, s_U, s_X, a, b\}$ induces $F_6$, a contradiction. So $b$ is in one of $U, X$, say $b \in X \setminus U$ (if $b$ is in $U \setminus X$ the argument is similar). Since $W$ is in $L$, there is a vertex $c \in S_W \setminus S_X$. Vertex $c$ is adjacent to $a, b, s_U, s_Q$ and not to $x$. Then $\{x, a, b, u, s_U, c, s_Q, q\}$ induces $F_8$, $F_9$, or $F_{10}(8)$, a contradiction. Therefore $U \cap X = \emptyset$. 

*Journal of Graph Theory* DOI 10.1002/jgt
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 \( X \cap Q = \emptyset \). Let \( c \in W \cap Q \). Then \( c \in L \) and \( c \notin X \). Let \( d \in X \cap S_W \); so \( d \in L \), \( d \notin Q \), \( d \notin U \). If \( c \) is adjacent to \( u \), then \( \{q, u, x, s_Q, s_U, s_X, c, d\} \) induces \( F_{6} \), else \( \{q, u, x, s_Q, s_U, s_X, a, c, d\} \) induces \( F_{7} \), a contradiction. So \( a \in Q \). Let \( e \in X \cap S_W \); so \( e \in L \). If \( e \notin Q \), then \( \{q, u, x, s_Q, s_U, s_X, a, e\} \) induces \( F_6 \), a contradiction. So \( e \in Q \). Let \( f \in S_W \setminus S_Q \) (\( f \) exists because \( q \) is special and co-special). Since \( U \cap X = \emptyset \), \( f \) is adjacent to at most one of \( u, x \), and then \( \{q, u, x, s_U, s_X, a, e, f\} \) induces \( F_9 \) or \( F_{10} \), a contradiction. This completes the proof of Theorem 1.

6. RECOGNITION ALGORITHM

Our proof above yields a new recognition algorithm for path graphs, which takes any graph \( G \) as input and either builds a clique path tree for \( G \) or finds one of \( F_0 \), ..., \( F_{16} \) as an induced subgraph of \( G \). We have not analyzed the exact complexity of such a method but it is easy to see that it is polynomial in the size of the input graph. More efficient algorithms were already given by Gavril [8], Schäffer [19], and Chaplick [4], with complexity, respectively, \( O(n^3) \), \( O(nm) \), and \( O(nm) \) for graphs with \( n \) vertices and \( m \) edges. Another algorithm was proposed in [5] and claimed to run in \( O(n+m) \) time, but it has only appeared as an extended abstract (see comments in [4, Section 2.1.4]).

There are classical linear time recognition algorithms for triangulated graphs [18], and, following [2], there have been several linear time recognition algorithms for interval graphs, of which the most recent is [10]. We hope that the work presented here will be helpful in the search for a linear time recognition algorithm for path graphs.

ACKNOWLEDGMENTS

The authors thank the anonymous referees for their useful remarks. We are especially grateful to a referee whose suggestions lead to a significant simplification of the proof of Theorem 2.

REFERENCES


