To appear in *Information Processing Letters*

Subtree filament graphs are subtree overlap graphs

Jessica Enright and Lorna Stewart*

Department of Computing Science, University of Alberta†

July 4, 2007

Abstract

We show that the class of intersection graphs of subtree filaments in a tree is identical to the class of overlap graphs of subtrees in a tree.

Key Words: overlap graph; subtree filament graph; subtree overlap graph; combinatorial problems; graph algorithms

1 Introduction

Subtree filament graphs, interval filament graphs, and other types of filament graphs were introduced by Gavril in [5]. Independently, Čenek and Stewart studied subtree overlap graphs [2]. Since interval filament graphs contain circle graphs and several other graph classes, it seems reasonable to expect that subtree overlap graphs would form a proper subset of subtree filament graphs. However, in this paper, we show that the classes of subtree filament graphs and subtree overlap graphs are identical. This result first appeared in Enright's MSc thesis [3].

We consider finite, simple graphs. Two sets A and B are said to *intersect* if $A \cap B \neq \emptyset$, and to *overlap* if $A \cap B \neq \emptyset$, $A \not\subseteq B$, and $B \not\subseteq A$. For nonempty sets $A, B, C,$ and D, if A overlaps B and C overlaps D, or $A \cap B = \emptyset$ and $C \cap D = \emptyset$, or ($A \subseteq B$ or $B \subseteq A$) and ($C \subseteq D$ or $D \subseteq C$) then we say that sets A, B and sets C, D are *similarly related* and we write $A, B \sim C, D$.

Let $\mathcal{S} = \{S_1, S_2, \ldots, S_n\}$ be a multiset of sets. The *intersection graph* (respectively *overlap graph*) of $\mathcal{S}$ is the graph $G = (V, E)$ where $V = \{v_1, v_2 \ldots, v_n\}$ and, for all $1 \leq i, j \leq n$, $v_i v_j \in E$ if and only if S_i and S_j intersect (respectively overlap). The *containment graph* of $\mathcal{S}$ is the graph $G = (V, E)$ where $V = \{v_1, v_2 \ldots, v_n\}$ and, for all $1 \leq i, j \leq n$, $v_i v_j \in E$ if and

*stewart@cs.ualberta.ca

†Edmonton, Alberta, Canada T6G 2E8

only if $S_i \subseteq S_j$ or $S_j \subseteq S_i$. If G is the intersection, overlap, or contaiment graph of $\mathcal{S}$ then $\mathcal{S}$ is called an intersection, overlap, or containment representation for G. Every graph has both an intersection representation [12] and an overlap representation (obtained by adding a unique new element to each set of an intersection representation, as observed in [2]). We assume without loss of generality that all sets in the representations that we consider are nonempty. Note that, for $\mathcal{S} = \{S_1, \ldots, S_n\}$ and $\mathcal{S}' = \{S'_1, \ldots, S'_n\}$, where $S_i, S_j \sim S'_i, S'_j$ for all $1 \leq i, j \leq n$, the intersection (respectively overlap, containment) graphs of $\mathcal{S}$ and $\mathcal{S}'$ are isomorphic.

Interval graphs are the intersection graphs of intervals on a line and circle graphs are the overlap graphs of intervals on a line. Chordal graphs are graphs in which every cycle of length greater than three has a chord or, equivalently, the intersection graphs of subtrees in a tree. Comparability graphs are graphs whose edges can be transitively oriented. Equivalently, comparability graphs are the containment graphs of subtrees in a tree, the containment graphs of substars in a star, and the set of all containment graphs [9]. For more information on these and other graph classes, the reader is referred to [1] and [7].

Gavril [5] defines interval filament graphs as follows. Let $\mathcal{I} = \{I_1, I_2, \ldots, I_n\}$ be a set of intervals on a line L and let P be a plane containing L. We construct a set $\mathcal{F} = \{f_1, f_2, \ldots, f_n\}$ of filaments on the intervals of $\mathcal{I}$ where, for each $1 \leq i \leq n$, f_i is a curve in P on and above L, connecting the endpoints of I_i such that if two intervals are disjoint their curves do not intersect. Pairs of filaments corresponding to overlapping intervals will necessarily intersect. For a given pair of intervals, one of which is contained in the other, the corresponding filaments may or may not intersect. Now, *interval filament graphs* are the intersection graphs of interval filaments.

The intersection graphs of other types of filaments are also defined in [5]. In particular, we are concerned with subtree filament graphs. Consider a tree T in a plane P and let $\mathcal{T} = \{t_1, t_2, \ldots, t_n\}$ be a set of subtrees of T. Subtree filaments on the elements of $\mathcal{T}$ are constructed in a surface perpendicular to P whose intersection with P is exactly T. In this surface, in and above T, we construct filaments $\mathcal{F} = \{f_1, f_2, \ldots, f_n\}$ where each f_i, $1 \leq i \leq n$, is a curve connecting the leaves of t_i such that (i) filaments corresponding to disjoint subtrees do not intersect, and (ii) filaments corresponding to overlapping subtrees intersect. Filaments corresponding to subtrees, one of which is contained in the other, may or may not intersect. The leaves of t_i are referred to as the *leaves* or the *endpoints* of f_i. Figure 1 depicts a set of filaments on subtrees of a tree, and the corresponding subtree filament graph. It is necessary to explicitly require (ii) for subtree filaments, whereas this requirement is always satisfied by interval filaments. To see this, consider two subtrees that overlap in exactly one vertex of T such that the common vertex is an internal node in both subtrees.

All of the graph classes defined in this section are *hereditary*, that is, every induced subgraph of a graph in the class is also in the class.

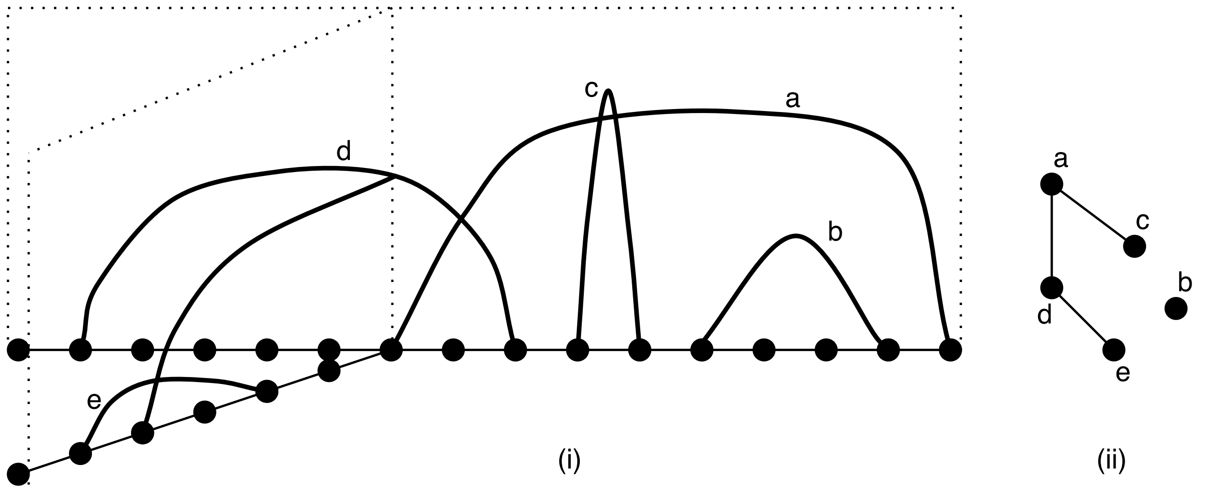

Figure 1: (i) A set of filaments on subtrees of a tree and (ii) the corresponding subtree filament graph. Labels a, b, c, d, and e indicate the correspondence between filaments and vertices.

2 Equivalence of subtree filament graphs and subtree overlap graphs

In this section, we present a proof of the equivalence of subtree filament graphs and subtree overlap graphs. Since the intersection graphs of interval filaments is a much larger graph class than that of interval overlap graphs (i.e. circle graphs), it is surprising that the graph family defined by subtree filament intersection is exactly the class of subtree overlap graphs.

We first identify some operations on trees and subtrees that preserve relationships (disjointness, overlapping, and containment) among subtrees. We start with a technical lemma which shows that relationships among subtrees are preserved by the addition of a new leaf to a given subtree provided the leaf is also added to all subtrees that contain the given subtree.

Lemma 1 *Let $T = (V_T, E_T)$ be a tree and $\mathcal{T} = \{t_1 = (V_{t_1}, E_{t_1}), \ldots, t_n = (V_{t_n}, E_{t_n})\}$ be a multiset of subtrees of T. Let $x \notin V_T$. For any given $1 \le k \le n$, and any given $v \in t_k$, let $T' = (V_T \cup \{x\}, E_T \cup \{vx\})$ and $\mathcal{T}' = \{t'_1, \ldots, t'_n\}$ where, for all $1 \le i \le n$, t'_i is defined as*

$$t'_i = \begin{cases} (V_{t_i} \cup \{x\}, E_{t_i} \cup \{vx\}) & \text{if } t_k \subseteq t_i \\ t_i & \text{otherwise} \end{cases}$$

Then, for all $1 \le i, j \le n$, $t_i, t_j \sim t'_i, t'_j$.

Proof. It is clear that T' is a tree and $\mathcal{T}'$ is a multiset of subtrees of T'. Note that the relationship between two nonempty sets A and B (i.e. whether they are equal, disjoint, overlap, or one is properly contained in the other) is completely determined by whether each of the following is empty: $A \cap B$, $A \backslash B$, and $B \backslash A$. We consider three cases.

- If $v \in t_i \cap t_j$ then $t_i \cap t_j \neq \emptyset$, $t'_i \cap t'_j \neq \emptyset$, $t'_i \backslash t'_j = t_i \backslash t_j$, and $t'_j \backslash t'_i = t_j \backslash t_i$.
- If $v \in t_i \backslash t_j$ then $t_i \cap t_j = t'_i \cap t'_j$, $t_i \backslash t_j \neq \emptyset$, $t'_i \backslash t'_j \neq \emptyset$, and $t'_j \backslash t'_i = t_j \backslash t_i$.
- If $v \notin t_i \cup t_j$ then $t'_i = t_i$ and $t'_j = t_j$.

In each case, the conclusion that $t_i, t_j \sim t'_i, t'_j$ follows. □

The next lemma shows that, given a tree and a set of subtrees, the operation of adding a new leaf to an arbitrary vertex v of the tree and to exactly the subtrees that contain v, preserves the disjointness, overlapping, and containment relationships among the subtrees. Repeated application of the lemma results in a representation in which every subtree contains an edge and every two intersecting subtrees share an edge.

Lemma 2 *Let $T = (V_T, E_T)$ be a tree and $\mathcal{T} = \{t_1 = (V_{t_1}, E_{t_1}), \ldots, t_n = (V_{t_n}, E_{t_n})\}$ be a multiset of subtrees of T. Let $x \notin V_T$. For any given $v \in V_T$, let $T' = (V_T \cup \{x\}, E_T \cup \{vx\})$ and $\mathcal{T}' = \{t'_1, \ldots, t'_n\}$ where, for all $1 \leq i \leq n$, t'_i is defined as*

$$t'_i = \begin{cases} (V_{t_i} \cup \{x\}, E_{t_i} \cup \{vx\}) & \text{if } v \in V_{t_i} \\ t_i & \text{otherwise} \end{cases}$$

Then, for all $1 \leq i, j \leq n$, $t_i, t_j \sim t'_i, t'_j$.

Proof. Add a new subtree $t_{n+1} = \{v\}$ to $\mathcal{T}$, apply Lemma 1 with $t_k = t_{n+1}$, and then remove t'_{n+1} from $\mathcal{T}'$. □

It has often been pointed out in the literature that, for any multiset of intervals on a line, there is a set of intervals exhibiting the same intersection, overlapping, and containment relationships, in which no two intervals share an endpoint. The next lemma shows that a similar observation holds for subtrees of a tree. Specifically, the proof of the lemma shows how to transform an arbitrary multiset of subtrees of a tree into a set of subtrees which preserves the disjointness, overlapping, and containment relationships among the subtrees and in which no vertex of the tree is a leaf in two distinct subtrees.

Lemma 3 *Let $T = (V_T, E_T)$ be a tree and $\mathcal{T} = \{t_1, \ldots, t_n\}$ be a multiset of subtrees of T. There exists a tree $T' = (V_{T'}, E_{T'})$ and set $\mathcal{T}' = \{t'_1, \ldots, t'_n\}$ of subtrees of T' such that, for all $1 \leq i, j \leq n$, $t_i, t_j \sim t'_i, t'_j$ and no element of $V_{T'}$ is a leaf of two distinct elements of $\mathcal{T}'$.*

Proof. We assume without loss of generality that every subtree of $\mathcal{T}$ contains at least two vertices. If this were not the case, we could apply the transformation of Lemma 1 to each single vertex subtree of $\mathcal{T}$.

First, for each vertex $v \in V_T$, for each subtree t_i that contains v as a leaf, add a new leaf ℓ_i adjacent to v in T, and add ℓ_i to t_i and to all subtrees that contain t_i or are equal to t_i. Let L be the entire set of new leaves added after processing all vertices of V_T, and let T' and $\mathcal{T}' = \{t'_1, \ldots, t'_n\}$ be the resulting tree and multiset of subtrees. Now, all leaves of subtrees in $\mathcal{T}'$ are leaves of T' and, by repeated applications of Lemma 1, for $1 \leq i, j \leq n$, $t_i, t_j \sim t'_i, t'_j$.

Now, for each leaf ℓ of T' that is contained in two or more subtrees of $\mathcal{T}'$, let $t'_{\ell,1}, t'_{\ell,2}, \ldots, t'_{\ell,k}$ be the subtrees, in nondecreasing order of size, of $\mathcal{T}'$ that contain ℓ. Attach a path of $k-1$ new vertices: $p_1, \ldots, p_{k-1}$ to ℓ by adding the edge ℓp_1 to T'. Now, add vertices $p_1, \ldots, p_{i-1}$ and the edges of the $\ell, \ldots, p_{i-1}$ path to $t'_{\ell,i}$, for $2 \leq i \leq k$. This step finally ensures that no two subtrees share a common leaf. Disjointness is not affected by this operation; proper containment is guaranteed to be preserved because of the order in which the $t'_{\ell,i}$'s are handled, and overlapping cannot be destroyed by the addition of new vertices. Any two subtrees that were originally equal are now such that one is properly contained in the other. □

We make use of Lemmas 2 and 3 in the proof of Theorem 1.

Theorem 1 *A graph is a subtree overlap graph if and only if it is a subtree filament graph.*

Proof. We first prove that every subtree overlap graph is a subtree filament graph. Given a set of subtrees $\mathcal{T} = \{t_1 = (V_{t_1}, E_{t_1}), \ldots, t_n = (V_{t_n}, E_{t_n})\}$ of a tree T, we construct a set of filaments on T whose intersection graph is the overlap graph of $\mathcal{T}$. We may assume that every subtree contains at least two vertices and that every two subtrees are either disjoint or share an edge, since this could be achieved by applying the transformation of Lemma 2 to each vertex of T. In addition, by Lemma 3 we may assume that no two subtrees share an endpoint, which implies that no two subtrees are equal.

Let the elements of $\mathcal{T}$ be indexed such that $i < j$ implies $|V_{t_i}| \leq |V_{t_j}|$. We first construct a set $\mathcal{F} = \{f_1, \ldots, f_n\}$ of filaments, for i from 1 to n, as follows: every point of f_i is above a point of t_i and f_i is entirely above every f_j, $j < i$ for which $t_j \subset t_i$. In addition, we choose f_i to be a function, i.e. such that no point of f_i is above any other point of f_i.

Now, $\mathcal{F}$ contains non-intersecting filaments for all pairs of subtrees that are disjoint and for all pairs in which one contains the other. However, there may be overlapping subtrees with corresponding filaments that do not intersect. Thus, to complete the construction, we may need to change some elements of $\mathcal{F}$ to ensure that overlapping subtrees correspond to intersecting filaments.

For every $1 \leq i, j, \leq n$ such that t_i overlaps t_j but f_i does not intersect f_j, we alter filaments as follows. Assume without loss of generality that f_i is entirely below f_j. As previously mentioned, there is an edge uv that is in both t_i and t_j. Let p be a point on f_i that is directly above the edge uv but not directly above either u or v. Thus, p is not directly above an endpoint of any subtree. We now draw p upwards so that it intersects f_j. Imagine that, as p is drawn upwards, the neighbourhood of p on f_i is also drawn upwards, so that f_i is stretched in such a way that it remains a curve and none of its points is above any other point of f_i. For each filament f_k that we encounter when drawing p up, if t_i overlaps t_k, we draw p across f_k so that f_i intersects f_k; otherwise ($t_i \subset t_k$), we draw the point q of f_k that is directly above p, along with q's neighbourhood on f_k, up such that f_i and f_k remain non-intersecting. Note that f_k will intersect f_j in this case. We say that f_k was *pushed up.*

Now, disjoint subtrees are still represented by non-intersecting filaments and overlapping subtrees are guaranteed to have intersecting filaments. Before any of the filaments were drawn up, pairs of subtrees exhibiting containment were represented by non-intersecting filaments. We must show that this is still the case.

Suppose that, during the drawing upwards of a point of f_i to cross f_j, an intersection is created between two filaments f_k and f_ℓ, where $t_k \subset t_\ell$. We know that $f_k \neq f_i$, since the processing of the f_i, f_j pair results in correct intersections between f_i and f_j and between f_i and all filaments encountered between f_i and f_j. Further, since f_k is pushed up by f_i, it must be that $t_i \subset t_k$ and therefore $t_i \subset t_\ell$. But, since f_i pushed f_k upwards to intersect f_ℓ, t_i must overlap t_ℓ, a contradiction. Thus, the intersection graph of the modified filaments $\mathcal{F}$ is the overlap graph of $\mathcal{T}$.

To complete the proof of the theorem, we now show that every subtree filament graph is a subtree overlap graph. Given filaments $\mathcal{F} = \{f_1, \ldots, f_n\}$ on tree $T = (V_T, E_T)$, we construct a tree T' and a set $\mathcal{T}' = \{t'_1, \ldots, t'_n\}$ of subtrees of T' such that, for all $1 \leq i, j \leq n$, t'_i and t'_j overlap if and only if f_i and f_j intersect.

The *subtree of T induced by the endpoints of filament f_i*, denoted by T_{f_i}, is defined to be the subtree of T induced by $X_i \cup \{x | x \in V_T$ is on a path in T between two vertices of $X_i\}$ where X_i is the set of endpoints of f_i.

First, we define T' to be the tree T with n additional nodes $X = \{x_1, x_2, \ldots, x_n\}$ attached as follows: for $1 \leq i \leq n$, x_i is adjacent in T' to a node of V_T that, in T, is an endpoint of f_i. Next, we define the subtrees of $\mathcal{T}'$. For $1 \leq i \leq n$, t'_i is the subtree of T' induced by the vertices of T_{f_i} plus $\{x_i\} \cup \{x_j \mid T_{f_j} \subset T_{f_i}$ and f_i does not intersect $f_j\}$.

It remains to show that the overlap graph of $\mathcal{T}'$ is identical to the intersection graph of $\mathcal{F}$. Suppose $f_i \cap f_j \neq \emptyset$. Then $T_{f_i} \cap T_{f_j} \neq \emptyset$ and therefore $t'_i \cap t'_j \neq \emptyset$. By the construction of the subtrees, since f_i intersects f_j, $x_i \in t'_i \backslash t'_j$ and $x_j \in t'_j \backslash t'_i$ and, therefore, t'_i overlaps t'_j.

Suppose $f_i \cap f_j = \emptyset$. If $T_{f_i} \cap T_{f_j} = \emptyset$ then $t'_i \cap t'_j = \emptyset$ as required. If one of T_{f_i} and T_{f_j} is contained in the other then assume, without loss of generality, that $T_{f_j} \subset T_{f_i}$. Then, by the construction of t'_i, x_j and the vertices of T_{f_j} are in t'_i. The vertices of t'_j, other than x_j and the vertices of T_{f_j}, are $\{x_k \mid T_{f_k} \subset T_{f_j}$ and f_k does not intersect $f_j\}$. However, all such vertices x_k are also in t'_i since $T_{f_k} \subset T_{f_j} \subset T_{f_i}$ and f_k cannot intersect f_i (otherwise, one of f_k or f_i would intersect f_j, a contradiction). Therefore, $t'_j \subset t'_i$ as required. □

Theorem 1 may be helpful in the search for an efficient recognition algorithm, and in the development of exact and approximation algorithms for various optimization problems, for subtree filament graphs. It is known that the following problems are polynomially solvable for subtree overlap graphs, given a subtree overlap representation: maximum weight clique and maximum weight independent set [2] [5], and finding holes and antiholes of a given parity, and minimum dominating holes [6]. The problems of vertex colouring [4], clique cover [11], and dominating set [10], among others, remain NP-complete when restricted to circle graphs and therefore to subtree overlap graphs. Approximation algorithms for the following problems on subtree filament graphs are given in [11], where it is assumed that a subtree filament representation is given: minimum vertex colouring, maximum k-colourable subgraph, minimum clique cover, and maximum h-coverable subgraph.

Edge intersection graphs of subtrees in a tree were defined in [8]; we give an analogous definition of edge overlap graphs of subtrees in a tree. The *edge intersection* (respectively *overlap*) graph of subtrees $\mathcal{T} = \{t_1, t_2, \ldots, t_n\}$ in a tree T is the graph $G = (V, E)$ where $v = \{v_1, \ldots, v_n\}$ and, for all $1 \leq i, j \leq n$, $v_i v_j \in E$ if and only if t_i and t_j have an edge of T in common, i.e., if and only if t_i and t_j share at least two vertices of T (respectively, T_i and T_j overlap and share at least two vertices of T).

By Lemma 2, we can transform an arbitrary set of subtrees into an equivalent set (i.e. one which preserves the disjointness, overlapping, and containment relationships among subtrees) such that every two subtrees are either disjoint or share an edge, by appending a new leaf to each node of the tree and expanding all subtrees that contain the original node to also contain the new leaf. Thus, subtree overlap graphs are edge overlap graphs of subtrees in a tree. This raises the question of whether all edge overlap graphs of subtrees in a tree are subtree overlap graphs. We answer this question in the negative by noting that not all graphs are subtree overlap graphs [2], and showing that every graph is an edge overlap graph of subtrees in a tree.

Golumbic and Jamison [8] showed that every graph is the edge intersection graph of subtrees in a tree. It turns out that a small alteration to the construction of [8] shows that every graph is the edge overlap graph of subtrees in a tree. Let $G = (V, E)$ be an arbitrary graph. Define $T = (V_T, E_T)$ where $V_T = V \cup E \cup \{x\}$ and $E_T = \{xv \mid v \in V\} \cup \{xe \mid e \in E\}$. Now, for each $v \in V$, let t_v be the subtree of T induced by $\{v\} \cup \{e \mid e$ is incident to v in $G\} \cup \{x\}$. Now, all subtrees contain x, and each subtree contains a vertex that no other subtree contains. Therefore, all subtrees overlap and two subtrees share an edge of T if and only if the associated vertices are adjacent in G.

Acknowledgements

The authors thank two anonymous referees for helpful suggestions, and NSERC and iCORE for financial support.

References

[1] A. Brandstädt, V.B. Le, and J.P. Spinrad, *Graph Classes: A Survey*, SIAM Monographs on Discrete Mathematics and Applications 3, SIAM, 1999.

[2] E. Čenek and L. Stewart, Maximum independent set and maximum clique algorithms for overlap graphs, *Discrete Applied Mathematics*, 131 (2003) 77-91.

[3] J.A. Enright, Subtree overlap graphs – towards recognition, MSc thesis, Department of Computing Science, University of Alberta, 2006.

[4] M.R. Garey, D.S. Johnson, G.L. Miller, and C.H. Papadimitriou, The complexity of coloring circular arcs and chords, *SIAM Journal on Algebraic and Discrete Methods*, 1 (1980) 216-227.

[5] F. Gavril, Maximum weight independent sets and cliques in intersection graphs of filaments, *Information Processing Letters*, 73 (2000) 181-188.

[6] F. Gavril, Algorithms on interval filament graphs, Technical Report CS-2007-03, Computer Science Department, Technion, Haifa, 2007.

[7] M.C. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York, 1980; Second edition, North Holland, 2004.

[8] M.C. Golumbic and R.E. Jamison, The edge intersection graphs of paths in a tree, *Journal of Combinatorial Theory, Series B*, 38 (1985) 8-22.

[9] M.C. Golumbic and E.R. Scheinerman, Containment graphs, posets, and related classes of graphs, *Combinatorial Mathematics: Proceedings of the Third International Conference* (New York, 1985) 192-204, *Annals of the New York Academy of Sciences*, 555, New York Academy of Sciences, New York, 1989.

[10] J.M. Keil, The complexity of the domination problems in circle graphs, *Discrete Applied Mathematics*. 42 (1993) 51-63.

[11] J.M. Keil and L. Stewart, Approximating the minimum clique cover and other hard problems in subtree filament graphs, *Discrete Applied Mathematics*, 154 (2006) 1983-1995.

[12] E. Marczewski, Sur deux propriétés des classes d'ensembles, *Fundamenta Mathematicae*, 33 (1945) 303-307.