All graphs with at most seven vertices are Pairwise Compatibility Graphs

T. Calamoneri[a,1], D. Frascaria[a,1], B. Sinaimeri[a,1,]

[a]*Department of Computer Science, "Sapienza" University of Rome, Via Salaria 113, 00198 Roma, Italy*

Abstract

A graph G is called a pairwise compatibility graph (PCG) if there exists an edge-weighted tree T and two non-negative real numbers d_{min} and d_{max} such that each leaf l_u of T corresponds to a vertex $u \in V$ and there is an edge $(u, v) \in E$ if and only if $d_{min} \leq d_{T,w}(l_u, l_v) \leq d_{max}$ where $d_{T,w}(l_u, l_v)$ is the sum of the weights of the edges on the unique path from l_u to l_v in T.

In this note, we show that all the graphs with at most seven vertices are PCGs. In particular all these graphs exept for the wheel on 7 vertices W_7 are PCGs of a particular structure of a tree: a centipede.

Keywords: Pairwise Comparability Graphs, Caterpillar, Centipede, Wheel.

1. Introduction

A graph $G = (V, E)$ is a *pairwise compatibility graph* (PCG) if there exists a tree T, an edge-weight function w that assigns positive values to the edges of T and two non-negative real numbers d_{min} and d_{max}, with $d_{min} \leq d_{max}$, such that each vertex $u \in V$ is uniquely associated to a leaf l_u of T and there is an edge $(u, v) \in E$ if and only if $d_{min} \leq d_{T,w}(l_u, l_v) \leq d_{max}$ where $d_{T,w}(l_u, l_v)$ is the sum of the weights of the edges on the unique path from l_u to l_v in T. In such a case, we say that G is a PCG of T for d_{min} and d_{max}; in symbols, $G = PCG(T, w, d_{min}, d_{max})$.

It is clear that if a tree T, an edge-weight function w and two values d_{min} and d_{max} are given, the construction of a PCG is a trivial problem. We focus on the reverse of this problem, i.e., given a graph G we have to find out a tree T, an edge-weight function w and suitable values, d_{min} and d_{max}, such that $G = PCG(T, w, d_{min}, d_{max})$. Such a problem is called the *pairwise compatibility tree construction problem*.

The concept of pairwise compatibility was introduced in [7] in a computational biology context and the weight function w has positive values, as it represents a not null distance. There are several known specific

graph classes of pairwise compatibility graphs, e.g., cliques and disjoint union of cliques [1], chordless cycles and single chord cycles [11], some particular subclasses of bipartite graphs [10], some particular subclasses of split matrogenic graphs [4]. Furthermore a lot of work has been done concerning some particular subclasses of PCG as leaf power graphs [1], exact leaf power graphs [2] and lately a new subclass has been introduced, namely the min-leaf power graphs [4].

Initially, the authors of [7] conjectured that every graph is a PCG, but this conjecture has been confuted in [10], where a particular bipartite graph with 15 nodes has been proved not to be a PCG. This latter result has given rise to this research as it is natural to ask for the smallest graph that is not a PCG.

A *caterpillar* Γ_n is an n-leaf tree for which any leaf is at a distance exactly one from a central path called *spine*. A *centipede* is an n-leaf caterpillar, in which the edges incident to the leaves produce a perfect matching. Deleting from an n-leaf centipede the degree two vertices and merging the two edges incident to each of these vertices into a unique edge, results in a new caterpillar that we will call *reduced centipede* and denote by Π_n (as an example, Π_5 is depicted at the top left of Fig. 1).

Caterpillars are interesting trees in the context of PCGs, as in most of the cases, the pairwise compatibility tree construction problem admits as solution a tree that is in fact a caterpillar. For this reason, we focus on this special kind of tree. In this note, we prove that all the graphs with at most seven vertices are PCGs. More precisely, we demonstrate the following results:

- If $G = PCG(\Gamma_n, w, d_{min}, d_{max})$, then there always exist a new edge-weight function w', and a new value d'_{max} such that it also holds: $G = PCG(\Pi_n, w', d_{min}, d'_{max})$.

- It is well known that graphs with five vertices or less are all PCGs and the witness trees – not all caterpillars – are shown in [9]. For each one of these graphs we prove that it is PCG of a reduced centipede, providing accordingly, an edge-weight function w and the two values d_{min} and d_{max}.

- All the graphs with six and seven vertices, except for the wheel W_7 (i.e. the graph formed by connecting a single vertex to all vertices of a cycle of length six – see Figure 2.a), are PCGs of a reduced centipede and, for each of them, we provide the edge-weight function w and the two values d_{min} and d_{max} such that it is $PCG(\Pi_n, w, d_{min}, d_{max})$, $n = 6, 7$.

- For what concerns the wheel W_7, it is known [3] that W_7 is not PCG of the reduced centipede Π_7 (and hence it is not PCG of a caterpillar). We show that W_7 is PCG of a tree different from a caterpillar.

2

2. Preliminaries

In this section we list some results that will turn out to be useful in the rest of the paper.

Let T be a tree such that there exist an edge-weight function w and two non-negative values d_{min} and d_{max} such that $G = PCG(T, w, d_{min}, d_{max})$. Observe that if T has at least 4 vertices and contains a vertex v of degree 2, then we can construct a new tree T' in which v is eliminated, the two edges (x, v) and (v, y) incident to v are merged into a unique edge (x, y) and a new function w' is defined from w only modifying the weight of the new edge, that is set equal to the sum of the weights of the old edges: $w'((x, y)) = w((x, v)) + w((v, y))$. It is easy to see that $G = PCG(T', w', d_{min}, d_{max})$. For this reason, from now on, we will assume that all the trees we handle do not contain vertices of degree two.

Proposition 1. *[5] Let $G = PCG(T, w, d_{min}, d_{max})$, where d_{min}, d_{max} and the weight $w(e)$ of each edge e of T are positive real numbers. Then it is possible to choose $\hat{w}, \hat{d}_{min}, \hat{d}_{max}$ such that for any e, the quantities $\hat{d}_{min}, \hat{d}_{max}$ and $\hat{w}(e)$ are natural numbers and $G = PCG(T, \hat{w}, \hat{d}_{min}, \hat{d}_{max})$.*

We prove here the following useful lemma:

Lemma 1. *Let $G = PCG(T, w, d_{min}, d_{max})$. It is possible to choose $\hat{w}, \hat{d}_{min}, \hat{d}_{max}$ such that $\min \hat{w}(e) = 1$, where the minimum is computed on all the edges of T, and $G = PCG(T, \hat{w}, \hat{d}_{min}, \hat{d}_{max})$.*

PROOF. According to Proposition 1, we can assume that the edge weight w and the two values d_{min}, d_{max} are integers. Let $e_1, \ldots, e_n$ be the edges of T incident to the leaves. Without loss of generality, we can assume $w(e_1) = \min_i w(e_i)$.

We define $\hat{w}$ as follows: $\hat{w}(e_1) = 1$ and for each $i = 2, \ldots, n$ define $\hat{w}(e_i) = w(e_i) - w(e_1) + 1$. Clearly, the function $\hat{w}$ is well defined as all its values are positive.

As the weight of any edge incident to a leaf has been decreased by exactly $w_1 - 1$ and the rest of the weights remained unchanged, then for of any two leaves l_i, l_j it holds that $d_{T,\hat{w}}(l_i, l_j) = d_{T,w}(l_i, l_j) - 2w(e_1) + 2$. Let $\hat{d}_{min} = \max\{d_{min} - 2w(e_1) + 2, 0\}$ and $\hat{d}_{max} = d_{max} - 2w(e_1) + 2$. It is easy to see that $G = PCG(T, \hat{w}, \hat{d}_{min}, \hat{d}_{max})$ indeed, if $\hat{d}_{min} = 0$ then it means that there was no path weight below d_{min}, with respect to w. □

The previous results imply that it is not restrictive to assume that the weights and d_{min} and d_{max} are integers and that the smallest weight is 1. Thus, in the rest of the paper we will use these assumptions.

3. PCGs of Caterpillars

In this section we will prove that we can get rid of different kinds of caterpillar structures and restrict to consider only reduced centipedes.

Theorem 1. *Let G be an n vertex graph, Γ_n and Π_n be an n-leaf caterpillar without degree 2 vertices and an n-leaf reduced centipede, respectively.*

Let $G = PCG(\Gamma_n, w, d_{min}, d_{max})$. It is possible to choose w' and d'_{max} such that $G = PCG(\Pi_n, w', d_{min}, d'_{max})$.

PROOF. In order not to overburden the exposition, let $\Gamma = \Gamma_n$ and $\Pi = \Pi_n$.

If Γ is a reduced centipede, the claim is trivially proved, so assume it is not. We lead the proof into two steps. First we define a non-negative edge-weight function w'' proving that Γ weighted by w and Π weighted by w'' generate the same PCG G for the same values d_{min} and d_{max}. Then we modify w'' into a positive weight function w' and introduce two new values d'_{min} and d'_{max} proving that G is also $PCG(\Pi, w', d'_{min}, d'_{max})$.

Draw Γ so that: i) the spine lies on a horizontal line, ii) all the leaves lie on a parallel line and iii) the edges between the spine and the leaves are represented as non-crossing line segments; number the leaves and the vertices of the spine from left to right $l_1, \ldots, l_n$ and $s_1, \ldots s_k$, $k < n$, respectively. By drawing the reduced centipede Π in a similar way, we number the leaves and the vertices of the spine from left to right by $m_1, \ldots, m_n$ and $t_2, \ldots t_{n-1}$.

We define the edge-weight function w'' as follows:

- let $p(l_i)$ the unique adjacent vertex of l_i in Γ; for each $1 < i < n$, define $w''((m_i, t_i)) = w((l_i, p(l_i)))$;

- define $w''((m_1, t_2)) = w((l_1, p(l_1)))$ and $w''((m_n, t_{n-1})) = w((l_n, p(l_n)))$;

- for each $2 \leq i \leq n - 2$, define $w''((t_i, t_{i+1})) = 0$ if and only if $p(l_i) = p(l_{i+1})$ in Γ;

- for each $2 \leq i \leq n - 2$, define $w''((t_i, t_{i+1})) = w((p(t_i), p(t_{i+1})))$ if and only if $p(l_i) \neq p(l_{i+1})$ in Γ.

Observe that w'' is well defined, as Γ has no degree 2 vertices.

It is quite easy to convince oneself that for each pair of leaves in Γ, l_i and l_j, $d_{\Gamma, w}(l_i, l_j)$ is exactly the same as $d_{\Pi, w''}(m_i, m_j)$ and that d_{min} and d_{max} remain unchanged, so $G = PCG(\Gamma, w'', d_{min}, d_{max})$.

It remains to show that we can reassign the edge-weights of Π in a way that any edge gets a positive weight and Π is the pairwise compatibility tree of G. To this purpose, we denote by $E(H)$ the edge set of any

4

graph H, and we introduce the following two quantities:

$$L = \min_{(u,v)\notin E(G)} \{|d_{min} - d_{\Pi,w''}(l_u, l_v)|, |d_{max} - d_{\Pi,w''}(l_u, l_v)|\}, \qquad N = |\{e : e \in E(\Pi), w(e) = 0\}|,$$

L is the smallest distance between the quantities d_{min}, d_{max} and the weighted distances on the tree of the paths corresponding to non-edges of G; N is the number of edges of Π of weight 0.

Observe that, unless G coincides with the clique K_n (which trivially is PCG of the reduced centipede), there always exists a pair of leaves such that their distance on Π falls out of the interval $[d_{min}, d_{max}]$ and hence $L > 0$. Furthermore, as any edge incident to a leaf in Π is strictly greater than 0 and in view of the hypothesis that the caterpillar Γ is not a reduced caterpillar, it holds $1 \leq N \leq n - 3$ (the bound $n - 3$ is reached when Γ is a star). So, the value $\epsilon = \frac{L}{N+1}$ is well defined.

Now define a new weight function w' on Π by assigning the weight ϵ to any edge of weight 0. More formally, $w'(e) = w''(e)$ if $w''(e) \neq 0$ and $w'(e) = \epsilon$ otherwise. In this way the distance between any two leaves in Π can result increased by a value upper bounded by $\epsilon N < L$.

Set the new value $d'_{max} = d_{max} + \epsilon N$.

The following three observations conclude the proof:

- any distance between leaves in Π that was strictly smaller than d_{min} with respect to the weight function w'' remains so after this transformation in view of the fact that $\epsilon N < L$;

- any distance that was strictly greater than d_{max} with respect to the weight function w'' is strictly greater than d'_{max} due to the definition of L;

- any distance that was in the interval $[d_{min}, d_{max}]$ with respect to the weight function w'' is now in the interval $[d_{min}, d'_{max}]$. $\qquad\qquad\square$

Observe that the previous statement suggests not to consider all kinds of caterpillars, but to restrict to reduced centipedes, only. In the next section we exploit this result.

4. Graphs on at most seven vertices

In this section we show that all graphs with at most seven vertices, except for the wheel W_7, are PCGs of a reduced centipede.

Analogously to what we did in the proof of Theorem 1, name the leaves of Π_n from left to right with $l_1, \ldots, l_n$ and the vertices of the spine from left to right with $s_2, \ldots s_{n-1}$. As, for any n, there exists a unique
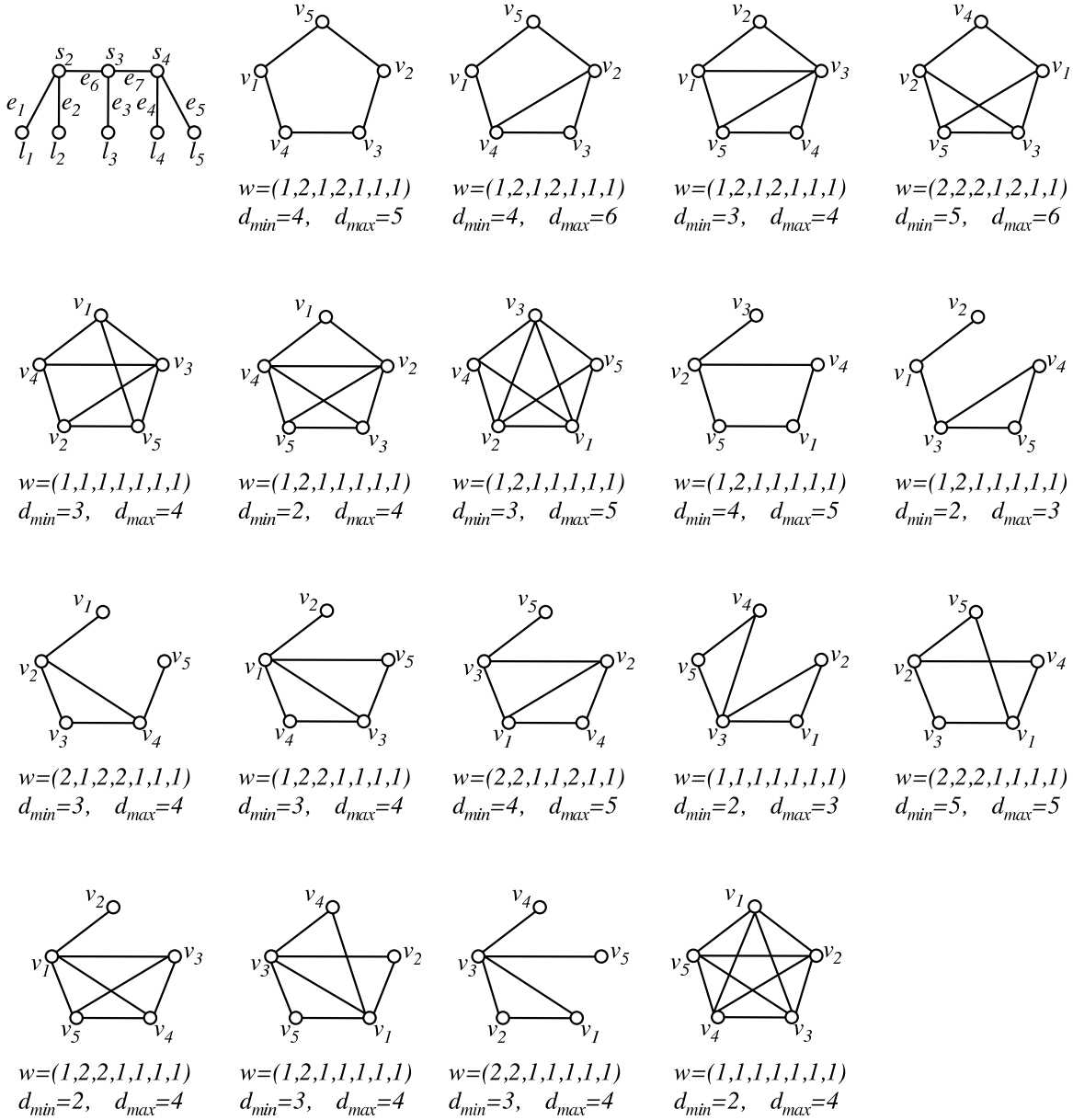
5

Figure 1: All the non isomorphic connected cyclic graphs with 5 vertices with their representation as PCGs of the reduced centipede (top left).

unlabeled reduced centipede with n leaves Π_n, in the following we consider the edges of Π_n as ordered in the following way: $e_1 = (l_1, s_2)$; $e_i = (l_i, s_i)$ for each $2 \leq i \leq n - 1$; $e_n = (l_n, s_{n-1})$; finally, $e_{n+i-1} = (s_i, s_{i+1})$ for each $2 \leq i \leq n - 2$.

Now, the edge-weight function w can be expressed as a $(2n - 3)$ long vector $\vec{w}$, where the component w_i is a positive integer representing the weight assigned to edge e_i.

In Figure 1 all the 18 connected non isomorphic cyclic graphs with 5 vertices are depicted, together with the vector $\vec{w}$ and the values of d_{min} and d_{max} that witness that all of them are PCGs of Π_5. Observe that the connected non isomorphic graphs on 5 vertices are 21, we have omitted the 3 graphs that are trees, which are trivially PCGs. We remind that it is already proved in [9] that all the graphs with at most five vertices are PCG, but the provided trees were all different and not all caterpillars.

For what concerns graphs with 6 and 7 vertices, except for the wheel W_7, we get a similar result. For the sake of brevity we do not depict all these graphs (there are 112 connected non isomorphic graphs with 6 vertices and 853 with 7 vertices), but the values of $\vec{w}$, d_{min} and d_{max} we got with the help of an enumerative C program can be found at the web page `https://sites.google.com/site/pcg6and7vertices/`. Thus, we obtain the following result:

Lemma 2. *All graphs with at most 7 vertices except for the wheel W_7 are PCGs of a reduced centipede.*

Lemma 3. *The graph W_7 is a PCG.*

PROOF. Consider the edge-weighted tree T depicted in Figure 2.b and the two values $d_{min} = 5$ and $d_{max} = 7$. It is immediate to see that $W_7 = PCG(T, w, d_{min}, d_{max})$. □
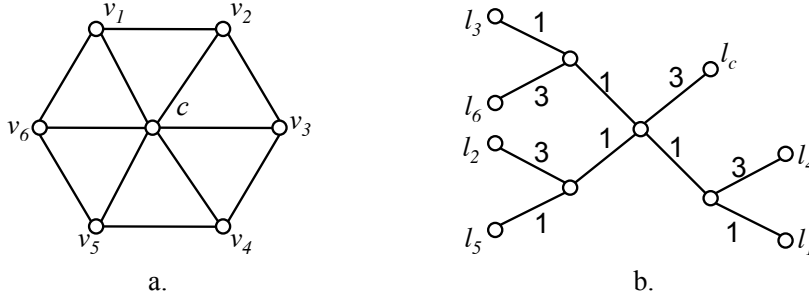

Figure 2: (a.) The wheel W_7 and (b.) the edge-weighted tree T such that $W_7 = PCG(T, w, 5, 7)$.

This result is in agreement with the negative result in [3], stating that it is not possible to find any edge-weight function w and any two values d_{min} and d_{max} such that $W_7 = PCG(\Pi_7, w, d_{min}, d_{max})$.

From Lemmas 2 and 3 it immediately derives the main result of this note:

Theorem 2. *All graphs with at most 7 vertices are PCGs.*

References

[1] A. Brandstädt. : On Leaf Powers. Technical report, University of Rostock, (2010).

[2] A. Brandstädt, V. B. Le and D. Rautenbach: Exact leaf powers, *Theor. Comput. Sci.* 411, (2010) 2968–2977.

[3] Calamoneri,T., Frangioni, A., Sinaimeri, B. On Pairwise Compatibility Graphs of Caterpillars (submitted).

[4] Calamoneri,T., Petreschi, R., Sinaimeri, B. On relaxing the constraints in pairwise compatibility graphs, In: Md. S. Rahman and S.-i. Nakano (Eds.), WALCOM 2012, Lecture Notes in Computer Science, vol. 7157, pp. 124–135, Springer, Berlin (2012).

[5] Calamoneri, T., Montefusco, E., Petreschi, R., Sinaimeri, B. Exploring Pairwise Compatibility Graphs. (submitted).

[6] J. Felsenstein: Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* 27, (1978) 401-410.

[7] P.E. Kearney, J. I. Munro and D. Phillips: Efficient generation of uniform samples from phylogenetic trees. In: Benson, G., Page, R.D.M. (eds.) WABI. Lecture Notes in Computer Science, vol. 2812, pp. 177–189. Springer, Berlin (2003).

[8] N. Nishimura, P. Ragde, D.M. Thilikos, On graph powers for leaf-labeled trees, J. Algorithms 42 (2002) 69-108.

[9] Phillips, D.: Uniform sampling from phylogenetics trees. Masters Thesis, University of Waterloo (2002).

[10] Nur Yanhaona, M., Bayzid, Md.S., Rahman, Md. S.: Discovering Pairwise compatibility graphs. J. Appl. Math. Comput. 30, 479–503 (2009).

[11] Nur Yanhaona, M., Hossain, K.S.M. T. Rahman, M. S.: Pairwise compatibility graphs. J. Appl. Math. Comput. 30, 479–503 (2009).