



2015

Path-tables of trees: a survey and some new results

Kevin Asciak

University of Malta, kevin.j.asciak@um.edu.mt

Follow this and additional works at: <https://digitalcommons.georgiasouthern.edu/tag>



Part of the [Discrete Mathematics and Combinatorics Commons](#)

Recommended Citation

Asciak, Kevin (2015) "Path-tables of trees: a survey and some new results," *Theory and Applications of Graphs*: Vol. 2 : Iss. 2 , Article 4.
DOI: 10.20429/tag.2015.020204

Available at: <https://digitalcommons.georgiasouthern.edu/tag/vol2/iss2/4>

This article is brought to you for free and open access by the Journals at Digital Commons@Georgia Southern. It has been accepted for inclusion in Theory and Applications of Graphs by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.

Path-tables of trees: a survey and some new results

Cover Page Footnote

Acknowledgement I am grateful to Virgilio Pannone for pointing out to me the problem of path-tables for trees and for his many helpful suggestions and to Josef Lauri for his valuable contribution in writing this paper.

Abstract

The (vertex) path-table of a tree T contains quantitative information about the paths in T . The entry (i, j) of this table gives the number of paths of length j passing through vertex v_i . The path-table is a slight variation of the notion of path layer matrix. In this survey we review some work done on the vertex path-table of a tree and also introduce the edge path-table. We show that in general, any type of path-table of a tree T does not determine T uniquely. We shall show that in trees, the number of paths passing through edge xy can only be expressed in terms of paths passing through vertices x and y up to a length of 4. In contrast to the vertex path-table, we show that the row of the edge path-table corresponding to the central edge of a tree T of odd diameter, is unique in the table. Finally we show that special classes of trees such as caterpillars and restricted thin trees (RTT) are reconstructible from their path-tables.

1 Historical background on tables involving number of paths

M. Randić has been a major contributor to the development of mathematical chemistry. In particular he wanted to give the concept of atomic paths a formal mathematical development based on the use of Graph Theory. In [19], he discussed in some detail the uses of the enumeration of paths and neighbours in molecular graphs and tried to characterise molecular graphs through the use of the atomic path code of a molecule. Randić conjectured that the list of path numbers determines the graph G uniquely and verified this assertion for graphs up to 11 vertices.

Later Bloom et al. [2] defined $d_{i,j}$ as the number of vertices in a graph G of diameter $d(G)$ that are at a distance j from vertex v_i . The sequence $(d_{i,0}, d_{i,1} \dots d_{i,j} \dots d_{i,d(G)})$ is then called the *distance degree sequence* of v_i in G . The $|G|$ -tuple of distance degree sequences of the vertices of G with entries arranged in lexicographic order is called the *Distance Degree Sequence of G* (DDS(G)).

Similarly they also defined the *path degree sequence* of v_i in G as the sequence $(p_{i,0}, p_{i,1} \dots p_{i,j} \dots p_{i,d(G)})$ where $p_{i,j}$ is the number of paths in G starting at vertex v_i and having length j . The ordered set of all such sequences arranged in lexicographic order is called the *Path Degree Sequence of G* (PDS(G)).

In his study, Randić remarked that since there is a unique path between pairs of atoms in acyclic structures, the number of paths of a given length corresponds to the number of neighbours at a given distance. Quintas and Slater [18] used Randić's remark in order to prove that for a connected graph G , $DDS(G) = PDS(G)$ if and only if G is a tree. They also pointed out that when G is thought of as a molecular graph, PDS(G) is precisely the lexicographically ordered list of atomic codes for the atoms (vertices) of G . Slater [24] first showed that Randić's conjecture is not valid by constructing pairs of non-isomorphic trees having the same path degree sequence and then together with Quintas [18] they constructed also non-tree graphs having a variety of properties and which also invalidate the conjecture.

Since a tree is not, in general, characterised by its Path(Distance) Degree Sequence, researchers turned their focus on finding the least possible order for the existence of pairs of non-isomorphic graphs having the same PDS. They found out that the least order for which there exists a pair of non-isomorphic trees with the same PDS is greater than 15

for trees with no vertices of degree greater than 4 and less than 19 for trees without any vertex degree restrictions [18].

V. A. Skorobogatov and A.A.Dobrynin [23] introduced the *path layer matrix*, a matrix characterising the paths that can be found in a graph. The path layer matrix of a graph G of order p is the $p \times (p-1)$ matrix $\tau(G) = ||\tau_{i,j}||$, where $\tau_{i,j}$ is the number of simple paths in G starting at vertex v_i and having length j . By ordering the rows of $\tau(G)$ in decreasing lengths (“length” here being the number of nonzero elements) and then by lexicographically arranging the rows with the same length, one can obtain a canonical form of $\tau(G)$. This path layer matrix is also known as the *path degree sequence* of a graph or the *atomic path code* of a molecule. This invariant and its modifications have found some important applications in chemistry especially in the characterisation of branching in molecules, establishing similarity of molecular graphs and for drug design [20, 21, 22]

As pointed out earlier, the path layer matrix of graph G having a sufficiently large order, does not characterise every graph G uniquely. The problem of bounding this order has been studied for over twenty years. Dobrynin and Mel’nikov [9] noticed that mathematical investigations of this matrix deal with finding a pair of non-isomorphic graphs having some specific properties and such that both graphs have the same path layer matrix. Among these properties we can have the girth, cyclomatic number and planarity of graphs [6, 9, 14, 15, 16]. Thus an interesting problem is to determine the least possible order of a pair of non-isomorphic graphs in such a class with the same path layer matrix. In fact in [18], Quintas and Slater proposed the following problem :

Does there exist a pair of connected non-isomorphic r -regular graphs (graphs in which every vertex has degree r) having the same path degree sequences?
If the answer is yes, then for each $r \geq 3$, what is the least order $p(r)$ possible for graphs in such a pair?

Defining f_1 to be the minimal order such that there exist non-isomorphic graphs of order f_1 having the same path layer matrix, Dobrynin [4] and Randić [19] showed that $12 \leq f_1 \leq 14$ for general graphs. Dobrynin restricted the problem to certain sub-classes of graphs by first considering r -regular graphs and then r -regular graphs without cut-vertices (a cut-vertex is one whose removal disconnects the graph).

Balaban et al. were the first researchers who found a pair of cubic graphs of order 142 that have the same path layer matrix [1]. Then Dobrynin [5] showed that for every $r \geq 3$, r -regular graphs with the same path layer matrix can be constructed and the least order for these pairs of graphs is a linear function of r when $r \geq 5$ while for cubic graphs $f_1 \leq 116$ and for 4-regular graphs $f_1 \leq 114$. In [7] the upperbound for the order of cubic graphs has been improved to 62.

More recently it has been discovered that if $p(r)$ is the least order of pairs of non-isomorphic r -regular graphs having the same path layer matrix then

$$20 \leq p(3) \leq 36; [8]$$

$$16 \leq p(4) \leq 18; [25]$$

$$12 \leq p(5) \leq 48; [26]$$

$$12 \leq p(6) \leq 51; [26]$$

Now since the key feature of all graphs with same path layer matrix was that they

contained cut-vertices then this prompted the investigation of finding a pair of non-isomorphic r -regular graphs without cut-vertices and having the same path layer matrix.

Defining f_2 to be the least order for pairs of non-isomorphic graphs without cutvertices and having the same path layer matrix, Dobrynin [8] proved that for cubic graphs $f_2 \leq 31$ and Yuansheng et al. [25] discovered a construction for a pair of non-isomorphic 4-regular graphs having the same path layer matrix and proved that $f_2 \leq 18$.

Recently H.Chung [3] constructed a pair of non-isomorphic 5-regular graphs without cut-vertices of order 20 having the same path layer matrix in which f_2 is reduced from 48 to 20.

After proving a converse of Kelly's Lemma in graph reconstruction, Dulio and Pannone [10] defined a slightly different kind of table which they call *path-table*. They stated that there must be a class of graphs F in between the empty class and the class of all graphs of order less than n with the following property. Let X and Y be trees of order n in which there is a bijection $f : V(X) \rightarrow V(Y)$ such that, if for every graph Q in F , the number of subgraphs of X containing a vertex x of $V(X)$ and isomorphic to Q is the same as the number of subgraphs of Y containing $f(x)$ and isomorphic to Q , then $X \simeq Y$.

There can be many such classes F but the first and most natural class to study for trees is the class F of all paths. Thus the statement to be tested becomes:

If X and Y are trees of order n and there is a bijection $f : V(X) \rightarrow V(Y)$ such that, for every path in F , the number of paths of X containing a vertex x of $V(X)$ of length l is the same as the number of paths of Y containing $f(x)$ of length l , then $X \simeq Y$.

The assumption in the above statement is tantamount to saying that X and Y have the same path-table (up to re-ordering of the rows).

Thus rather than considering paths starting at a vertex, they considered paths passing through a vertex (that is, containing a given vertex, possibly as its endvertex). Like Slater they also showed that trees having the same path-tables need not be isomorphic. In [11], they proved that there exist infinitely many pairs of non-isomorphic trees having the same path- and path layer-tables. The smallest tree that can be obtained with their construction has twenty vertices.

In this paper we shall take a closer look at these two path-tables for trees, and we shall also introduce an analogous table involving the number of paths passing through edges. We shall study possible relationships between these tables and we shall also try to identify classes of trees which are uniquely determined by some type of path-table. We shall review some known results and present some new ones.

2 Definitions

We shall start by giving uniform definitions for the three path-tables which we shall be studying. Let T be a tree and let its vertices be v_1, v_2, \dots, v_n and edges e_1, e_2, \dots, e_m . Let $d = \text{diam}(T)$ be the length of the longest path in T . For any v_i , let $\mathbf{s}_T(v_i)$ be the d -tuple whose j^{th} entry equals the number of paths of length j starting at v_i and let $\mathbf{p}_T(v_i)$ be the d -tuple where j^{th} entry equals the number of paths of length j passing through v_i . In both cases we shall drop the suffix T when the tree in question is clear from the context. We shall then denote the j^{th} entry of $\mathbf{s}(v)$ or $\mathbf{p}(v)$ by $s_j(v)$ and $p_j(v)$, respectively.

Note that, in both cases, the first entry equals the degree of v_i . The Randić table, which we shall here denote by $S(T)$, is the $n \times d$ matrix whose i^{th} row is $\mathbf{s}_T(v_i)$. Dulio and Pannone's table, which we shall here call the *Vertex path-table* of T and denoted by $VP(T)$ will be the $n \times d$ matrix where i^{th} row is $\mathbf{p}_T(v_i)$.

Now, given an edge e_i of T , let $\mathbf{p}_T(e_i)$ be the d -tuple whose j^{th} entry equals the number of paths of length j passing through e_i . The *Edge path-table* $EP(T)$ of T is the $(n-1) \times d$ matrix whose i^{th} row is $\mathbf{p}_T(e_i)$.

Again, we shall also drop the suffix T and denote the j^{th} entry of $\mathbf{p}(e_i)$ by $p_j(e_i)$. If e_i is the edge xy we shall denote this by $p_j(x, y)$. We shall also write $\mathbf{p}(x, y)$ for $\mathbf{p}(e_i)$.

Clearly, for any given T , $S(T)$, $VP(T)$ and $EP(T)$ are unique up to re-ordering of its rows, and if two Randić tables or vertex path-tables or edge path-tables are such that one can be obtained from the other by a re-ordering of its rows, we shall consider the two tables to be the same.

3 A Path-table does not determine its tree, in general

It is not a priori evident that the “passing” vector $\mathbf{p}_T(v)$ of a vertex v contains more information than the “starting” vector $\mathbf{s}_T(v)$. By definition, the first component is $\deg(v)$ for both and the i^{th} component of the former contains the i^{th} component of the latter as a summand.

But one can verify that the example pairs of non-isomorphic trees T_1, T_2 with $S(T_1) = S(T_2)$ provided by Slater [24] have $VP(T_1) \neq VP(T_2)$.

However, in [11], Dulio and Pannone provided other families of pairs of non-isomorphic trees U_1 and U_2 with the same vertex path-table.

These considerations give rise to the interesting open problems of classifying non-isomorphic trees that share a given vertex path-table and of identifying classes of trees which are uniquely determined by their path-tables. The main aim of this paper is to survey the known results in the area and to present some new ones of our own.

Figure 1 shows a pair of Slater's counterexamples on 18 vertices showing non-isomorphic trees with the same Randić-table. Table 1 shows their Randić-table but Table 2 shows their VP tables which are different.

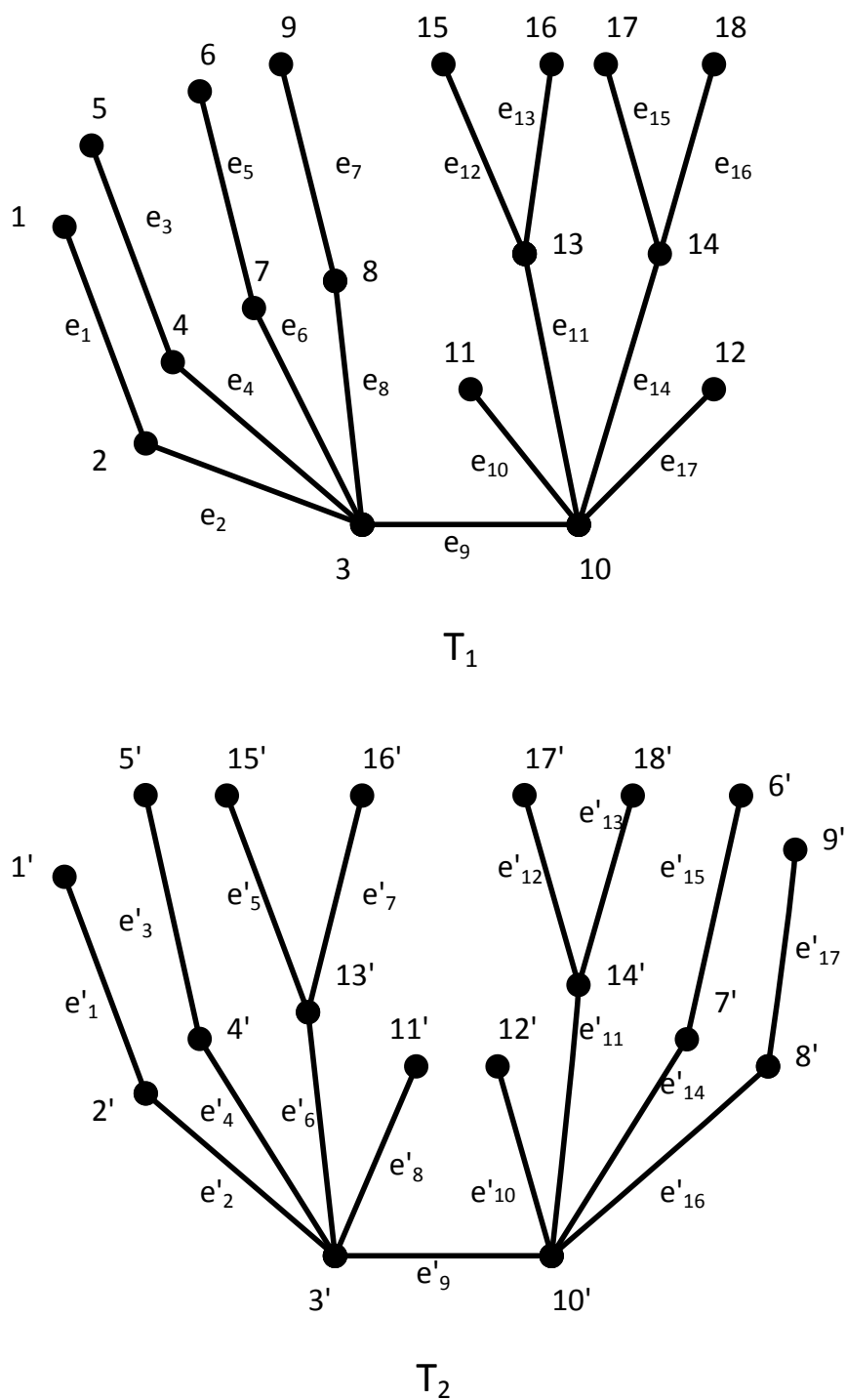


Figure 1: Slater's minimal pair of non-isomorphic trees with the same Randić-table.

$$S(T_1) = S(T_2)$$

v \ l	1	2	3	4	5
1 = 1'	1	1	4	7	4
2 = 2'	2	4	7	4	0
3 = 3'	5	8	4	0	0
4 = 4'	2	4	7	4	0
5 = 5'	1	1	4	7	4
6 = 6'	1	1	4	7	4
7 = 7'	2	4	7	4	0
8 = 8'	2	4	7	4	0
9 = 9'	1	1	4	7	4
10 = 10'	5	8	4	0	0
11 = 11'	1	4	8	4	0
12 = 12'	1	4	8	4	0
13 = 13'	3	4	6	4	0
14 = 14'	1	4	8	4	0
15 = 15'	1	2	4	6	4
16 = 16'	1	2	4	6	4
17 = 17'	1	2	4	6	4
18 = 18'	1	2	4	6	4

Table 1: The Randić-table $S(T_1) = S(T_2)$ of the smallest Slater pair in Figure 1

$$VP(T_1) \neq VP(T_2)$$

$v \setminus l$	1	2	3	4	5
1	1	1	4	7	4
2	2	5	11	11	4
3	5	18	36	38	16
4	2	5	11	11	4
5	1	1	4	7	4
6	1	1	4	7	4
7	2	5	11	11	4
8	2	5	11	11	4
9	1	1	4	7	4
10	5	18	36	36	16
11	1	4	8	4	0
12	1	1	4	7	4
13	3	7	14	16	8
14	1	1	4	7	4
15	1	2	4	6	4
16	1	2	4	6	4
17	1	2	4	6	4
18	1	2	4	6	4

$v \setminus l$	1	2	3	4	5
1'	1	1	4	7	4
2'	2	5	11	11	4
3'	5	18	36	37	16
4'	2	5	11	11	4
5'	1	1	4	7	4
6'	1	1	4	7	4
7'	2	5	11	11	4
8'	2	5	11	11	4
9'	1	1	4	7	4
10'	5	18	36	37	16
11'	1	4	8	4	0
12'	1	1	4	7	4
13'	3	7	14	16	8
14'	1	1	4	7	4
15'	1	2	4	6	4
16'	1	2	4	6	4
17'	1	2	4	6	4
18'	1	2	4	6	4

Table 2 : The vertex path-tables $VP(T_1), VP(T_2)$ of the smallest Slater pair in Figure 1.

Figure 2 shows the smallest examples on 20 vertices (found by Dulio and Pannone) of non-isomorphic trees with the same VP table. They also have the same *Randić-table*. Table 3 gives the Randić-table and the vertex path-table of U_1 and U_2 . In U_1 there are 3 pairs of vertices (8,18), (3,8) and (3,17) of degree 3 whose distance is 3 while in U_2 there exist only 2 such pairs (3', 17') and (8', 17'). Therefore they are not isomorphic.

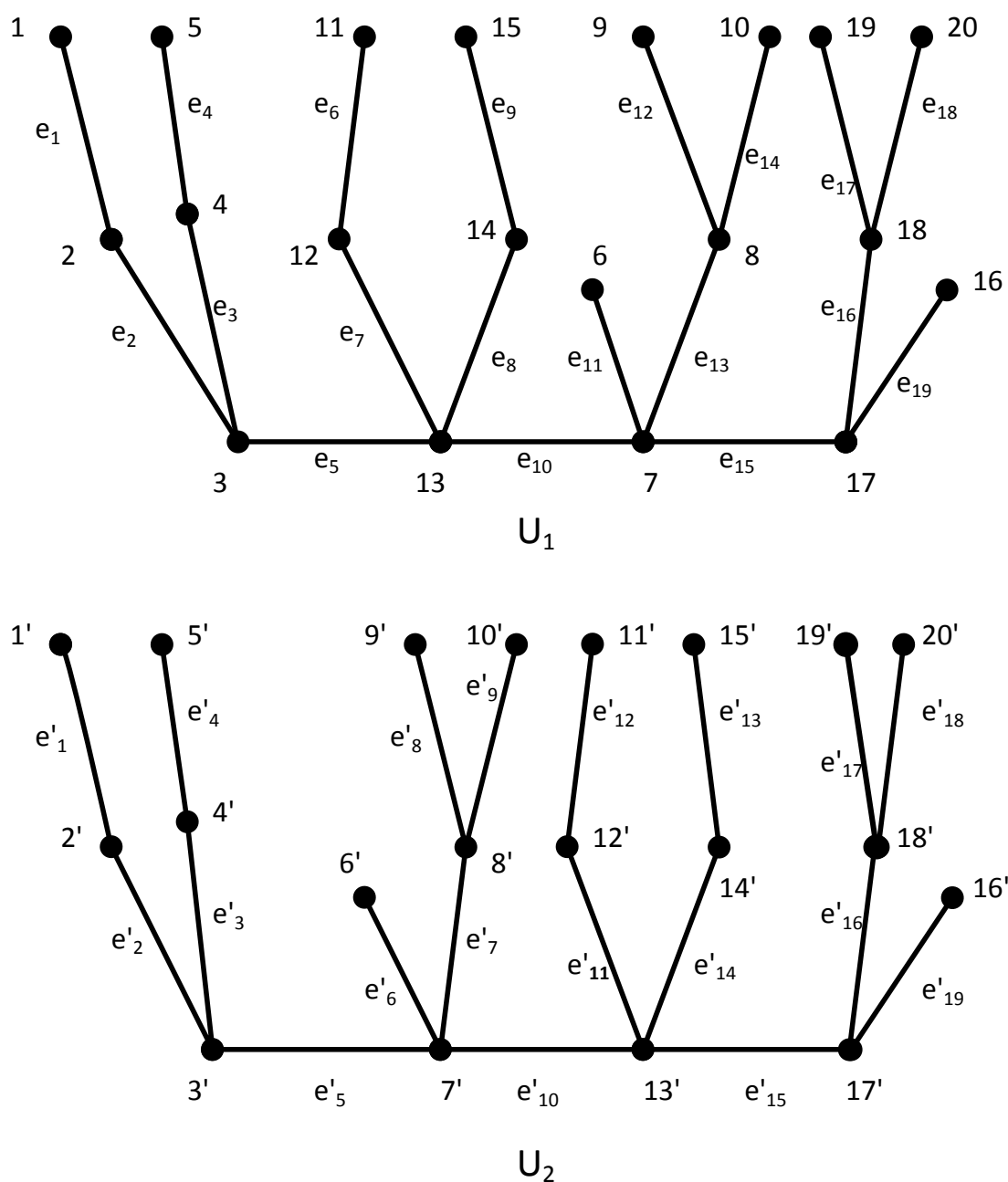


Figure 2: Minimal pair of non-isomorphic trees with the same Randić-table and path-table.

v \ l	1	2	3	4	5	6	7
1 = 1'	1	1	2	4	5	4	2
2 = 2'	2	2	4	5	4	2	0
3 = 3'	3	5	5	4	2	0	0
4 = 4'	2	2	4	5	4	2	0
5 = 5'	1	1	2	4	5	4	2
6 = 6'	1	3	7	6	2	0	0
7 = 7'	4	7	6	2	0	0	0
8 = 8'	3	3	5	6	2	0	0
9 = 9'	1	2	3	5	6	2	0
10 = 10'	1	2	3	5	6	2	0
11 = 11'	1	1	3	5	6	2	0
12 = 12'	2	3	6	6	2	0	0
13 = 13'	4	7	6	2	0	0	0
14 = 14'	2	3	6	6	2	0	0
15 = 15'	1	1	3	6	6	2	0
16 = 16'	1	2	5	5	4	2	0
17 = 17'	3	5	5	4	2	0	0
18 = 18'	3	2	3	5	4	2	0
19 = 19'	1	2	2	3	5	4	2
20 = 20'	1	2	2	3	5	4	2

v \ l	1	2	3	4	5	6	7
1 = 1'	1	1	2	4	5	4	2
2 = 2'	2	3	6	9	9	6	2
3 = 3'	3	8	15	21	20	12	4
4 = 4'	2	3	6	9	9	6	2
5 = 5'	1	1	2	4	5	4	2
6 = 6'	1	3	7	6	2	0	0
7 = 7'	4	13	27	36	32	16	4
8 = 8'	3	6	11	16	14	4	0
9 = 9'	1	2	3	5	6	2	0
10 = 10'	1	2	3	5	6	2	0
11 = 11'	1	1	3	5	6	2	0
12 = 12'	2	4	9	12	8	2	0
13 = 13'	4	13	27	37	32	16	4
14 = 14'	2	4	9	12	8	2	0
15 = 15'	1	1	3	6	6	2	0
16 = 16'	1	2	5	5	4	2	0
17 = 17'	3	8	15	20	20	12	4
18 = 18'	3	5	7	11	14	10	4
19 = 19'	1	1	2	3	5	4	2
20 = 20'	1	1	2	3	5	4	2

$$S(U_1) = S(U_2)$$

$$VP(U_1) = VP(U_2)$$

TABLE 3 : The Randić-table and the vertex path-table of the smallest pair of non-isomorphic trees in Figure 2

Tables 4 and 5 give two examples of edge path-tables for the trees T_1 , T_2 and U_1 , U_2 we showed in Figures 1 and 2. Therefore, just as for the Randić-table and the vertex path-table, the edge path-table does not distinguish between T_1 and T_2 or between U_1 and U_2 .

e \ l	1	2	3	4	5
1 = 1'	1	1	4	7	4
2 = 2'	1	5	11	11	4
3 = 3'	1	1	4	7	4
4 = 4'	1	5	11	11	4
5 = 15'	1	1	4	7	4
6 = 14'	1	5	11	11	4
7 = 17'	1	1	4	7	4
8 = 16'	1	5	11	11	4
9 = 9'	1	8	24	32	16
10 = 8'	1	4	8	4	0
11 = 6'	1	6	14	16	8
12 = 5'	1	2	4	6	4
13 = 7'	1	2	4	6	4
14 = 11'	1	6	14	16	8
15 = 12'	1	2	4	6	4
16 = 13'	1	2	4	6	4
17 = 10'	1	4	8	4	0

Table 4 : The edge path-table for trees T_1 and T_2 .

e \ l	1	2	3	4	5	6	7
1 = 1'	1	1	2	4	5	4	2
2 = 2'	1	3	6	9	9	6	2
3 = 3'	1	3	6	9	9	6	2
4 = 4'	1	1	2	4	5	4	2
5 = 5'	1	5	13	20	20	12	4
6 = 12'	1	1	3	6	6	2	0
7 = 11'	1	4	9	12	8	2	0
8 = 14'	1	4	9	12	8	2	0
9 = 13'	1	1	3	6	6	2	0
10 = 10'	1	6	17	28	28	16	4
11 = 6'	1	3	7	6	2	0	0
12 = 8'	1	2	3	5	6	2	0
13 = 7'	1	5	11	16	14	4	0
14 = 9'	1	2	3	5	6	2	0
15 = 15'	1	5	13	20	20	12	4
16 = 16'	1	4	7	11	14	10	4
17 = 17'	1	2	2	3	5	4	2
18 = 18'	1	2	2	3	5	4	2
19 = 19'	1	2	5	5	4	2	0

TABLE 5 : The edge path-table for trees U_1 and U_2 .

4 Relation between Randić , Vertex and Edge path-tables

The introduction of the edge path-table adds further topics which this paper discusses. Apart from investigating which trees can be reconstructed from a knowledge of one of its tables, we shall also consider the relationships between tables of the same tree, in particular, between $VP(T)$ and $EP(T)$, particularly the question of determining $\mathbf{p}(x, y)$ from a knowledge of $\mathbf{p}(x)$, $\mathbf{p}(y)$ and possibly $\mathbf{p}(v)$ for some neighbours v of x and y .

Tables 2 and 4 show that the edge path-table does not determine the vertex path-table. However, we do not know whether the vertex path-table determines the edge path-table. Also, in Slater's pairs of trees T_1, T_2 , we can see from Table 4 that edges e_9 and e'_9 representing (v_3, v_{10}) and (v'_3, v'_{10}) respectively, give an equal path vector of $(1, 8, 24, 32, 16)$ but from Table 2 one can see that $\mathbf{p}(v_3) \neq \mathbf{p}(v'_3)$ and $\mathbf{p}(v_{10}) \neq \mathbf{p}(v'_{10})$. Therefore, $\mathbf{p}(a, b)$ does not determine $\mathbf{p}(a)$, $\mathbf{p}(b)$ in general. We therefore pose the following natural problems.

Problem 4.1 *If $\mathbf{p}_{T_1}(a) = \mathbf{p}_{T_2}(a')$ and $\mathbf{p}_{T_1}(b) = \mathbf{p}_{T_2}(b')$ (where possibly $T_1 = T_2$) when does it follow that $\mathbf{p}_{T_1}(a, b) = \mathbf{p}_{T_2}(a', b')$? In general, given two non-isomorphic trees T_1 and T_2 with $VP(T_1) = VP(T_2)$, when does it follow that $EP(T_1) = EP(T_2)$?*

Problem 4.2 *If $\mathbf{s}_{T_1}(a) = \mathbf{s}_{T_2}(b)$ when does it follow that $\mathbf{p}_{T_1}(a) = \mathbf{p}_{T_2}(b)$? In general, given two non-isomorphic trees T_1 and T_2 with the same Randić table, when does it follow that $VP(T_1) = VP(T_2)$?*

We shall first start by considering this problem.

4.1 $\mathbf{p}(x, y)$ is determined by $\mathbf{s}(x)$ and $\mathbf{s}(y)$

In this section we shall prove the following theorem

Theorem 4.1 *Let T, T' be two trees (possibly $T = T'$). Let x, y be two adjacent vertices in T and x', y' be two adjacent vertices in T' . Assume that $\mathbf{s}(x) = \mathbf{s}(x')$ and $\mathbf{s}(y) = \mathbf{s}(y')$. Then for edge xy in T_1 and edge $x'y'$ in T_2 $\mathbf{p}(x, y) = \mathbf{p}(x', y')$.*

But first we need to develop some notation and a few results.

4.1.1 Definitions and notation

Given a tree T and one of its edges xy , we may consider T to be the union of a positive subtree T^+ and a negative subtree T^- . Thus for any edge xy of a tree T we define the *positive-subtree* T^+ of T with respect to edge xy to be the maximal subtree of T containing vertex y but not vertex x .

Analogously we define the *negative-subtree* T^- of T . The decision of which subtree to call positive or negative is, of course, arbitrary.

Recall that $s_l(x)$ is defined to be the number of paths of length l that start from vertex x so that we denote $s_l^-(x)$ to be the number of paths of length l that start from vertex x and lie within the subtree T^- . Analogously we define $s_l^+(y)$. In various formulae it is useful to allow the length l to be equal to zero so that conventionally $s_0^-(x) = s_0^+(y) = 1$.

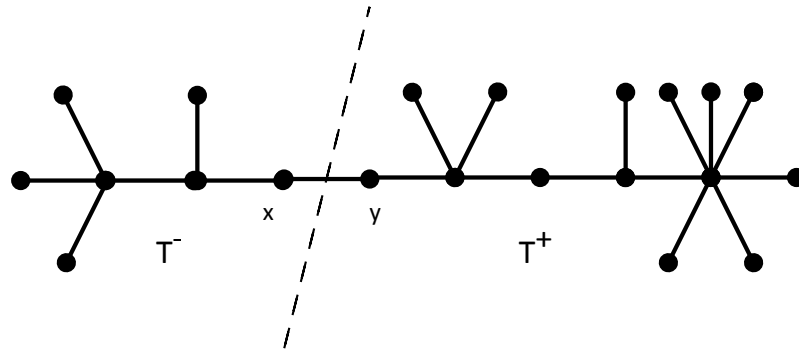


Figure 3: The positive-subtree T^+ and negative-subtree T^- of a tree T .

We shall define $p_l(x, \bar{y})$ to be the number of paths of length l passing through x but not through y and analogously define $p_l(\bar{x}, y)$ to be the number of paths of length l passing through y but not through x so that

$$p_l(x) = p_l(x, y) + p_l(x, \bar{y})$$

and

$$p_l(y) = p_l(x, y) + p_l(\bar{x}, y)$$

4.1.2 A useful formula for the number of paths passing through an edge in terms of starting vectors

Given a tree T and an edge xy in T we want to find the number of paths of length l passing through edge xy in terms of $\mathbf{p}(x)$ and $\mathbf{p}(y)$. Suppose that the number of paths of length i ($1 \leq i \leq l$) starting from x and belonging to T^- is equal to $s_i^-(x)$. Then for each $s_i^-(x)$ we determine the number of paths of length j starting from y and lying within T^+ where $j = l - i - 1$, and which is equal to $s_j^+(y)$. Then the number of paths of length l passing through edge xy can be determined by summing $s_i^-(x)s_{l-i-1}^+(y)$ for $1 \leq i \leq l$. Thus

$$p_l^T(x, y) = \sum_{i, j \geq 0} s_i^-(x)s_j^+(y)$$

where $i + j = l - 1$ and $l \geq 1$.

If the edge considered is an end-edge then the formula for determining the number of paths passing through this edge becomes simpler :

$$p_l^T(x, y) = s_l(x)$$

where $i = l - 1$ and $l \geq 1$.

Lemma 4.1 Let $s_l(x)$ be the number of paths of length l starting from vertex x and $s_l^-(x)$ be the number of paths of length l starting from vertex x and belonging to T^- as previously

described. Let x and y be two adjacent vertices in the tree T . Then

$$s_{2n}^-(x) = \sum_{i=0}^n s_{2i}(x) - \sum_{i=0}^{n-1} s_{2i+1}(y) \quad (1)$$

and

$$s_{2n+1}^-(x) = \sum_{i=0}^n s_{2i+1}(x) - \sum_{i=0}^n s_{2i}(y) \quad (2)$$

Proof. To prove statement (1), we use induction on the number of paths of length l . Consider $s_2^-(x)$ which represents the number of paths of length 2 starting from x but not containing y . Now $s_2(x)$ gives the number of paths of length 2 starting from x so that it includes those paths that pass through y . To eliminate such paths we consider paths of length 1 starting from y and belonging to T^+ . These are represented by $s_1^+(y) = \deg(y) - 1 = s_1(y) - s_0(x)$. Thus $s_2^-(x) = s_2(x) - s_1^+(y) = s_2(x) - s_1(y) + s_0(x)$. By the induction hypothesis $s_{2k}^-(x) = \sum_{i=0}^k s_{2i}(x) - \sum_{i=0}^{k-1} s_{2i+1}(y)$.

Consider $s_{2k+2}^-(x)$ that represents the number of paths of length $2k+2$ that start from x but do not contain vertex y . Now $s_{2k+2}(x)$ gives the number of paths of length $2k+2$ that start from x . Some of these paths belong to T^- which contribute to the total sum of $s_{2k+2}^-(x)$ while the remaining paths that pass through y must not be considered. As a result we consider paths of length $2k+1$ starting from y and belonging to T^+ . Thus $s_{2k+2}^-(x) = s_{2k+2}(x) - s_{2k+1}^+(y)$.

Similar arguments on $s_{2k+1}^+(y)$ give $s_{2k+1}^+(y) = s_{2k+1}(y) - s_{2k}^-(x)$ so that $s_{2k+2}^-(x) = s_{2k+2}(x) - s_{2k+1}(y) + s_{2k}^-(x)$.

By the induction hypothesis

$$s_{2k+2}^-(x) = s_{2k+2}(x) - s_{2k+1}(y) + \sum_{i=0}^k s_{2i}(x) - \sum_{i=0}^{k-1} s_{2i+1}(y)$$

so that

$$s_{2k+2}^-(x) = \sum_{i=0}^{k+1} s_{2i}(x) - \sum_{i=0}^k s_{2i+1}(y)$$

A similar proof holds for statement (2). \square

Now using the above results and notations we can prove Theorem 4.1.

Proof of Theorem 4.1 We use the formula for the number of paths of length l passing through edge xy

$$p_l^T(x, y) = \sum_{i,j \geq 0} s_i^-(x) s_j^+(y)$$

where $i+j = l-1$ and $l \geq 1$. Assume $s_0^-(x) = 1$ and use the fact that $s_1^-(x) = \deg(x) - 1$ together with the results obtained in lemma 4.1, namely,

$$s_{2n}^-(x) = \sum_{i=0}^n s_{2i}(x) - \sum_{i=0}^{n-1} s_{2i+1}(y)$$

and

$$s_{2n+1}^-(x) = \sum_{i=0}^n s_{2i+1}(x) - \sum_{i=0}^n s_{2i}(y)$$

for $n \geq 1$. Analogous formulae hold for $s_j^+(y)$. For every path length l , $p_l(x, y)$ can be expressed in terms of $\mathbf{s}(x) = (s_0(x), s_1(x), \dots, s_l(x))$ and $\mathbf{s}(y) = (s_0(y), s_1(y), \dots, s_l(y))$. Similarly $p_l(x', y')$ can be expressed in terms of $\mathbf{s}(x') = (s_0(x'), s_1(x'), \dots, s_l(x'))$ and $\mathbf{s}(y') = (s_0(y'), s_1(y'), \dots, s_l(y'))$. Thus from the assumption $\mathbf{s}(x) = \mathbf{s}(x')$ and $\mathbf{s}(y) = \mathbf{s}(y')$ the result follows. \square

Remark. Slater's pair of non-isomorphic trees on 18 vertices, shown in Figure 1, gives an example of non-isomorphic trees T_1, T_2 such that $S(T_1) = S(T_2)$ but $VP(T_1) \neq VP(T_2)$ as can be seen by studying Tables 1 and 2.

We can also notice from Table 2 that although $\mathbf{p}(v_{10}) \neq \mathbf{p}(v_{3'})$ and $\mathbf{p}(v_{11}) = \mathbf{p}(v_{11'})$ yet, from Table 4 and Figure 1 we can see that $\mathbf{p}(e_{10}) = \mathbf{p}(e_{8'})$ where edge $e_{10} = v_{10}v_{11}$ and edge $e_{8'} = v_{3'}v_{11'}$. This follows from the fact that $\mathbf{s}(v_{10}) = \mathbf{s}(v_{3'})$ and $\mathbf{s}(v_{11}) = \mathbf{s}(v_{11'})$, confirming Theorem 4.1.

4.2 $p_l(x, y)$ is not, in general, determined by $\mathbf{p}(x)$ and $\mathbf{p}(y)$

We have shown in Theorem 4.1 that knowing the number of paths *starting from* vertex x and *from* vertex y gives the number of paths *passing through* edge xy . The situation is quite different when we turn to Problem 4.1, that is, if we start from the number of paths passing through vertices x and y . We shall show that in general trees, $p_l(x, y)$ can only be expressed in terms of $p_i(x)$ and $p_i(y)$ for up to $l = 4$ where $1 \leq i \leq 4$.

Let xy be an edge in a graph T . Then it is obvious that

$$p_1(x, y) = 1 \tag{1}$$

It is also easy to see that

$$p_2(x, y) = (\deg(x) - 1) + (\deg(y) - 1) = p_1(x) + p_1(y) - 2 \tag{2}$$

To express $p_3(x, y)$ in terms of $s_i(x)$ and $s_i(y)$ for $i = 1, 2$ we need to consider all paths of length 2 starting from x and belonging to T^- , paths of length 2 starting from y and belonging to T^+ , paths of length 1 starting from x and belonging to T^- and paths of length 1 starting from y and belonging to T^+ . Thus

$$p_3(x, y) = (s_2(x) - s_1(y) + 1) + (s_2(y) - s_1(x) + 1) + (s_1(x) - 1)(s_1(y) - 1) \tag{*}$$

But as

$$p_1(x) = s_1(x)$$

and

$$p_2(x) = s_2(x) + \binom{s_1(x)}{2}$$

then

$$p_2(x) = s_2(x) + \binom{p_1(x)}{2}$$

and analogously

$$p_1(y) = s_1(y)$$

and

$$p_2(y) = s_2(y) + \binom{p_1(y)}{2}$$

Substituting the formulae derived of $s_1(x)$, $s_1(y)$, $s_2(x)$ and $s_2(y)$ in (*) we obtain

$$\begin{aligned} p_3(x, y) &= (p_2(x) - \binom{p_1(x)}{2}) - p_1(y) + 1 \\ &+ (p_2(y) - \binom{p_1(y)}{2}) - p_1(x) + 1 \\ &+ (p_1(x) - 1)(p_1(y) - 1) \end{aligned} \quad (3)$$

Similarly to what we have done with $p_3(x, y)$,

$$\begin{aligned} p_4(x, y) &= (s_3(x) - s_2(y) + s_1(x) - 1) \\ &+ (s_3(y) - s_2(x) + s_1(y) - 1) \\ &+ (s_2(x) - s_1(y) + 1)(s_1(y) - 1) \\ &+ (s_2(y) - s_1(x) + 1)(s_1(x) - 1) \end{aligned} \quad (**)$$

in which

$$\begin{aligned} p_3(x) &= s_3(x) - s_2(x) + s_2(x)s_1(x) \\ &= s_3(x) + s_2(x)[s_1(x) - 1] \end{aligned}$$

so that

$$s_3(x) = p_3(x) - [p_2(x) - \binom{p_1(x)}{2}][p_1(x) - 1]$$

Recalling formulae for $s_1(x)$, $s_2(x)$ and $s_3(x)$ and analogous formulae for $s_1(y)$, $s_2(y)$ and $s_3(y)$ and substituting in (**) we get an expression for $p_4(x, y)$ in terms of $p_i(x)$ and $p_i(y)$ for $i = 1, 2, 3$ as shown.

$$\begin{aligned}
p_4(x, y) &= p_3(x) - [p_2(x) - \binom{p_1(x)}{2}][p_1(x) - 1] \\
&- p_2(y) + \binom{p_1(y)}{2} + p_1(x) - 1 \\
&+ p_3(y) - [p_2(y) - \binom{p_1(y)}{2}][p_1(y) - 1] \\
&- p_2(x) + \binom{p_1(x)}{2} + p_1(y) - 1 \\
&+ [p_2(x) - \binom{p_1(x)}{2} - p_1(x) + 1][p_1(y) - 1] \\
&+ [p_2(y) - \binom{p_1(y)}{2} - p_1(y) + 1][p_1(x) - 1]
\end{aligned} \tag{4}$$

We are therefore able to determine up to $p_4(x, y)$ in terms $p_i(x)$ and $p_i(y)$ for smaller values of i . However the following example shows that it is not possible to express $p_l(x, y)$ in terms of $p_i(x)$ and $p_i(y)$, $i \leq l$, for paths of length l greater than 4. In fact, the following example shows that $p_l(x, y)$ cannot be expressed even in terms of $\mathbf{p}(x)$ and $\mathbf{p}(y)$.

Example 4.1 let T'' and T''' be two identical trees as shown in Figure 4.

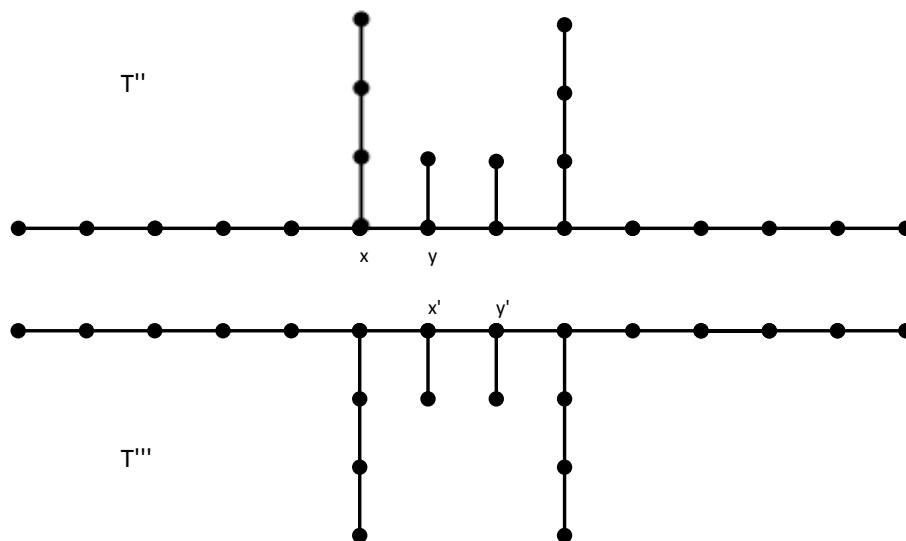


Figure 4: Two identical trees T'' and T''' ; $\mathbf{p}(x) = \mathbf{p}(x')$ and $\mathbf{p}(y) = \mathbf{p}(y')$ but $p_5(x, y) \neq p_5(x', y')$

Path length	1	2	3	4	5	6	7	8	9	10	11	12	13
$\mathbf{p}(x) = \mathbf{p}(x')$	3	7	12	16	19	20	19	16	12	8	5	2	1
$\mathbf{p}(y) = \mathbf{p}(y')$	3	7	12	16	19	20	19	16	12	8	5	2	1

Table 6 : Path vectors for vertices x and y in tree T'' and for vertices x' and y' in tree T''' in Figure 4.

Path length	1	2	3	4	5	6	7	8	9	10	11	12	13
$\mathbf{p}(x, y)$	1	4	8	12	15	17	17	15	12	8	5	2	1
$\mathbf{p}(x', y')$	1	4	8	12	16	18	17	16	12	8	5	2	1

Table 7 : Path vectors for edges xy and $x'y'$ in trees T'' and T''' respectively in Figure 4.

One can see that $\mathbf{p}(x) = \mathbf{p}(x')$ and $\mathbf{p}(y) = \mathbf{p}(y')$ but $p_5(x, y) = 15$ while $p_5(x', y') = 16$. Thus in this example $p_l(x, y) = p_l(x', y')$ only for $1 \leq l \leq 4$ and $9 \leq l \leq 13$.

4.3 Determination of $p_l(x, y)$ in terms of $\mathbf{p}(x)$ and $\mathbf{p}(y)$ in some special cases of trees

We have shown that, at least for general trees, it is impossible to express $p_l(x, y)$ in terms of $\mathbf{p}(x)$ and $\mathbf{p}(y)$. However, it may be possible that for some special cases of trees, such as caterpillars and thin trees (defined below), we are able to express $p_l(x, y)$ in terms of $\mathbf{p}(x)$ and $\mathbf{p}(y)$ and the path-vectors of other nearby vertices of x and y .

Case(i) The Caterpillar case

Recall that a caterpillar is a tree such that the deletion of its endvertices results in a path called the *spine* of the caterpillar. A caterpillar whose spine is the path v_1, v_2, \dots, v_s and such that the vertex v_i is adjacent to a_i endvertices will be denoted by $C(a_1, a_2, \dots, a_s)$. Here we shall show that in a caterpillar it is possible to express $p_l(x, y)$ in terms of just $\mathbf{p}(x)$ and $\mathbf{p}(y)$.

Proposition 4.1 *Let C be a caterpillar, x and y are two adjacent vertices in which $\deg(x) = h$ and $\deg(y) = k$. Then for any $t \geq 1$, the number $p_t(x, y)$ depends only on $p_i(x)$ and $p_i(y)$ for $i = 1, 2, 3, \dots, t - 1$.*

Proof. For every path of length l , we have shown in section 4.1 that

$$\begin{aligned} p_l(x) &= p_l(x, \bar{y}) + p_l(x, y) \\ &= s_l^-(x) + [h - 2]s_{l-1}^-(x) + \sum_{\substack{i+j=l-1 \\ i,j \geq 0}} s_i^-(x)s_j^+(y) \end{aligned}$$

and

$$\begin{aligned} p_l(y) &= p_l(\bar{x}, y) + p_l(x, y) \\ &= s_l^+(y) + [k - 2]s_{l-1}^+(y) + \sum_{\substack{i+j=l-1 \\ i,j \geq 0}} s_i^-(x)s_j^+(y) \end{aligned}$$

Now $p_l(x)$ and $p_l(y)$ are known for $l = 1, 2, \dots, t - 1$. Thus we can express $s_l^-(x)$ and $s_l^+(y)$ in terms of analogous unknowns $s_r^-(x)$ and $s_r^+(y)$ where $r < l$. Since we know that $s_1^-(x) = h - 1$ and $s_1^+(y) = k - 1$ we can inductively calculate $s_i^-(x)$ and $s_i^+(y)$ for $i = 1, 2, \dots, t - 1$. Then $p_t(x, y)$ can be obtained from $\sum_{\substack{i+j=t-1 \\ i,j \geq 0}} s_i^-(x)s_j^+(y)$. \square

Case(ii) The Thin Tree case

A tree T is a *thin tree* if, removing every endvertex of T leaves a caterpillar. The *trunk*

of a tree is the intersection of all the longest paths of the tree; a vertex on the trunk is called a *trunk vertex*. A *restricted thin tree* (RTT) is a thin tree all of whose trunk vertices have degree at most 3. Figure 5 shows a thin tree and a restricted thin tree.

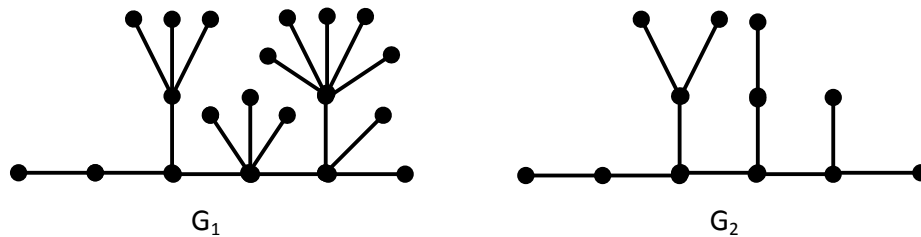


Figure 5: A thin tree G_1 and a restricted thin tree G_2 .

Figure 4 gave an example of a non-thin tree for which $p_5(x, y)$ cannot be expressed in terms of $p_i(x)$ and $p_i(y)$ for $i < 5$. The simplest tree which we can now consider which is not a caterpillar, is either a thin or a restricted thin tree. But in both situations the problem of whether $p_5(x, y)$ can be expressed in terms of $p_i(x)$ and $p_i(y)$ for $i < 5$, is still open.

5 Uniqueness of the path-vector of the central vertex and central edge

In this section we shall first present Dulio and Pannone's result which states that a central and a non-central vertex in a tree T may have the same path-vector so that given its vertex path-table, it is not always possible to recognise the path-vector representing the centre of the tree T . In contrast we shall prove that the path-vector of the central edge of a tree of odd diameter is unique in its edge path-table. Thus this leads us to investigate whether edges having different eccentricities can have different path-vectors.

First we need to recall some definitions which will be used throughout this section. Suppose G is a finite graph of order n and v is a vertex in G . Then the *eccentricity* of the vertex v denoted by $\text{ecc}(v)$ is defined as the maximum distance from v to any vertex in G , that is, $\text{ecc}(v) = \max\{d(v, x) : x \in V(G)\}$ where d is the natural metric. Thus the *diameter* $\text{diam}G$ and *radius* $\text{rad}G$ can be respectively defined as the maximum and minimum eccentricity of the vertices of G . If T is a tree, then we recall that the *trunk* of T , denoted by $\text{Tr}(T)$ is the set of those vertices of T contained in all paths of length equal to $\text{diam}T$. The centre of T , denoted by $Z(T)$ is the set of all vertices with minimum eccentricity and can have either one or two adjacent vertices depending on whether $\text{diam}T$ is even or odd, respectively.

Given any $v \in V(\text{Tr}(T))$ we define the *branch* from v , denoted by $\text{Br}(v)$, to be the maximal connected subgraph H of T containing v such that $H \cup \text{Tr}(T) = \{v\}$. The *ramification* $\text{ram } v$ of $v \in V(\text{Tr}(T))$ is defined to be the eccentricity within $\text{Br}(v)$.

In [12], Dulio and Pannone showed that if T is a tree with $\text{diam } T \leq 7$, then the path-vector of $Z(T)$ is unique. In general we cannot determine which vector in a given vertex path-table of a tree represents the centre since there can be two identical vectors,

one which represents the centre and the other representing a non-central vertex. The following Example 5.1 confirms the previous statement.

Example 5.1 *The tree H with diameter 10 in which three consecutive vertices x, c, y have the same path-vector, where c is the central vertex.*

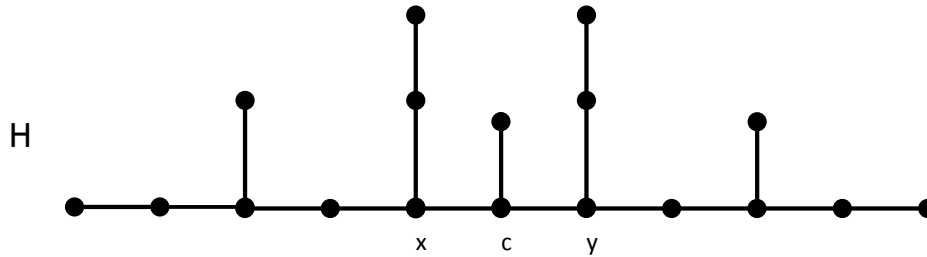


Figure 6: The tree H in which the path-vector of the centre c and two of its adjacent vertices x and y are $(3, 7, 12, 16, 18, 16, 12, 8, 4, 1)$.

Also Examples 4.1 and 5.1 show also that two vertices in a tree with different eccentricities can have equal path-vectors.

Dulio and Pannone [12] discovered that the vectors of the centre $Z(T) = \{c_L, c_R\}$ (possibly $c_L = c_R$) are(is) unique if at least one of the following holds:

- (i) $|Tr(T)| = 1$ (which implies $|Z(T)| = 1$)
- (ii) $\min\{\text{ram } c_L, \text{ram } c_R\} \geq \text{ram } x$ for all $x \in Tr(T)$
- (iii) $\text{ram } c_L = \text{ram } c_R = \lfloor \frac{D}{2} \rfloor - 1$ where D is the diameter of T
- (iv) $\deg c = 2$ where c is the unique central vertex

We shall now show that the path-vector of the central edge (when T is bicentral) is unique.

Proposition 5.1 *Let xy be the central edge of the bicentral tree T and vw be another edge in a tree T . If $p_l(x, y) = p_l(v, w)$, then both vertices v and w belong to the trunk of T .*

Proof. Since xy is the central edge, then all paths of maximum length pass through it. Hence by assumption all paths of maximum length also pass through vw . Therefore v and w belong to the trunk of T . \square

Proposition 5.2 *For any length l , the following statements are equivalent:*

- (i) $p_l(x, y) = p_l(v, w)$
- (ii) *the number of paths of length l passing through xy but not through vw equals the number of paths of length l passing through vw but not through xy .*

Proof. Define N_1 to be the number of paths of length l passing through xy but not through vw and N_2 to be the number of paths of length l passing through vw but not through xy . Since $p_l(x, y) = N_1 + M$ and $p_l(v, w) = N_2 + M$ where M is the number of paths of length l passing through both xy and vw then result follows. \square

Theorem 5.1 *The path-vector of the central edge of a tree of odd diameter is unique in its path-table.*

Proof. Let xy be the central edge and vw be another edge in a tree T . Assume for contradiction that, for any length l , $p_l(x, y) = p_l(v, w)$. Then, by Proposition 5.1, edge vw belongs to the trunk of T . Suppose that P is one of the longest paths passing through vw but not through xy . Let S be a subtree attached to the trunk vertex s that contains an end-segment of P , which we call P_s as shown in Figure 7. Then path P is made up of path H' consisting of vertices in the trunk and P_s such that path H' is smaller than half the diameter of the tree.

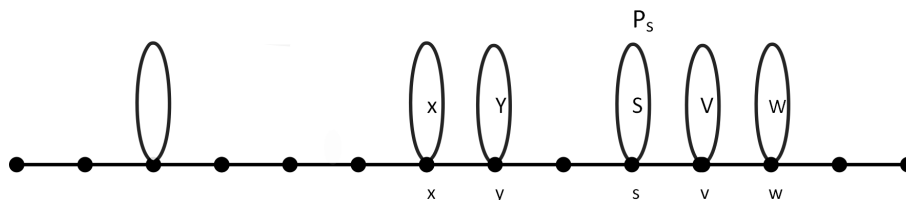


Figure 7: A tree of odd diameter with central edge xy .

Let us construct path Q by splicing three paths : a path H connecting vertex y to a peripheral vertex such that it passes through xy but not through vw , path $[y, \dots, s]$ (of length ≥ 0) consisting of vertices in the trunk and path P_s . Path Q is strictly longer than path P since H is longer than half the diameter of the tree. Suppose l_Q is the length of Q . Then there exist no paths of length l_Q passing through vw but not through xy so that by Proposition 5.2, $p(x, y) \neq p(v, w)$. \square

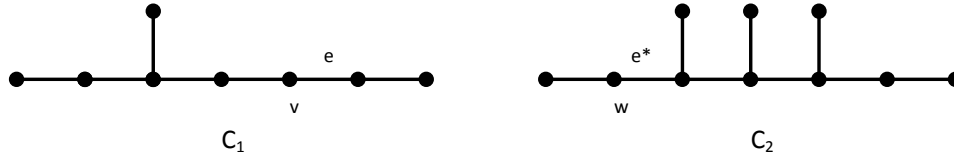
5.1 Vertex and edge eccentricities

In [13], Dulio and Pannone gave the following three results.

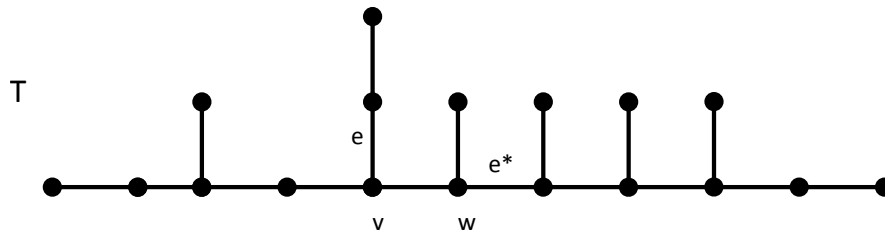
Result 1 : Two vertices v and w in two distinct caterpillars of equal diameter, may have the same path-vector even though they have different eccentricities in their respective caterpillars.

Result 2 : Consider a vertex v in a caterpillar C_1 and a vertex w in another caterpillar C_2 in which $\text{ecc}(v) < \text{ecc}(w)$ but $\mathbf{p}_{C_1}(v) = \mathbf{p}_{C_2}(w)$. Then a tree T can be constructed by **joining** v and w by a new edge wherein v and w still have different eccentricities but equal path-vectors.

Example 5.2 *In Figure 8 there exists a vertex v in caterpillar C_1 and a vertex w in caterpillar C_2 in which $\text{ecc}(w) > \text{ecc}(v)$ and $\mathbf{p}_{C_1}(v) = \mathbf{p}_{C_2}(w) = (2, 3, 4, 4, 3, 1)$.*


 Figure 8: The two caterpillars C_1 and C_2 .

Since v and w become adjacent in T , then for each length l , the additional new paths of length l occurring in T must pass through both v and w so that $\mathbf{p}_T(v) = \mathbf{p}_T(w) = (3, 7, 12, 16, 18, 16, 12, 8, 4, 1)$.


 Figure 9: Joining the two caterpillars C_1 and C_2 by adding a new edge vw to form a thin tree T .

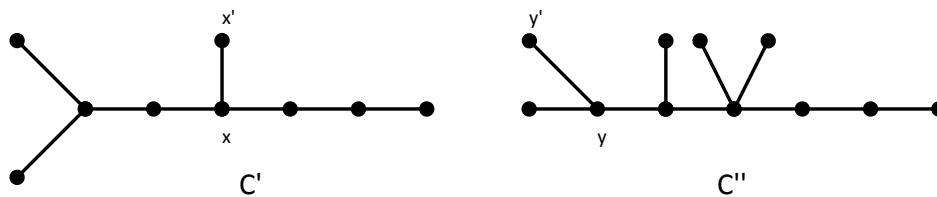
Before stating Dulio and Pannone's third result, we shall define the concept of *fusing* two vertices in order to construct a new tree from two given trees.

Definition 5.1 Consider two trees T_1 and T_2 in which vertex u_1 is in T_1 while vertex u_2 is in T_2 . The tree T obtained by **fusing** u_1 from T_1 and u_2 from T_2 is obtained by identifying vertices u_1 and u_2 which results in a vertex u . This will be denoted as $T = T_1^{u_1} \cup T_2^{u_2}$.

Clearly the number of vertices in the tree T is equal to the sum of the vertices in T_1 and T_2 minus one while the number of edges in T is equal to the sum of edges in T_1 and T_2 .

Result 3 : Suppose that a new tree G' is constructed by fusing a pendant vertex x' from caterpillar C' with another pendant vertex y' from a caterpillar C'' , both of which caterpillars have the same diameter. Then, if vertices x and y in C' and C'' respectively, are adjacent to these pendant vertices and have the same path-vectors, then, after fusion, they would still have equal path-vectors.

Example 5.3 Figure 10 shows two caterpillars C' and C'' in which $\text{ecc}(x) < \text{ecc}(y)$ and $\mathbf{p}_{C'}(x) = \mathbf{p}_{C'}(y) = (3, 5, 7, 7, 3, 2)$


 Figure 10: The two caterpillars C' and C'' .

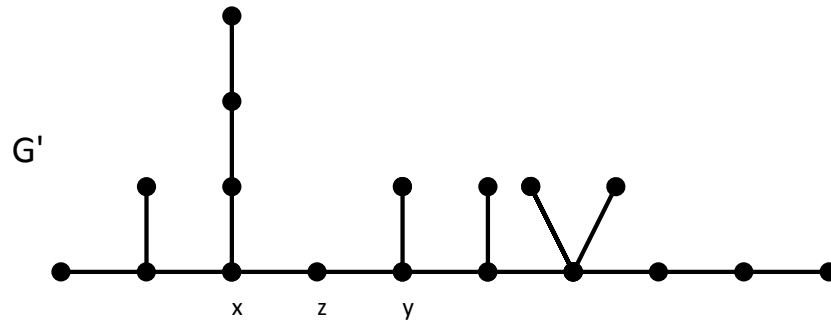


Figure 11: Fusing two pendant vertices x' and y' to obtain a tree G' in which the vertices x' and y' become a single vertex z in G' and where $\mathbf{p}_{G'}(x) = \mathbf{p}_{G'}(y) = (3, 6, 11, 15, 17, 19, 15, 15, 5, 3)$.

Question 5.1 *Can the results obtained for vertices be repeated for edges? That is, given two prescribed eccentricities, can we find two caterpillars of the same diameter and two edges (one in each caterpillar) with the given eccentricities, having the same path-vector? And, if there exist such caterpillars, then by using the notion of either joining or fusing, can we get a single tree in which the two edges **still** have the same path-vectors (even though possibly modified) ?*

We shall now investigate these questions. But first we need to define the *eccentricity of an edge* in a graph G .

Definition 5.2 *Let G be a finite graph of order n . Then the eccentricity of the edge e denoted by $\text{ecc}(e)$ is the minimum eccentricity of the two adjacent vertices forming the edge e .*

Consider Example 5.2 in which we select edge e in C_1 and edge e^* in C_2 where they have equal eccentricities and $\mathbf{p}_{C_1}(e) = \mathbf{p}_{C_2}(e^*) = (1, 3, 4, 4, 3, 1)$. When the resultant tree is obtained by **joining** vertices x and y , their eccentricities become different and $\mathbf{p}_T(e) = (1, 3, 5, 7, 7, 5, 3, 1, 0, 0)$ while $\mathbf{p}_T(e^*) = (1, 4, 8, 12, 15, 15, 12, 8, 4, 1)$ so that $\mathbf{p}_T(e) \neq \mathbf{p}_T(e^*)$. So we can conclude that Dulio and Pannone's second result does not hold for the edge case.

Example 5.4 *Figure 12 shows two caterpillars G_1 and G_2 in which $\text{ecc}(e) = \text{ecc}(e')$ and $\mathbf{p}_{G_1}(e) = \mathbf{p}_{G_2}(e') = (1, 3, 4, 4, 3, 1)$*

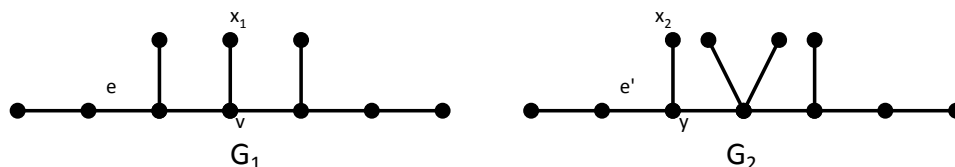
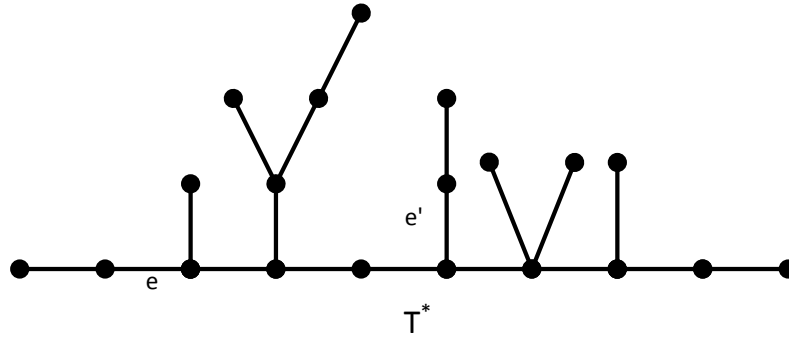


Figure 12: The two caterpillars G_1 and G_2 .


 Figure 13: The tree T^* .

Path length	1	2	3	4	5	6	7	8	9
$\mathbf{p}(e)$	1	3	4	5	6	8	6	3	1
$\mathbf{p}(e')$	1	3	6	7	9	7	2	0	0

 Table 8 : Path-vectors for edges e and e' in tree T^* .

As in the case of joining two vertices, when **fusing** two pendant vertices to obtain a tree T^* as shown in Figure 13, the path-vectors of edges e and e' in T^* are not equal, even though they were initially equal to each other before the fusion. This also contrasts with the third result of Dulio and Pannone obtained when considering vertex path-vectors.

The following theorem gives a necessary condition in order to obtain two equal edge path-vectors after fusing two vertices of two given trees.

Theorem 5.2 Consider two trees T_1 and T_2 and two edges u_1v_1 in T_1 and u_2v_2 in T_2 . Let $T = T_1^{u_1} \cup T_2^{u_2}$. Assume that for every l , $s_l^-(v_1)$ in T_1 is equal to $s_l^-(v_2)$ in T_2 . Then for the two adjacent edges uv_1 and uv_2 in the tree T , $\mathbf{p}(u, v_1) = \mathbf{p}(u, v_2)$.

Proof. As $p_l(x, y)$, the number of paths of length l passing through edge xy is given by

$$\sum_{i=0}^n s_i^-(x) s_{n-j}^+(y)$$

for $l \geq 1$ and $n = l - 1$,

then the number of edges passing through edge uv_1 in T

$$p_l(u, v_1) = \sum_{i=0}^n s_i^-(v_1) s_{n-j}^+(u)$$

while the number of edges passing through edge uv_2 in T

$$p_l(u, v_2) = \sum_{i=0}^n s_i^-(v_2) s_{n-j}^+(u)$$

But from assumption, $s_l^-(v_1) = s_l^-(v_2)$ for any path length l in T , so that $\mathbf{p}(u, v_1) = \mathbf{p}(u, v_2)$ \square

So far, we have proved that for a tree with odd diameter, the central edge and a non-central edge do not have the same path-vector. But Figures 14 and 15 show that, even in the case of caterpillars, two edges with different eccentricities can have equal edge path-vectors. This result is similar to the vertex case as can be seen from Examples 4.1 and 5.1. We believe that this can happen only in cases similar to those shown in Figures 14 and 15, that is, if two edges of different eccentricities in a caterpillar have the same path-vectors, then they must lie on opposite sides of the central vertex (or edge), their eccentricities must differ by 1, and the edge with the smaller eccentricity is incident to the (or to a) central vertex.

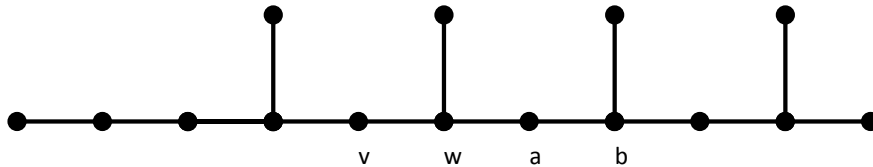


Figure 14: The caterpillar $C(1, 0, 1, 0, 1, 0, 1, 0, 2)$ of even diameter; $\mathbf{p}(vw) = \mathbf{p}(ab) = (1, 3, 5, 8, 8, 10, 7, 7, 3, 2)$ and $\text{ecc}(vw) \neq \text{ecc}(ab)$.

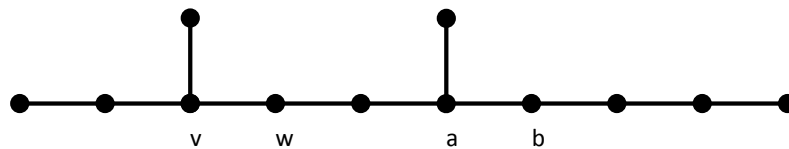


Figure 15: The caterpillar $C(1, 1, 0, 0, 1, 0, 0, 1)$ of odd diameter; $\mathbf{p}(vw) = \mathbf{p}(ab) = (1, 3, 4, 5, 6, 5, 4, 3, 1)$ and $\text{ecc}(vw) \neq \text{ecc}(ab)$.

In [17], Pannone proposed the following conjecture for general trees.

Conjecture 5.1 *If three vertices (respectively edges) of a tree T have mutually distinct eccentricities, then they cannot have the same path-vectors.*

6 Reconstruction of trees from path-tables

We finally consider the reconstruction of a tree given its path-table. As we have seen from various counterexamples, in general the information obtained from a path-table is not sufficient to reconstruct its tree. We may therefore ask the following question: What do two trees with the same path-table have in common? In other words by looking at a path-table (vertex or edge) of a tree, ‘how much’ of the tree can we draw?

Although in general, we will not be able to draw the complete tree, by studying carefully the *rows* of the given path-table which represent the path-vectors of vertices (or edges), and making suitable calculations, we can recognise the type of vertices (or edges) to which some of the rows in the path-table refer.

Now if v is a vertex of a tree T we often refer to the row corresponding to it in the path-table as the row \mathbf{v} ; the i^{th} entry of row \mathbf{v} is denoted by $\mathbf{v}[i]$.

Analogous notations hold for an edge e of a tree T .

(1) *The vertices (edges) of the trunk:*

By definition trunk vertices or edges form part of all the longest paths in a tree so that in the path-table they correspond to those rows with the largest entries in the last column. Obviously the row of the central vertices/edges do have maximal entries in the last column but these need not be the only ones. In fact, it is known that in general, we cannot tell which row (or rows) represent the centre since there can be two identical rows representing both a central and a non-central vertex as can be confirmed by Example 5.1.

On the other hand, Theorem 5.1 shows that the path-row of the central edge of a tree of odd diameter in an edge path-table is unique.

(2) *Peripheral vertices (edges):*

A vertex is *peripheral* if it has degree one and there is a path of length $\text{diam } T$ passing through it. The edge incident to the peripheral vertex is called a *peripheral edge*.

Peripheral vertices (edges) are easily recognisable from the path-tables. Their rows must have a 1 in the first entry and a non-zero last entry (i.e. counting paths of diametrical length). All peripheral vertices (edges) have the same row-sum of $n - 1 = m$ (no. of edges) and can be split into two subsets: left and right peripherals. The choice of left and right is arbitrary; what is important is to recognise how peripherals are divided.

Suppose h is the number of right peripheral vertices and k is the number of left peripheral vertices, then h and k can be found by solving the equations $h + k = p$ where p is the total number of peripheral vertices in T and $hk = l$ where l is the largest entry in their last column of the path-table.

If the entries h or k appear more than required then we only choose those vertices(edges) having a row-sum of $n - 1 = m$ (no. of edges). (NOTE: the path-vector of an end-edge in an edge path-table is equal to the corresponding endvertex in a vertex path-table).

(3) *Endvertices(edges) which are not peripheral:*

These are rows which have a 1 in the first entry and a zero last entry. They all have the same row-sum of $n - 1 = m$ (no. of edges).

Since we cannot reconstruct trees in general, a natural question to consider is whether we can find a class of trees for which equality of vertex (or edge) path-tables implies isomorphism. A natural choice can be the class of (i) caterpillars and (ii) restricted thin trees (RTT).

It is important to point out that Dulio and Pannone's example described in Section 3 excludes the possibility of studying the reconstructibility of general thin trees. Their example, shown in Figure 2, gives a pair of thin trees on 20 vertices that are non-isomorphic and have the same path-table.

6.1 Path-tables of caterpillars

Recall that a tree T is a caterpillar if, removing every endvertex of T results in a path, called the *spine* of the caterpillar.

Theorem 6.1 *If a caterpillar C and a tree T have the same vertex (edge) path-table, then $C \simeq T$.*

Proof. Let v_1, v_2, \dots, v_n be the vertices of the trunk of the tree T in the order in which they appear on the trunk. By definition, a tree is a caterpillar if and only if every vertex not belonging to the trunk has degree 1. Since the caterpillar C and the tree T have the same vertex path-table then T is also a caterpillar which can be completely determined once we manage to uniquely order the set of vertices of the trunk. This can be done by looking at the row of any peripheral vertex v of T in which the i^{th} term ($i \geq 2$) of its row is equal to the $\deg(v_{i-1}) - 1$ where v_{i-1} is a vertex of the trunk.

If the caterpillar C and the tree T have the same edge path-table then T is also a caterpillar which is completely determined once we manage to uniquely order the set of edges of the trunk. Then we can proceed as in the vertex case. \square

6.2 Path-tables of restricted thin trees

As described in section 4.3, a *restricted thin tree* (RTT) is a thin tree all of whose trunk vertices have degree at most 3. Since we need to study RTTs we have to define a *near endvertex* which is a vertex all of whose neighbours with at most one exception are endvertices. Moreover a *trunk near endvertex* is a near endvertex whose only neighbour which is not an endvertex is a trunk vertex.

The strategy used to reconstruct an RTT is to first show that they are recognisable from path-tables and then proceed to reconstruct them.

For the first part of our strategy we need to show that, from the path-table of a tree T , we can determine when these properties hold:

- (i) every trunk vertex has degree at most 3;
- (ii) every vertex which is not a trunk vertex is either an endvertex or a trunk near endvertex.

Observation 6.1 : *The vertex v of T lies on a longest path if and only if $v[d] > 0$.*

Throughout, d will denote, as usual, the diameter of T which is determined by the position of the last column in any path-table of T .

Observation 6.2 : *The vertex v of T is a trunk vertex if and only if $v[d]$ is maximal amongst all entries in the d^{th} column of the path-table.*

Observation 6.3 : *The degree of v of T is equal to $v[1]$.*

It is therefore clear that we can determine from the vertex path-table whether or not every trunk vertex has degree at most 3. We therefore see that we can recognise from the vertex path-table whether or not T is a restricted thin tree.

Observation 6.4 : *Let T be a tree in which every trunk vertex has degree at most 3. If v is a trunk near endvertex, then $v[2] = 2 + \binom{\deg(v)}{2}$.*

Therefore if T is a restricted thin tree then, for every vertex v which is not an endvertex nor a trunk vertex, $v[2]$ satisfies the above equality.

Observation 6.5 : *Suppose that T is not a thin tree. Then there is some vertex w which is not an endvertex nor a trunk vertex such that $w[2] > 2 + \binom{\deg(w)}{2}$.*

Proof. Since T is not a thin tree, there is a trunk vertex x adjacent to a vertex w not on the trunk such that w is not a near endvertex, as in Figure 16.

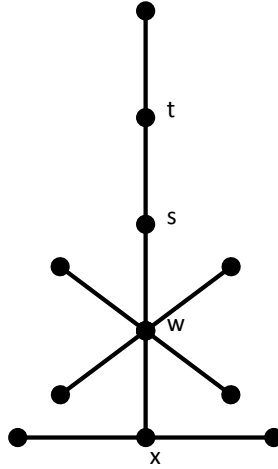


Figure 16: A trunk vertex x adjacent to vertex w which is not a near endvertex.

Note that, ignoring any paths containing vertex t , there are certainly at least $2 + \binom{\deg(w)}{2}$ paths of length 2 containing w . But then, the path wst shows that $w[2] \geq 1 + 2 + \binom{\deg(w)}{2}$ as required. \square

From the last two observations it follows that, for a tree whose trunk vertices all have degree at most 3 we can recognise from its path-table whether or not it is a restricted thin tree. We have therefore proved the following theorem.

Theorem 6.2 *Restricted thin trees are recognisable from their path-tables.*

For the second part of our strategy we need to show that given the path table of T and knowing that T is an RTT, we need to prove that the table determines T uniquely.

Observation 6.6 : *A vertex v is a peripheral vertex if and only if $v[1] = 1$ and $v[d] > 0$.*

Therefore, the number p of peripheral vertices of T can be determined from the path-table. Let v be the row corresponding to a peripheral vertex, let $h = v[d]$ and $k = p - h$. Then h and k are the numbers of peripheral endvertices at the two ends of T , respectively. We therefore have the following observation.

Observation 6.7 : *The total number p of peripheral vertices of T and the numbers h, k of left and right, respectively, peripheral endvertices of T can all be determined from the path-table. Also, we can determine from a row of an endvertex v whether v is peripheral or not.*

Let S be the union of all longest paths in T . We call S the *skeleton* of T . The first step will be to determine the skeleton of T . Note that the skeleton contains all peripheral vertices and all non-endvertices which lie on a longest path. Since we can recognise from the path table all such vertices and since we know the number of peripheral vertices at each end of T , we practically have S , except that we need to know how the peripheral

vertices are distributed at each end. For example, Figure 17 shows two situations where there are five peripheral vertices at one end.

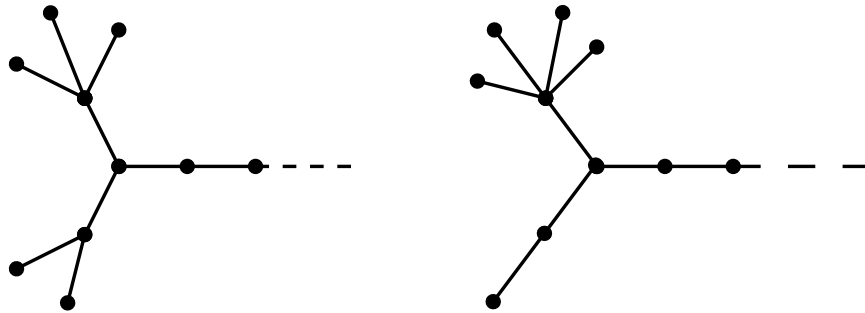


Figure 17: Two different cases with five peripheral vertices at one end of the skeleton.

Let us call the subtree consisting of k peripheral vertices adjacent to a common neighbour a k -peripheral cluster. Therefore, in Figure 17, we see k -peripheral clusters for $k = 3, 2, 4$ and 1 .

Note that, since T is an RTT, therefore no trunk vertex can have degree greater than 3, these facts are obvious:

- (i) If there is only one peripheral cluster at an end, then $k \leq 2$.
- (ii) There can be at most two peripheral clusters at one end.

Also, suppose we identify a row corresponding to a peripheral vertex v . Then $v[2] = k$ if and only if v is in a k -peripheral clusters in the skeleton S .

Since we know the number of peripheral vertices at each end, it follows from these considerations that the distribution of peripheral vertices into structures at each end can be determined. Since we know the diameter d of T , we have the skeleton S . Also, we can identify two row vectors \mathbf{a} , \mathbf{b} from the path table corresponding to two peripheral vertices coming from different ends of T . If the two ends of T are identical then we can let $\mathbf{a} = \mathbf{b}$; if they are different we can associate correctly the two vectors \mathbf{a} , \mathbf{b} , with the appropriate end.

We now need to determine which vertices in S have degree 3 and, for these vertices v that do, the number of endvertices incident to the third neighbour of v . We call this the *configuration* at v , and it can therefore be one of the types shown in Figure 18.

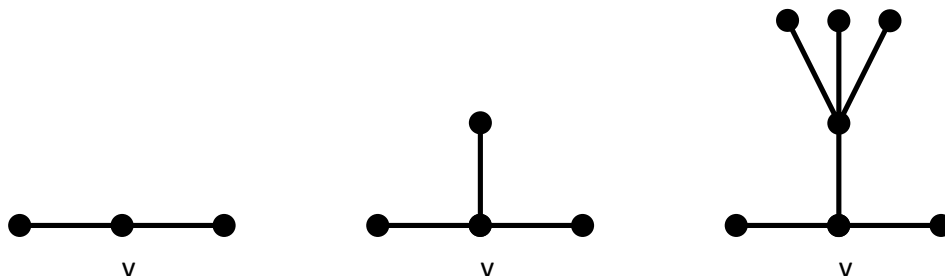


Figure 18: The types of configuration at v .

We shall first assume that both ends of S have only one k -peripheral cluster. In this case we label the non-peripheral vertices in S as v_2, v_3, \dots, v_d in the order in which they

appear in S . Suppose the peripheral vertex corresponding to the row vector \mathbf{a} corresponds to the end of S starting from v_2 . Then $\mathbf{a} = (1, a_2, a_3, \dots, a_{d-1}, a_d)$ and therefore a_i equals the number of paths of length i passing through the peripheral vertex. We shall associate a_i with v_i in our proof below. Also, let us reverse the order of the elements of \mathbf{b} so that $\mathbf{b} = (b_2, b_3, \dots, b_d, 1)$. Therefore, b_i equals the number of paths of length $d - i + 2$ passing through the peripheral vertex corresponding to \mathbf{b} . Again, we shall associate b_i with the vertex v_i of S .

Our claim is the following: knowing the skeleton S , the vectors \mathbf{a} , \mathbf{b} as above, and the configurations at $v_2, v_3, \dots, v_j (j \geq 3)$ one can determine the configuration at v_{j+1} .

Note that the configuration at v_3 is easily determined. If $a_3 = 1$ then $\deg(v_3) = 2$. Otherwise, $\deg(v_3) = 3$ and v_3 is incident to an endvertex (otherwise we would have more than one peripheral cluster at the end corresponding to v_2). Also, we know the configuration at v_2 since we know exactly the number of peripheral vertices at this end. Therefore, proving the above claim would show that the RTT T is reconstructible.

So, suppose we know the configurations up to v_j . Figure 19 shows the configurations at v_{j-1}, v_j, v_{j+1} when their degrees are 3.

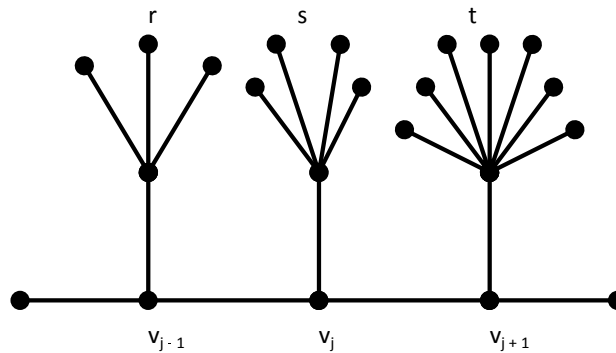


Figure 19: Configurations at v_{j-1}, v_j and v_{j+1} .

Therefore, referring to this figure, we know r and $s (\geq 0)$ if they exist and we need to determine whether $\deg(v_{j+1}) = 2$ or 3 and, in the latter case, the value of $t (\geq 0)$. (We are considering here only the case when $\deg(v_{j-1}) = \deg(v_j) = 3$; the case when either or both of these degrees equals 2 is easier to deal with).

Now, $a_{j+1} = s + \deg(v_{j+1}) - 1$, therefore $\deg(v_{j+1})$ can be determined. If $\deg(v_{j+1}) = 3$ then, $b_j = t + \deg(v_j) - 1$, therefore t can be determined. Hence one claim is verified.

In the case when we have two peripheral clusters at the end corresponding to v_3 we can proceed similarly. Suppose we have an h_1 -peripheral cluster and an h_2 -peripheral cluster, as illustrated in Figure 20, and the peripheral vertex corresponding to \mathbf{a} in the h_1 -peripheral cluster.

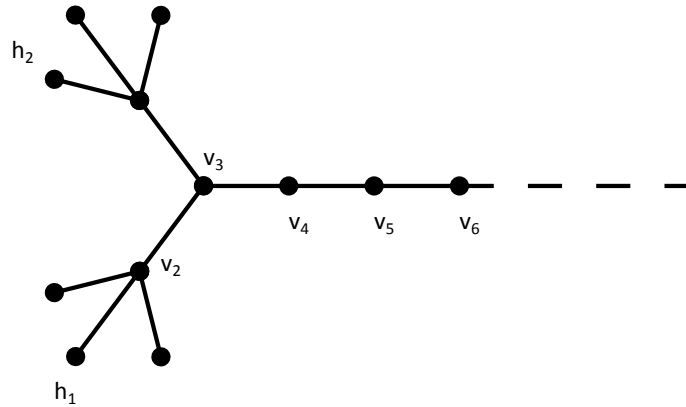
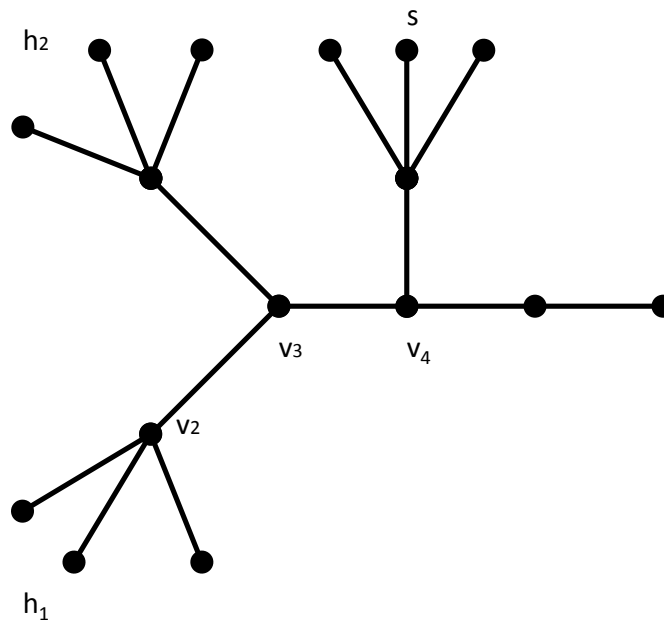


Figure 20: Two peripheral clusters at one end.

Firstly, note that $a_4 = \deg(v_4) - 1 + h_2$, therefore $\deg(v_4)$ is determined. If $\deg(v_4) = 3$ then we have the situation shown in Figure 21, and we need to determine s .

Figure 21: The case when $\deg(v_4) = 3$.

But $b_3 = 2 + s$ and therefore s can be determined. We therefore know the configuration at v_4 and we can proceed as before. Thus we have proved the following theorem.

Theorem 6.3 *Restricted thin trees are reconstructible from their path-tables.*

7 Conclusion

Let us now take stock of the situation on path-tables of trees in light of the results we have presented. We hereby list in table form statements which are either proved or disproved in our survey and we also offer some open problems.

Statement	Status	Remarks
1. A tree is reconstructible from its vertex path-table.	False	A counter-example can be found in Section 3 . Figure 2 gives an example of two non-isomorphic trees U_1 and U_2 having the same path-tables as shown in Table 3.
2. A tree is reconstructible from its edge path-table.	False	The same counter-example as in Section 3 . Here we have to consider Table 5.
3. Trees having equal vertex and edge path-tables are isomorphic so that the two tables together reconstruct a tree. T	False	Slater's trees T_1 and T_2 and Dulio/Pannone's trees U_1 and U_2 are two counter-examples given in Section 3 .
4. The edge path-table gives vertex path-table.	False	In Section 3 , Slater's pairs of trees T_1 and T_2 together with their path-tables shown in Tables 2 and 4 gives a counterexample.
5. The vertex path-table gives edge path-table for arbitrary tree.	Open Problem	
6. Trees having the same vertex path-table necessarily have the same Randić-table.	Open Problem	Slater example-pair on 18 vertices described in Section 3 shows that trees having the same Randić-table need not have the same vertex path-table.
7. $\mathbf{p}(x,y)$ can be determined from $\mathbf{p}(x)$ and $\mathbf{p}(y)$ for arbitrary trees.	It is valid only up to length 4	See Section 4.2 .
8. $\mathbf{p}(x,y)$ can be determined from $\mathbf{p}(x)$ and $\mathbf{p}(y)$ for caterpillars.	True	See Proposition 4.1 in Section 4.3 .
9. $\mathbf{p}(x,y)$ can be determined from $\mathbf{p}(x)$ and $\mathbf{p}(y)$ for thin or restricted thin trees.	Open problem	See Section 4.3 .
10. The path-vector of the central vertex is unique.	False	Dulio and Pannone give four sufficient conditions for uniqueness described in Section 5 .

Statement	Status	Remarks
11. The path-vector of the central edge is unique.	True	See Theorem 5.1 in Section 5 .
12. In an arbitrary tree, two vertices with different eccentricities have different path-vectors.	False	A counter-example for a tree with odd diameter is Example 4.1 in Section 4.2 while the counter-example for an even diameter is Example 5.1 in Section 5 .
13. Two edges in a caterpillar with different eccentricities have different path-vectors. (Statement 11 would be a special case).	False	Two counter-examples are shown in Figures 14 and 15 in section 5.1 .
14. A caterpillar is reconstructible from its vertex(edge) path-table.	True	See Theorem 6.1 in Section 6.1 .
15. A restricted thin tree is reconstructible from its vertex path-table.	True	See Theorem 6.3 in Section 6.2 . Theorem 6.3 can be adapted to prove that an RTT is also reconstructible from its edge path-table.
16. Three vertices of mutually different eccentricities cannot have the same path-vector.	Open Problem	See Conjecture 5.1 in Section 5.1 .
17. Five consecutive vertices on a path cannot have the same path-vector.	Open Problem	Example 4.1 in Section 4 shows that there exist four consecutive vertices on a path having the same vertex path-vectors.
18. Find a way to construct non-isomorphic trees with the same vertex and edge path-table. Such trees exist (see Item 3 in this table).	Open Problem	Dulio and Pannone suggest that such pairs can be constructed by attaching a number of branches A_1, A_2, \dots, A_m to a “hub” (a tree with vertices h_1, h_2, \dots, h_m) in two different ways but an exact method has not yet been found.

References

- [1] A.T. Balaban, G.S.Bloom and L.V.Quintas. unpublished result, 1982.
- [2] G.S.Bloom, J.W.Kennedy and L.V.Quintas. Some problems concerning distance and path degree sequences. *Graph Theory*, Lagów 1981, Lecture Notes in Mathematics, Vol 1018, Springer-Verlag, Berlin: 179-190, 1983.

- [3] H. Chung. Investigation of Graphs having identical path layer matrices. Research Science Inst at Massachusetts Inst. of Technology, 2004.
- [4] A.A. Dobrynin. Graphs having the same path layer matrix (in Russian), *Vychislitel'nye Sistemy* 119 : 13-33, 1987.
- [5] A.A. Dobrynin. Regular graphs having the same path layer matrix. *J.Graph Theory* 14 : 141-148, 1990.
- [6] A.A. Dobrynin. Path enumerations and similarity of graphs. *Graph Theory Notes NY* 23 : 7-9, 1992
- [7] A.A. Dobrynin. Cubic graphs with 62 vertices having the same path layer matrix. *J.Graph Theory* 17 : 1-4, 1993.
- [8] A.A. Dobrynin. Cubic graphs without cut-vertices having the same path layer matrix. *J.Graph Theory* 38 : 177-182, 2001.
- [9] A.A. Dobrynin and L.S.Mel'nikov. Recent results in the theory of path layer matrix. *Graph Theory Notes of New York XLIII*, New York Academy of Sciences : 48-56, 2006.
- [10] P.Dulio and V.Pannone. The Converse of Kelly's lemma and control-classes in graph reconstruction. Acta Univ. Palacki. Olomuc. Fac. rer. nat., *Mathematica* 44 : 25-38, 2005.
- [11] P.Dulio and V.Pannone. Trees with the same path-table. *Le Matematiche*, Catania, 60, no.1 : 59-65, 2005.
- [12] P.Dulio and V.Pannone. Trees with path-stable center. *Ars Combinatoria, LXXX* : 153-175, 2006.
- [13] P.Dulio and V.Pannone. Joining caterpillars and stability of the tree center. *Discrete Maths*, 308(7) : 1185-1190, 2008.
- [14] F.Y.Halberstam and L.V.Quintas. A note on tables of distance and path degree sequences for cubic graphs, Presented at Silver Jubilee Conference on Combinatorics, University of Waterloo, Waterloo, Ontario, Canada, June 14-July 2, 1982.
- [15] F.Y.Halberstam and L.V.Quintas. Distance and path degree sequences for cubic graphs, Mathematical Department, Pace University, New York, 1982.
- [16] F.Y.Halberstam and J.Zak. A note on planarity and distance degree sequences. *Lecture Notes Math.* 1018 : 286-289, 1983.
- [17] V. Pannone, Personal communication
- [18] L.V.Quintas and P.J.Slater, Pairs of non-isomorphic graphs having the same path degree sequences. *MATCH* 12 : 75-86, 1981.
- [19] M.Randić. Characterizations of atoms, molecules and classes of molecules based on path enumerations. *MATCH* 7 : 5 - 64, 1979.

- [20] M. Randić. Design of molecules with desired properties. In *Concepts and Application of Molecular Similarity*. Eds : M.A. Johnson and G.M.Maggiora, John Wiley and Sons, New York : 77-145, 1990.
- [21] M. Randić, G.A.Kraus and B. Džonova-Jerman-Blažić. Ordering of graphs as an approach to structure-activity studies. In: *Chemical Applications of Topology and Graph Theory* (a collection of papers from a symposium held at the University of Georgia, Athens, 1983), Elsevier, New York :192-205, 1983.
- [22] M.Randić and C.L.Wilkins, Graph theoretical approach to recognition of structural similarity in molecules. *J. Chem. Inf. Comput. Sci.* 19 : 23-31, 1979.
- [23] Skorobogatov V.A. and A.A. Dobrynin. Metric Analysis of Graphs. *MATCH* 23 : 105 - 151, 1988
- [24] P.J.Slater. Counterexamples to Randić's conjecture on distance degree sequences for trees.*J. Graph Theory* 6 : 89-91, 1982
- [25] Y.Yuansheng, L.Xiaohui, C.Zhiqiang and L. Weiming. 4-regular graphs without cut-vertices having the same path layer matrix. *J.Graph Theory* 44 : 304-310, 2003.
- [26] Y.Yuansheng, L.Jianhua and W.Chunli. Small regular graphs having the same path layer matrix. *J.Graph Theory* 39 : 219-221, 2002.