

Randić index and lexicographic order

Oswaldo Araujo and Juan Rada

Departamento de Matemáticas, Facultad de Ciencias, Universidad de Los Andes, 5101 Mérida, Venezuela
E-mail: {araujo; juanrada}@ciens.ula.ve

Received 8 March 2000

Let T be a tree and consider the Randić index $\chi(T) = \sum_{v_i-v_j} (1/\sqrt{\delta(v_i)\delta(v_j)})$, where v_i-v_j runs over all edges of T and $\delta(v_i)$ denotes the degree of the vertex v_i . Using counting arguments we show that the Randić index, is monotone increasing over the well (lexicographic order) ordered sequence of trees with unique branched vertex.

KEY WORDS: Randić index, m -tree, lexicographic order

1. Introduction and terminology

Molecular topology determines a large number of molecular properties, including not only those depending on molecular size and shape such as boiling points, molecular volume, solubilities, refractive indices, etc., but also the quantum mechanical characteristics of molecules, such as energy levels, electronic populations, etc., which depends, essentially, on the connectivity of the atoms [15].

Therefore, it would be of great interest to have a quantitative measure reflecting the essential features of a given structure. Such measures are usually called topological indices. These indices in some way reflect not only the size and shape of a molecule but also their connectivity, i.e. the way their atoms are linked. Among the numerous topological indices considered in chemical graph theory, only a few have been found noteworthy in practical applications [12]. One of these is the connectivity or Randić index χ , conceived by Randić to reflect the amount of branching present in a chemical species [14]. A complete list of scientific publications concerned with physical–chemical properties and pharmacology applications of χ includes several hundreds of papers, a few review articles and two books [10,11].

Recently, some results about the mathematical properties of χ have been published [1,2,7–9,13]. In particular, in [5] the authors demonstrated that among n -vertex trees, the path graph has maximal χ and Bollobás and Erdős [4] proved that among all connected graphs, the star has minimal Randić-value.

Let T be a tree with set of vertices $\{v_1, \dots, v_n\}$, the Randić index of T is defined as

$$\chi(T) = \sum_{v_i-v_j} \frac{1}{\sqrt{\delta(v_i)\delta(v_j)}},$$

where v_i-v_j runs over all the edges of T and $\delta(v_i)$ denotes the degree of the vertex v_i . It is easy to see that for $n \geq 3$, $\chi(S_n) = \sqrt{n-1}$ and $\chi(P_n) = (n-3)/2 + \sqrt{2}$, where S_n is the star and P_n is the path of n vertices. By our previous comments, we have the following inequality:

$$\chi(S_n) \leq \chi(T) \leq \chi(P_n).$$

On the other hand, the adjacency matrix of T is the $n \times n$ matrix A whose entries a_{ij} are given by

$$a_{ij} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

and the characteristic polynomial of A

$$\det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + \dots + c_{n-1}\lambda + c_n,$$

where $c_{2r+1} = 0$ for every $r \in \mathbb{N}$ and $c_2 = -(n-1)$ [3], is referred as the characteristic polynomial of T . The coefficients of this polynomial are used in [6] to give a (lexicographic) well-order “ \leq ” over the set of all trees with n vertices as follows:

$$T = T' \iff c_i(T) = c_i(T') \text{ for all } i$$

and

$$T < T' \iff \text{there exists } r \in \mathbb{N} \text{ such that } |c_r(T)| < |c_r(T')| \text{ and } c_i(T) = c_i(T') \text{ for } i < r.$$

It is well known that for every tree T with n vertices, $S_n \leq T \leq P_n$. So we have the following correspondence:

$$\begin{array}{ccccc} S_n & \leq & T & \leq & P_n \\ \updownarrow & & \updownarrow & & \updownarrow \\ \chi(S_n) & \leq & \chi(T) & \leq & \chi(P_n) \end{array}$$

A natural question arises: *is χ monotone increasing over the well (lexicographic) ordered sequence of trees with n vertices?* That is, if $\{T_i\}_{i=1}^k$ is the sequence of all trees with n vertices such that

$$S_n \leq T_1 \leq \dots \leq T_k \leq P_n,$$

then

$$\chi(S_n) \leq \chi(T_1) \leq \dots \leq \chi(T_k) \leq \chi(P_n)?$$

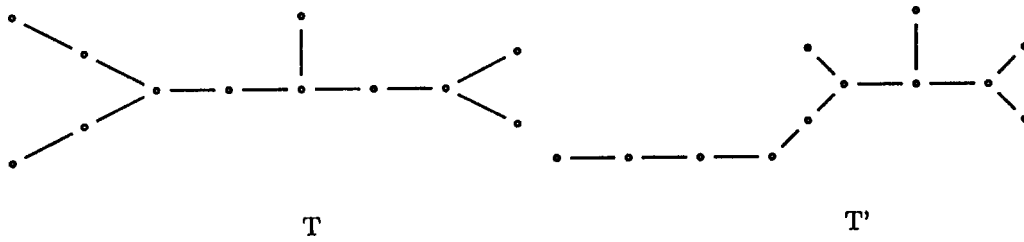


Figure 1.

Although this is true for $n = 1, \dots, 7$, in general, the answer is negative. For example, in figure 1 $T < T'$ but $\chi(T) = 5.5957 > 5.5914 = \chi(T')$. However, as we shall see, χ is increasing for certain classes of trees.

An i -vertex is a vertex of degree i , $k_i(T)$ is the number of i -vertices and for $i, j \in \mathbb{N}$, $e_{ij}(T)$ denotes the number of edges that joins a i -vertex with a j -vertex. Finally, we denote by $[v_i, v_j]$ the shortest walk joining v_i with v_j and recall that the number of edges in $[v_i, v_j]$ is called the distance in T between v_i and v_j and is denoted by $d(v_i, v_j)$.

We consider trees (with n vertices) which have exactly one branched vertex, that is, a unique vertex of degree greater than 2. If this unique vertex has degree m , then, for short, we say that T is an m -tree. In theorems 2.2 and 2.4 we use counting arguments to determine the coefficients c_4 and c_6 of the characteristic polynomial of an m -tree T , in terms of the degree structure of T . It becomes clear from here that these coefficients depends on n, m and $e_{1m}(T)$. It turns out that the Randić index of T can also be expressed in terms of these numbers (theorem 3.1), allowing us to establish the relationship between the coefficients of the characteristic polynomial and the Randić index of an m -tree. As a consequence, we prove in theorem 3.4 that χ is monotone increasing over the set of all m -trees ($2 \leq m \in \mathbb{N}$) with a fixed number of vertices.

2. First coefficients of the characteristic polynomial associated to an m -tree

Given an integer $m > 2$, an m -tree T has the form presented in figure 2, where v is a vertex of degree m . Our first goal is to find $c_4(T)$ and $c_6(T)$ in terms of the degree structure of T . We need the following result [3, p. 49]:

Proposition 2.1. If T is a tree with n vertices then the odd coefficients $c_{2r+1}(T)$ are zero, and the even coefficients $c_{2r}(T)$ are given by the rule that $(-1)^r c_{2r}$ is the number of ways of choosing r disjoint edges in the tree T .

An easy counting argument using this result shows that for every integer $k > 1$, $c_4(P_k) = (k - 3)(k - 2)/2$. Therefore, given an m -tree T as in figure 2, we can find the number of ways of choosing two disjoint edges in each “maximal path” $[v_i, v_j]_{i < j}$. Note that there are exactly $(m - 1)m/2$ maximal paths. Now, if we take two disjoint edges in T then there are two possibilities (see figure 3): the first, that the two edges are separated by the vertex v . In this case, both edges only belong to the maximal path $[v_i, v_j]$. The

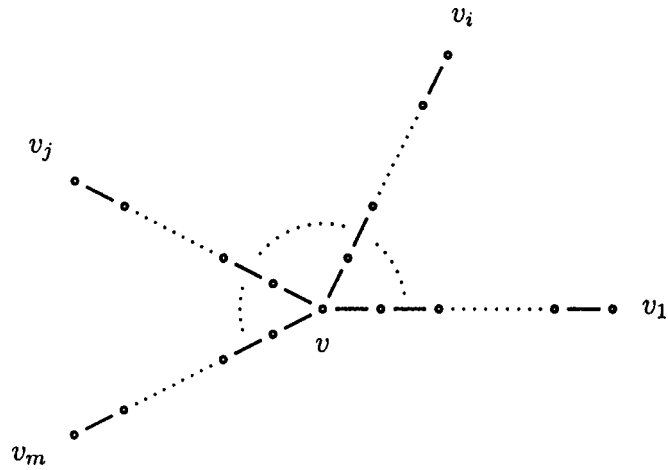


Figure 2.

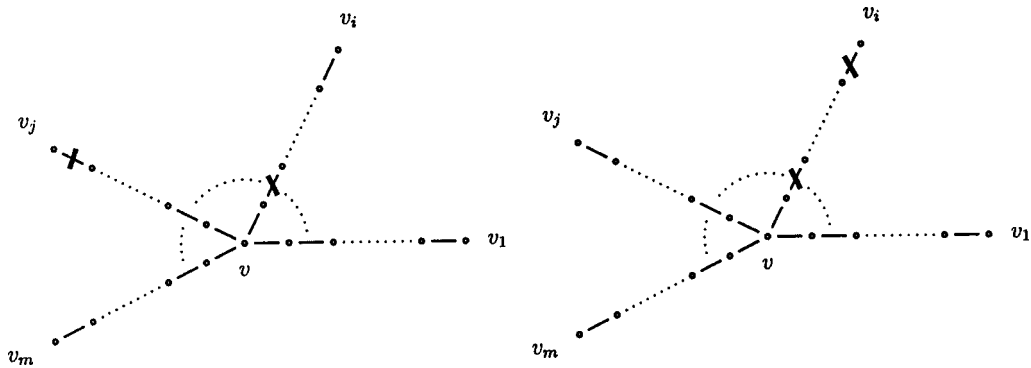


Figure 3.

second possibility is, of course, that both edges are not separated by v . Hence, these two edges belong to the $m - 1$ maximal paths $[v_i, v_j]_{i \neq j \in \{1, \dots, m\}}$.

Theorem 2.2. If T is an m -tree with n vertices then

$$c_4(T) = \frac{1}{2}[(n - 1)^2 - 3(n - 1) + m(3 - m)].$$

Proof. Set $a_i = d(v, v_i)$ for each $i = 1, \dots, m$. Then by proposition 2.1 and the previous observation we know that

$$c_4(T) = \sum_{i < j} c_4(P_{a_i+a_j+1}) - (m - 2) \sum_{i=1}^m c_4(P_{a_i+1}).$$

Now, since for every integer $k > 1$, $c_4(P_k) = (k - 3)(k - 2)/2$ and $\sum_{i=1}^m a_i = n - 1$ it follows

$$\begin{aligned}
 2c_4(T) &= (m - 1) \sum_{i=1}^m a_i^2 + 2 \sum_{i < j} a_i a_j - 3(m - 1) \sum_{i=1}^m a_i + \frac{(m - 1)(m - 2)}{2} \cdot 2 \\
 &\quad - (m - 2) \left[\sum_{i=1}^m a_i^2 - 3 \sum_{i=1}^m a_i + 2m \right] \\
 &= \sum_{i=1}^m a_i^2 + 2 \sum_{i < j} a_i a_j - 3 \sum_{i=1}^m a_i + m(3 - m) \\
 &= \left(\sum_{i=1}^m a_i \right)^2 - 3 \sum_{i=1}^m a_i + m(3 - m) \\
 &= (n - 1)^2 - 3(n - 1) + m(3 - m). \quad \square
 \end{aligned}$$

Corollary 2.3. If T and T' are m -trees with n vertices then $c_2(T) = c_2(T')$ and $c_4(T) = c_4(T')$.

Proof. $c_2(T) = c_2(T')$ since both trees have the same number of vertices and $c_4(T) = c_4(T')$ is immediate consequence of theorem 2.2. \square

In order to find $c_6(T)$ for an m -tree T as in figure 2, we need to count all possible ways of choosing three disjoint edges in T . This can be done by fixing three disjoint edges in T , then two situations can occur (see figure 4): (a) all three edges belong to a same maximal path $[v_i, v_j]$ or (b) three disjoint edges do not belong to a same maximal path.

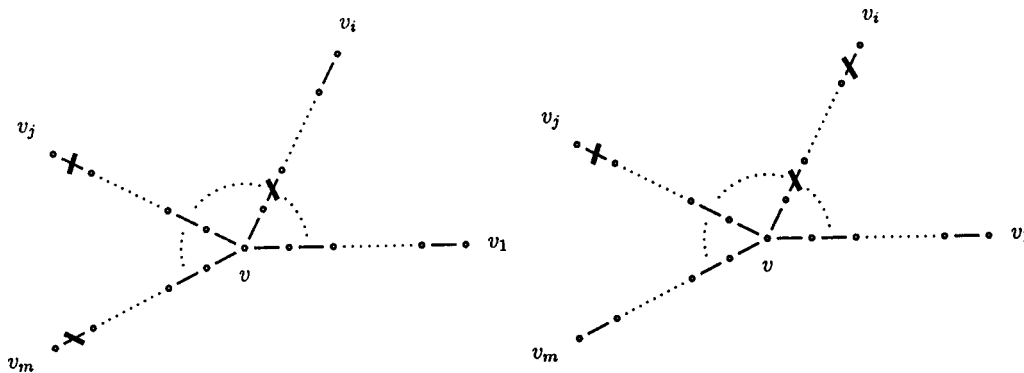


Figure 4.

Note that there are exactly

$$\sum_{i < k} c_6(P_{a_i+a_k+1}) - (m-2) \sum_{i=1}^m c_6(P_{a_i+1})$$

ways of choosing three disjoint edges in case (a) while in case (b), there are

$$\sum_{i < k < l} (a_i - 1)(a_k - 1)(a_l - 1)$$

when we consider edges apart from the center vertex v and

$$(m-2) \sum_{i < k} (a_i - 1)(a_k - 1)$$

when one of the edges has v as an endvertex.

Now, an easy application of proposition 2.1 shows that $-6c_6(P_k) = (k-5)(k-4)(k-3)$ for $k > 2$, consequently we have the following result:

Theorem 2.4. If T is an m -tree with n vertices then

$$c_6(T) = -\frac{1}{6} \alpha + (m-2)e_{1m}(T),$$

where α is given by

$$\begin{aligned} \alpha = & n^3 - 12n^2 + [47 - 3(m-2)^2 - 3(m-2)]n \\ & + 2(m-2)^3 + 21(m-2)^2 + 19(m-2) - 60. \end{aligned}$$

Proof. Set $a_i = d(v, v_i)$ for each $i = 1, \dots, m$. Then by our previous comments

$$\begin{aligned} -c_6(T) = & \sum_{i < k} c_6(P_{a_i+a_k+1}) - (m-2) \sum_{i=1}^m c_6(P_{a_i+1}) \\ & + \sum_{i < k < l} (a_i - 1)(a_k - 1)(a_l - 1) + (m-2) \sum_{i < k} (a_i - 1)(a_k - 1). \end{aligned}$$

Assume that $a_i = 1$ for $1 \leq i \leq j = e_{1m}(T)$.

$$\begin{aligned} -6 \sum_{i < k} c_6(P_{a_i+a_k+1}) = & (m-1) \sum_{i=j+1}^m a_i^3 + (m-1)j + 3 \sum_{i < k} (a_i a_k^2 + a_i^2 a_k) \\ & - 18 \sum_{i < k} a_i a_k - 9(m-1) \sum_{i=1}^m a_i^2 - 9(m-1)j \\ & + 26(m-1) \sum_{i=j+1}^m a_i + 26(m-1)j - 12m(m-1), \end{aligned}$$

$$\begin{aligned}
 -6 \sum_{i=1}^m c_6(P_{a_i+1}) &= \sum_{i=j+1}^m a_i^3 - 9 \sum_{i=j+1}^m a_i^2 + 26 \sum_{i=j+1}^m a_i - 24(m-j), \\
 \sum_{i < k < l} (a_i - 1)(a_k - 1)(a_l - 1) &= \sum_{j+1 \leq i < k < l} a_i a_k a_l - (m-j-2) \sum_{j+1 \leq i < k} a_i a_k \\
 &\quad + \frac{(m-j-2)(m-j-1)}{2} \sum_{i=j+1}^m a_i \\
 &\quad - \frac{1}{6}(m-j-2)(m-j-1)(m-j)
 \end{aligned}$$

and

$$\sum_{i < k} (a_i - 1)(a_k - 1) = \sum_{j+1 \leq i < k} a_i a_k - (m-j-1) \sum_{i=j+1}^m a_i + \frac{(m-j-2)(m-j)}{2}.$$

The result follows by replacing these equations in the first formula, grouping terms in a convenient way and bearing in mind that

$$\sum_{i=j+1}^m a_i + j = n - 1. \quad \square$$

Corollary 2.5. If T and T' are m -trees with n vertices then

$$c_6(T) - c_6(T') = (m-2)(e_{1m}(T) - e_{1m}(T')).$$

Proof. By theorem 2.4,

$$c_6(T) = -\frac{1}{6} \alpha + (m-2)e_{1m}(T)$$

and

$$c_6(T') = -\frac{1}{6} \alpha + (m-2)e_{1m}(T').$$

Hence,

$$c_6(T) - c_6(T') = (m-2)(e_{1m}(T) - e_{1m}(T')). \quad \square$$

3. Randić index of an m -tree

Given an m -tree T , we can find the Randić index $\chi(T)$ by looking at figure 5, where the numbers next to the edges indicate the product of the degrees of the adjacent vertices.

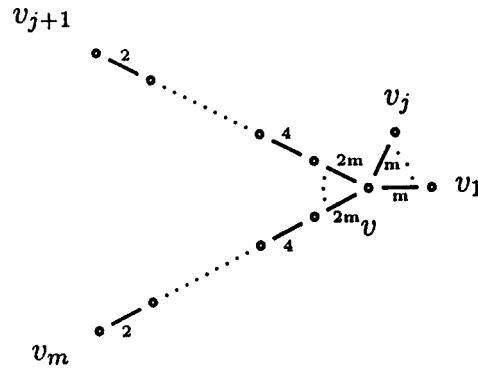


Figure 5.

Theorem 3.1. If T is an m -tree with n vertices then $\chi(T) = \lambda - \mu e_{1m}(T)$, where

$$\lambda = \frac{m}{\sqrt{2}} + \frac{m}{\sqrt{2m}} - m + \frac{1}{2}n - \frac{1}{2} \quad \text{and} \quad \mu = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2m}} - \frac{1}{\sqrt{m}} - \frac{1}{2}.$$

Proof. Keeping the notation of figure 5, set $a_i = d(v, v_i)$ for $i = 1, \dots, m$. Note that $a_i = 1$ for $i = 1, \dots, j = e_{1m}(T)$. Then

$$\begin{aligned} \chi(T) &= \sum_{i=1}^j \frac{1}{\sqrt{m}} + \sum_{i=j+1}^m \left(\frac{1}{\sqrt{2}} + \frac{1}{2}(a_i - 2) + \frac{1}{\sqrt{2m}} \right) \\ &= \frac{1}{\sqrt{m}}j + (m - j) \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2m}} - 1 \right) + \frac{1}{2}(n - j - 1) \\ &= \left(\frac{m}{\sqrt{2}} + \frac{m}{\sqrt{2m}} - m + \frac{1}{2}n - \frac{1}{2} \right) - \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2m}} - \frac{1}{\sqrt{m}} - \frac{1}{2} \right)j \\ &= \lambda - \mu e_{1m}(T). \end{aligned} \quad \square$$

Corollary 3.2. If T and T' are m -trees with n vertices then

- (1) $\chi(T) = \chi(T') \Leftrightarrow e_{1m}(T) = e_{1m}(T')$, and
- (2) $\chi(T) > \chi(T') \Leftrightarrow e_{1m}(T) < e_{1m}(T')$.

Proof. It follows from theorem 3.1, since $\chi(T) = \lambda - \mu e_{1m}(T)$, $\chi(T') = \lambda - \mu e_{1m}(T')$ and, clearly, $\mu > 0$. □

Proposition 3.3. Let T and T' be m -trees with n vertices. If $T \leq T'$ then $\chi(T) \leq \chi(T')$.

Proof. By corollary 2.3, $c_2(T) = c_2(T')$ and $c_4(T) = c_4(T')$. If $T \leq T'$ then we have two possibilities: either $c_6(T) = c_6(T')$ or $c_6(T) > c_6(T')$. In the first case it follows from corollary 2.5 that $e_{1m}(T) = e_{1m}(T')$. But then, by the first part of

corollary 3.2, $\chi(T) = \chi(T')$. In the latter case, again by corollary 2.5, $e_{1m}(T) > e_{1m}(T')$. Consequently, the second part of corollary 3.2 gives $\chi(T) < \chi(T')$. \square

We denote by Ω_n the set of all m -trees with n vertices, varying $2 \leq m \in \mathbb{N}$.

Theorem 3.4. The Randić index is monotone increasing over Ω_n .

Proof. Suppose that T is an m -tree, T' is an m' -tree with n vertices and $T \leq T'$. Clearly, $c_2(T) = c_2(T')$ since T and T' have the same number of vertices. Since $c_4(T) \leq c_4(T')$ then, by theorem 2.2, $m \geq m'$. Assume that

$$T_1 \leq T_2 \leq \dots \leq T_k$$

is the sequence (lexicographically ordered) of all m -trees with n vertices, then, by proposition 3.3, the Randić index of this sequence is monotone increasing with maximal Randić index $\chi(T_k) = \lambda - \mu e_{1m}(T_k)$. Hence,

$$\chi(T_k) \leq \lambda = \frac{m}{\sqrt{2}} + \frac{m}{\sqrt{2m}} - m + \frac{1}{2}n - \frac{1}{2},$$

where equality holds if and only if $e_{1m}(T_k) = 0$. Now, consider the sequence

$$U_1 \leq U_2 \leq \dots \leq U_l$$

of all m' -trees with n vertices. Again, by proposition 3.3, the minimal value of the Randić index in this sequence is

$$\chi(U_1) = \frac{m'}{\sqrt{2}} + \frac{m'}{\sqrt{2m'}} - m' + \frac{1}{2}n - \frac{1}{2} - \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2m'}} - \frac{1}{\sqrt{m'}} - \frac{1}{2} \right) (m' - 1).$$

It is easy to check that $\chi(U_1) \geq \lambda$. Consequently, $\chi(T) \leq \lambda \leq \chi(U_1) \leq \chi(T')$. \square

Example 3.5. Table 1 shows all m -trees with 9 vertices. The order in which these trees appear is the lexicographic order. First of all we note that the χ -values are monotone increasing as we should expect from theorem 3.4. Moreover, trees 9 and 10, or 14 and 15 or 16 and 17 have the same Randić index, this is a consequence of the first part of corollary 3.2, since each of these pairs have the same number of edges joining the 1-vertices with the m -vertex. Also we observe that, for example, all 4-trees have the same coefficient c_4 , this fact can be deduced by corollary 2.3, and among these trees, 9 and 10 have the same coefficient c_6 which is perfectly reasonable since they have the same number of edges joining the 1-vertices with the m -vertex (corollary 2.5). Furthermore, we can predict by corollary 2.5 the difference from the coefficient c_6 of trees 8 and 12 by just looking at the figure: $(4 - 2) \cdot (3 - 0) = 6$.

Table 1

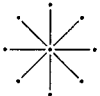
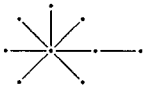
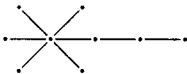

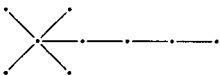
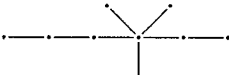




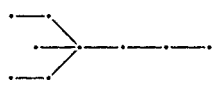
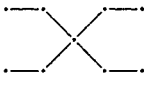
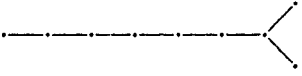
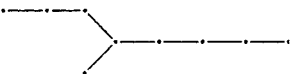
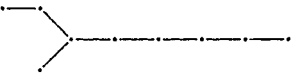
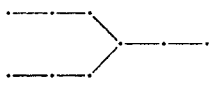
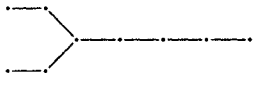
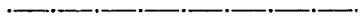
N°	Tree	m	$ c_4 $	$ c_6 $	e_{1m}	χ
1		8	0	0	8	2.8284
2		7	6	0	6	3.2421
3		6	11	0	5	3.5370
4		6	11	4	4	3.6245
5		5	15	4	4	3.8121
6		5	15	7	3	3.8883
7		5	15	10	2	3.9644
8		4	18	10	3	4.0606
9		4	18	12	2	4.1213

Table 1
(Continued.)

N°	Tree	m	$ c_4 $	$ c_6 $	e_{1m}	χ
10		4	18	12	2	4.1213
11		4	18	14	1	4.1819
12		4	18	16	0	4.2426
13		3	20	16	2	4.2700
14		3	20	17	1	4.3080
15		3	20	17	1	4.3080
16		3	20	18	0	4.3460
17		3	20	18	0	4.3460
18		2	21	20	2	4.4142

Acknowledgements

We are grateful to our colleague J.L. Chacón for his help in the construction of the figures.

References

- [1] O. Araujo and D.A. Morales, A theorem about the algebraic structure underlying orthogonal graph invariants, *J. Chem. Inf. Comput. Sci.* 36 (1996) 1051–1053.
- [2] O. Araujo and J.A. de la Peña, The connectivity index of a weighted graph, *Linear Algebra Appl.* 283 (1998) 171–177.
- [3] N. Biggs, *Algebraic Graph Theory* (Cambridge University Press, Cambridge, 1993).
- [4] B. Bollobás and P. Erdős, Graphs of extremal weights, *Ars Combin.* 50 (1998) 225–233.
- [5] G. Caporossi, I. Gutman and P. Hansen, Variable neighborhood search for extremal graphs. 4. Chemical trees with extremal connectivity index, to appear in *Comput. Chem.*
- [6] D.M. Cvetkovic, M. Doob and H. Sachs, *Spectra of Graphs, Theory and Applications* (Academic Press, New York, 1979).
- [7] S. Fajtlowicz, Written on the Wall, Conjectures derived on the basis of the program Galatea Gabriela Graffiti, University of Houston (1987).
- [8] I. Gutman, O. Miljkovic, G. Caporossi and P. Hansen, Alkanes with small and large Randić connectivity index, *Chem. Phys. Lett.* 306 (1999) 366–372.
- [9] I. Gutman, O. Araujo and D.A. Morales, Bounds for the Randić connectivity index, to appear in *J. Chem. Inf. Comput. Sci.*
- [10] L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research* (Academic Press, New York, 1976).
- [11] L.B. Kier and L.H. Hall, *Molecular Connectivity in Structure-Activity Analysis* (Wiley, New York, 1986).
- [12] Z. Mihalic and N. Trinajstić, A graph theoretical approach to structure-property relationship, *J. Chem. Educ.* 69 (1992) 701–712.
- [13] D.A. Morales and O. Araujo, The relationship between the connectivity index and the number and types of carbon atoms in a structure, *J. Mol. Struct. (Theochem)* 417 (1997) 241–246.
- [14] M. Randić, On characterization of molecular branching, *J. Am. Chem. Soc.* 97 (1975) 6609–6615.
- [15] D.H. Rouvray, Predicting chemistry from topology, *Sci. Amer.* 255 (1986) 40–47.