communications in mathematical
match
and in computer chemistry

no. 41, March 2000

ISSN 0340-6253

MATCDY (41) 57-70 (2000)

MOLECULES WITH SMALLEST CONNECTIVITY INDICES

Ivan Gutman and **Olga Miljković**

Faculty of Science, University of Kragujevac, P. O. Box 60, YU-34000 Kragujevac, Yugoslavia

(Received July 1999)

Abstract

Extending an earlier study [21], in which chemical trees with smallest, second-smallest and third-smallest connectivity indices were determined, we now solve the analogous problem for general molecular graphs, possessing n vertices and m edges, $n-1 \leq m \leq 2n$, representing hydrocarbons without any multiple bonds.

Dedicated to Professor Milan Randić
on the occasion of the 25th anniversary
of his invention of the connectivity index

1. INTRODUCTION

In 1975 Randić [1] introduced a novel graph invariant, aimed at being a measure of what intuitively is considered as the branching of the carbon-atom skeleton of an organic molecule. He called this invariant *the branching index*, but later it was renamed into *connectivity index* [2, 3] or *Randić index* [4, 5], and is usually denoted by χ. Randić himself [1] established correlations between χ and scores of physico-chemical properties of alkanes: boiling point, the Kováts chromatographic index,

enthalpy of formation, parameters in the Antoine equation for vapor pressure, surface area, solubility in water. Randić's article [1] was followed by almost countless QSPR and QSAR applications, making χ one of the most popular molecular structure-descriptors for predicting physico-chemical and, especially, pharmacologic properties of organic compounds [2, 3, 6]. For some of the most recent studies along these lines see [7]-[13].

Within classes of isomers, with increasing branching of the carbon-atom skeleton, the value of the connectivity index decreases. Thus, adopting Randić's original viewpoint, molecules possessing the smallest possible value of the connectivity index may be considered as being maximally branched. If so, then in this work we determine these maximally branched molecular systems.

The graph representation of the carbon-atom skeleton of an organic molecule (or more precisely: of a hydrocarbon without any multiple bonds) – the so-called molecular graph – is a connected graph whose all vertices have degrees four or less [4, 14]. If G is such a graph and $\delta(v)$ is the degree of its vertex v, then the connectivity index of G is defined as [1]

$$\chi = \chi(G) := \sum_{uv} \frac{1}{\sqrt{\delta(u)\,\delta(v)}}$$

with the summation ranging over all pairs of adjacent vertices, that is over all edges of the graph G.

Needless to say the above definition is not restricted to molecular graphs, but is applicable to all graphs, both connected and disconnected.

For some time the connectivity index is attracting the attention of mathematicians, who established a few of its fundamental mathematical properties. Fajtlowicz [15] and Araujo & de la Peña [16] characterized the n-vertex graphs with greatest χ-value: These are the graphs without isolated vertices in which each component is a regular graph; for all such graphs $\chi = n/2$. Bollobás & Erdős [17] proved that among n-vertex graphs without isolated vertices, the star has the smallest χ-value, equal to $\sqrt{n-1}$. Hansen et al. showed [18] that among n-vertex trees the path graph has greatest χ-value, equal to $(n-3)/2 + \sqrt{2}$, $n \geq 3$. Several lower and upper bounds for the connectivity index were reported [16, 19, 20].

In a recent work [18] a linear–programming method has been developed by means of which acyclic molecular graphs (with a fixed number of vertices), with smallest and greatest connectivity indices could be characterized. A proper extension of this method enabled also the characterization of acyclic molecular graphs with the second–smallest, third–smallest, greatest, second–greatest and third–greatest χ [21].

In this work we continue these studies for multi–cyclic graphs (with up to $n/2$ independent cycles), and for the n-vertex unicyclic case find molecular graphs with the three smallest χ-values, hence the three most branched families of molecular graphs.

2. THE METHOD

In what follows the molecular graph considered will be denoted by G, and the number of its vertices and edges by n and m, respectively. Sometimes we say that G is an (n,m)-molecular graph.

Denote by x_{ij} the number of edges uv of G, for which $\delta(u) = i$ and $\delta(v) = j$. Then

$$\chi(G) = \sum_{1 \le i \le j \le 4} \frac{x_{ij}}{\sqrt{i\,j}} \,. \tag{1}$$

Note that $x_{11} = 0$ whenever $n \ge 3$, and therefore the case $i = j = 1$ needs not be considered any further. Consequently, the right–hand side of Eq. (1) is a linear function of the following nine variables:

$$x_{12}\,,\, x_{13}\,,\, x_{14}\,,\, x_{22}\,,\, x_{23}\,,\, x_{24}\,,\, x_{33}\,,\, x_{34}\,,\, x_{44}\,.$$

Denote by n_i the number of vertices of G having degree i, $i = 1,2,3,4$. Then the following "book–keeping" relations are obeyed:

$$n_1 + n_2 + n_3 + n_4 = n \tag{2}$$

$$x_{12} + x_{13} + x_{14} = n_1 \tag{3}$$

$$x_{12} + 2\,x_{22} + x_{23} + x_{24} = 2\,n_2 \tag{4}$$

$$x_{13} + x_{23} + 2\,x_{33} + x_{34} = 3\,n_3 \tag{5}$$

$$x_{14} + x_{24} + x_{34} + 2\,x_{44} = 4\,n_4 \tag{6}$$

and, in addition to them:

$$n_1 + 2n_2 + 3n_3 + 4n_4 = 2m \, . \tag{7}$$

Recall that the left-hand side of Eq. (7) is the sum of the vertex degrees, known to be equal to twice the number of edges [4, 14].

Relations (2)–(7) are linearly independent.

Assuming that the parameters n and m are fixed, (2)–(7) may be understood as a system of six linear equations in thirteen unknowns: n_1, n_2, n_3, n_4, x_{12}, x_{13}, x_{14}, x_{22}, x_{23}, x_{24}, x_{33}, x_{34} and x_{44}. We solve them for the unknowns n_1, n_2, n_3, n_4 and two of the x_{ij}'s, say x_{pq} and x_{rs}. Then x_{pq} and x_{rs} are expressed as linear functions of n, m and the remaining seven x_{ij}'s. By substituting x_{pq} and x_{rs} back into Eq. (1) we then arrive at an expression for $\chi(G)$ in terms of n, m and the remaining seven parameters x_{ij}.

By trial-end-error it was found that the system (2)–(7) should be solved in the unknowns n_1, n_2, n_3, n_4, x_{14} and x_{44}. If so, then

$$x_{14} = \frac{4n - 2m}{3} - \frac{4}{3}x_{12} - \frac{10}{9}x_{13} - \frac{2}{3}x_{22} - \frac{4}{9}x_{23} - \frac{1}{3}x_{24} - \frac{2}{9}x_{33} - \frac{1}{9}x_{34} \tag{8}$$

and

$$x_{44} = \frac{5m - 4n}{3} + \frac{1}{3}x_{12} + \frac{1}{9}x_{13} - \frac{1}{3}x_{22} - \frac{5}{9}x_{23} - \frac{2}{3}x_{24} - \frac{7}{9}x_{33} - \frac{8}{9}x_{34} \tag{9}$$

which combined with Eq. (1) yields

$$\begin{aligned} \chi(G) &= \frac{4n+m}{12} + \left(\frac{1}{\sqrt{2}} - \frac{7}{12}\right)x_{12} + \left(\frac{1}{\sqrt{3}} - \frac{19}{36}\right)x_{13} \\ &+ \frac{1}{12}x_{22} + \left(\frac{1}{\sqrt{6}} - \frac{13}{36}\right)x_{23} + \left(\frac{1}{2\sqrt{2}} - \frac{1}{3}\right)x_{24} \\ &+ \frac{1}{36}x_{33} + \left(\frac{1}{2\sqrt{3}} - \frac{5}{18}\right)x_{34} \\ &\approx \frac{4n+m}{12} + 0.12377\,x_{12} + 0.04957\,x_{13} + 0.08333\,x_{22} \\ &+ 0.04714\,x_{23} + 0.02022\,x_{24} + 0.02778\,x_{33} + 0.01090\,x_{34} \, . \end{aligned} \tag{10}$$

Because all multipliers on the right-hand side of (10) are positive-valued, this expression is convenient for deducing the conditions that molecular graphs with small χ must obey.

3. MOLECULAR GRAPHS WITH SMALL CONNECTIVITY INDICES

Our considerations (which basically are a kind of integer–linear–programming reasoning) are based on the fact that all multipliers on the right–hand side of (10) are positive–valued. Define therefore an auxiliary quantity Δ as:

$$\begin{aligned} \Delta \ &:= \ \chi(G) - \frac{4n+m}{12} \\ &\approx \ 0.12377\, x_{12} + 0.04957\, x_{13} + 0.08333\, x_{22} + 0.04714\, x_{23} \\ &+ \ 0.02022\, x_{24} + 0.02778\, x_{33} + 0.01090\, x_{34} \end{aligned} \tag{11}$$

and note that it cannot be negative–valued.

Now, $\chi(G)$ will attain its smallest, second–smallest, third–smallest, ..., values if Δ is equal to zero or is as close to zero as possible. This will be achieved if the parameters x_{ij}, occurring on the right–hand side of (11) have non–negative integer values, close to zero. In addition, these parameters must be chosen in a "graphical" manner, namely so that there exist graphs pertaining to them [21]. Therefore, in order to minimize Δ, n_2 and n_3 are required to be equal to zero or as close to zero as possible.

For our considerations it is crucial that Δ, Eq. (11), depends on x_{12}, x_{13}, x_{22}, x_{23}, x_{24}, x_{33} and x_{34}, but does not (explicitly) depend on n and m. As a consequence, the below argument, already used in [21] in the case of chemical trees, $m = n - 1$, is equally applicable to unicyclic ($m = n$), bicyclic ($m = n + 1$), tricyclic ($m = n + 2$), etc., molecular graphs.

From Eq. (11) it is seen that if $n_2 = 3$ and $n_3 = 0$ then Δ is at least $6 \times 0.02022 = 0.12132$. If $n_2 = 2$ and $n_3 = 1$ then Δ is at least $4 \times 0.02022 + 3 \times 0.01090 = 0.11358$. If $n_2 = 1$ and $n_3 = 2$ then Δ is at least $2 \times 0.02022 + 6 \times 0.01090 = 0.10584$. If $n_2 = 0$ and $n_3 = 3$ then Δ is at least $9 \times 0.01090 = 0.09810$. In summary, if $n_2 + n_3 = 3$ then Δ cannot be less than 0.09810. Clearly, Δ will exceed the value 0.09810 also if $n_2 + n_3 > 3$.

We now consider the case when $n_2 + n_3 \leq 2$ and search for graphically feasible combinations of x_{12}, x_{13}, x_{22}, x_{23}, x_{24}, x_{33} and x_{34} for which Δ is less than 0.098.

There are exactly nine such combinations, given in Table 1.

n_2	n_3	non-zero x_{ij}'s	Δ
0	0		0
0	1	$x_{34} = 3$	0.03270
1	0	$x_{24} = 2$	0.04044
0	2	$x_{34} = 6$	0.06540
0	1	$x_{34} = 2$, $x_{13} = 1$	0.07137
0	2	$x_{34} = 4$, $x_{33} = 1$	0.07137
1	1	$x_{24} = 2$, $x_{34} = 3$	0.07314
2	0	$x_{24} = 4$	0.08088
1	1	$x_{24} = 1$, $x_{34} = 2$, $x_{23} = 1$	0.08916

Table 1. Graphically feasible combinations of the parameters x_{12} , x_{13} , x_{22} , x_{23} , x_{24} , x_{33} and x_{34} for which $\Delta < 0.09810$; if $n_2 + n_3 \geq 3$ then $\Delta \geq 0.09810$

From Eqs. (2) and (7) we readily obtain

$$2(m+n) = 3\,n_1 + 4\,n_2 + 5\,n_3 + 6\,n_4 = 3(n_1 + n_2 + 2\,n_3 + 2\,n_4) + n_2 - n_3$$

which implies

$$2(n+m) \equiv n_2 - n_3 \pmod 3$$

and therefore

$$n + m \equiv n_3 - n_2 \pmod 3 .$$

Thus, the congruence class modulo 3 to which $n + m$ belongs depends solely on the number of vertices of degree two and three. Bearing this in mind, from the data given in Table 1 we conclude that among the combinations for which $\Delta < 0.9810$, there are three with $n + m \equiv 0 \pmod 3$, three with $n + m \equiv 1 \pmod 3$, and three with $n + m \equiv 2 \pmod 3$. This means that the combinations given in Table 1 completely determine the (n, m)-molecular graphs with the smallest, second-smallest and third-smallest connectivity indices (except, possibly, for the first few values of n, see below).

Using the data given in Table 1 we arrive at our main results:

Theorem 1. If n is sufficiently large (cf. Table 2), then for any value of m , $n-1 \leq m \leq 2n$, the following is true.

(a) If $n+m \equiv 0 \pmod 3$, then among (n,m)-molecular graphs, the graphs without vertices of degree two and three (that is, the graphs possessing only vertices of degree one and four) have the smallest connectivity indices, equal to $(4n+m)/12$.

(b) If $n+m \equiv 1 \pmod 3$, then among (n,m)-molecular graphs, the graphs without vertices of degree two and with a single vertex of degree three, adjacent to three vertices of degree four, have the smallest connectivity indices, equal to $(4n+m)/12+(3\sqrt{3}-5)/6$.

(c) If $n+m \equiv 2 \pmod 3$, then among (n,m)-molecular graph, the graphs without vertices of degree three and with a single vertex of degree two, adjacent to two vertices of degree four, have the smallest connectivity indices, equal to $(4n+m)/12+(3\sqrt{2}-4)/6$.

Theorem 2. If n is sufficiently large (cf. Table 2), then for any value of m , $n-1 \leq m \leq 2n$, the following is true.

(a) If $n+m \equiv 0 \pmod 3$, then among (n,m)-molecular graphs, the graphs with a single vertex of degree two, adjacent to two vertices of degree four, and a single vertex of degree three, adjacent to three vertices of degree four, have the second-smallest connectivity indices, equal to $(4n+m)/12+(\sqrt{2}+\sqrt{3}-3)/2$.

(b) If $n+m \equiv 1 \pmod 3$, then among (n,m)-molecular graphs, the graphs without vertices of degree two and with a single vertex of degree three, adjacent to two vertices of degree four and a vertex of degree one, have the second-smallest connectivity indices, equal to $(4n+m)/12+(8\sqrt{3}-13)/12$.

(c) If $n+m \equiv 2 \pmod 3$, then among (n,m)-molecular graphs, the graphs without vertices of degree two and with two vertices of degree three, each adjacent to three vertices of degree four, have the second-smallest connectivity indices, equal to $(4n+m)/12+(3\sqrt{3}-5)/3$.

Theorem 3. If n is sufficiently large (cf. Table 2), then for any value of m , $n-1 \leq m \leq 2n$, the following is true.

(a) If $n+m \equiv 0 \pmod 3$, then among (n,m)-molecular graphs, the graphs with a single vertex of degree two, adjacent to a vertex of degree four and the vertex of degree three, and a single vertex of degree three, adjacent to two vertices of degree

four and the vertex of degree two, have the third–smallest connectivity indices, equal to $(4n + m)/12 + (3\sqrt{2} + 4\sqrt{3} + 2\sqrt{6} - 15)/12$.

(b) If $n + m \equiv 1 \pmod 3$, then among (n, m)-molecular graphs, the graphs without vertices of degree three and with two vertices of degree two, each adjacent to two vertices of degree four, have the third–smallest connectivity indices, equal to $(4n + m)/12 + (3\sqrt{2} - 4)/3$.

(c) If $n + m \equiv 2 \pmod 3$, then among (n, m)-molecular graphs, the graphs without vertices of degree two and with two adjacent vertices of degree three, each adjacent to two vertices of degree four, have the third–smallest connectivity indices, equal to $(4n + m)/12 + (8\sqrt{3} - 13)/12$.

What "sufficiently large n" is depends of the value of m and the congruence class of $n + m$. In Table 2 are given the smallest values of n for which Theorems 1–3 hold in the case of acyclic, unicyclic, bicyclic and tricyclic molecular graphs; the acyclic case has earlier been reported [21].

		$n + m \equiv 0$	$n + m \equiv 1$	$n + m \equiv 2$
acyclic	$m = n - 1$	5, 17, 17	13, 13, 13	9, 21, 21
unicyclic	$m = n$	9, 12, 12	11, 11, 11	7, 16, 16
bicyclic	$m = n + 1$	10, 10, 10	9, 9, 9	8, 14, 14
tricyclic	$m = n + 2$	8, 8, 8	7, 10, 10	9, 15, 15

Table 2. The smallest number of vertices of (n, m)-molecular graphs for which Theorems 1, 2 and 3, respectively, are applicable

Anyway, Theorems 1–3 do not cover the first few values of n, because molecular graphs with properties specified in these theorems do not exist if n is not large enough. The finding of the respective "exceptional" graphs (with n-values smaller than what in Theorems 1–3 is specified as "sufficiently large") needs to be done either by using some special graph–theoretical reasoning or by brute–force search. A complete list of chemical trees with smallest, second–smallest and third–smallest connectivity indices has been reported elsewhere [21]. Here we show the analogous results for unicyclic molecular graphs.

In Figs. 1a and 1b are depicted unicyclic n-vertex molecular graphs with the smallest (left), second–smallest (center) and third–smallest (right) connectivity indices, $3 \leq n \leq 20$. Asterisks indicate the smallest graphs to which one of the Theorems 1–3 is applicable. In the general case the graphs characterized by Theorems 1–3 are not unique. In Figs. 1a and 1b only one representative for each case is depicted.

4. ON MOLECULAR GRAPHS WITH LARGE CONNECTIVITY INDICES

In order to obtain molecular graphs with large χ-values, we may pursue a procedure analogous to what was described in Section 2: solving the system (2)–(7) in the unknowns n_1, n_2, n_3, n_4, x_{12} and x_{22}. This results in

$$x_{12} = 2n - 2m - \frac{2}{3}x_{13} - \frac{1}{2}x_{14} + \frac{1}{3}x_{23} + \frac{1}{2}x_{24} + \frac{2}{3}x_{33} + \frac{5}{6}x_{34} + x_{44}$$

$$x_{22} = 3m - 2n - \frac{1}{3}x_{13} - \frac{1}{2}x_{14} - \frac{4}{3}x_{23} - \frac{3}{2}x_{24} - \frac{5}{3}x_{33} - \frac{11}{6}x_{34} - 2x_{44}$$

and

$$\begin{aligned}\chi(G) &= \frac{(2\sqrt{2}-2)n + (3-2\sqrt{2})m}{2} - \frac{1+2\sqrt{2}-2\sqrt{3}}{6}x_{13} - \frac{\sqrt{2}-1}{4}x_{14} \\ &- \frac{4-\sqrt{6}-\sqrt{2}}{6}x_{23} - \frac{3-2\sqrt{2}}{4}x_{24} - \frac{3-2\sqrt{2}}{6}x_{33} \\ &- \frac{11-5\sqrt{2}-2\sqrt{3}}{12}x_{34} - \frac{3-2\sqrt{2}}{4}x_{44} \qquad (12) \\ &\approx \frac{(2\sqrt{2}-2)n + (3-2\sqrt{2})m}{2} - 0.06072\,x_{13} - 0.10355\,x_{14} - 0.02272\,x_{23} \\ &- 0.04289\,x_{24} - 0.02860\,x_{33} - 0.03874\,x_{34} - 0.04289\,x_{44}\,.\end{aligned}$$

All multipliers on the right–hand side of (12) are negative–valued. Thus, one may expect that the molecular graphs with greatest χ-values are those for which all the parameters x_{13}, x_{14}, x_{23}, x_{24}, x_{33}, x_{34} and x_{44} are either zero or close to zero, that is for which n_3 and n_4 are either zero or close to zero. This, however, is true only for (n,m)-molecular graphs for which $m \approx n$. With increasing number of edges there

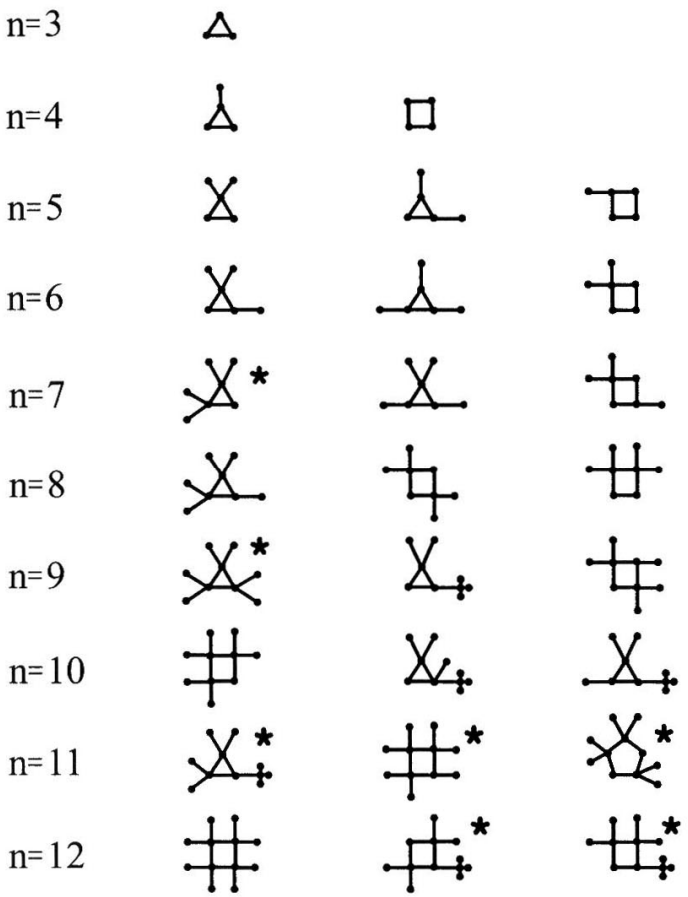

Fig. 1a

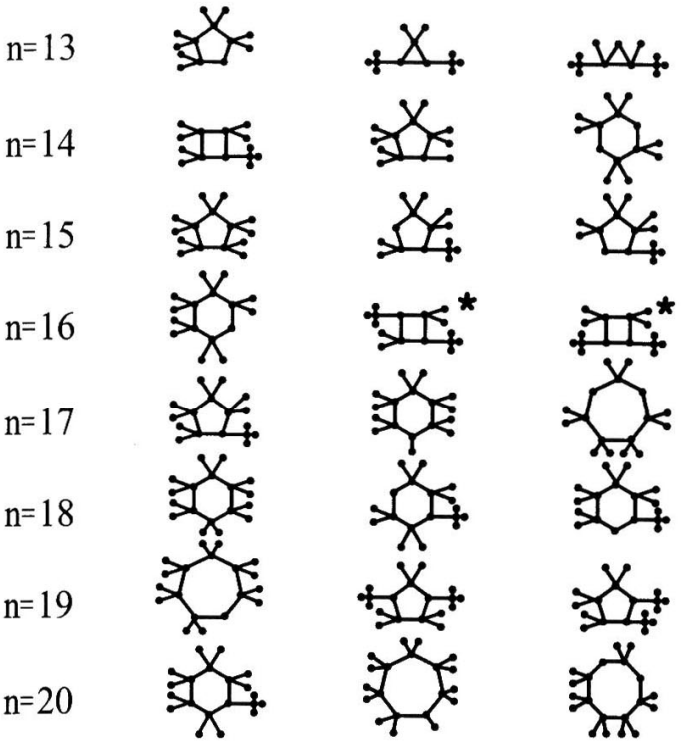

Fig. 1b

necessarily has to be more and more degree–three and degree–four vertices, making the analysis based on Eq. (12) difficult and/or infeasible.

Only for $m = n - 1$ (acyclic graphs) and $n = m$ (unicyclic graphs) Eq. (12) enables direct construction of molecular graphs with greatest connectivity indices. For acyclic graphs this has been reported earlier [21]. For monocyclic graphs we state the analogous results without a detailed proof.

Theorem 4. Among unicyclic n-vertex molecular graphs, $n \geq 3$, the cycle has greatest connectivity index, equal to $n/2$.

Theorem 5. Among unicyclic n-vertex molecular graphs, $n \geq 5$, the graphs without vertices of degree four, with a single vertex of degree three and a single vertex of degree one, that are not mutually adjacent, have the second–greatest connectivity indices, equal to $n/2 - (4 - \sqrt{2} - \sqrt{6})/2$.

Theorem 6. Among unicyclic n-vertex molecular graphs, $n \geq 5$, the graphs without vertices of degree four, with a single vertex of degree three and a single vertex of degree one, that are mutually adjacent, have the third–greatest connectivity indices, equal to $n/2 - (9 - 2\sqrt{3} - 2\sqrt{6})/6$.

Because cycles are regular graphs of degree two, Theorem 4 is just a special case of the more general (above mentioned) result of Fajtlowicz [15] and Araujo & de la Peña [16].

The structure of the graphs specified in Theorems 4–6 is shown in Fig. 2. The only "exceptional" case, not covered by Theorems 4–6 is the second–greatest χ-value for $n = 4$. This is immediately identified as methyl–cyclopropane; the respective graph is also shown in Fig. 2.

If $m > n$ then the characterization of (n, m)-molecular graphs possessing greatest connectivity indices cannot be accomplished by the presently elaborated method. Efforts are being made to approach this problem by other means.

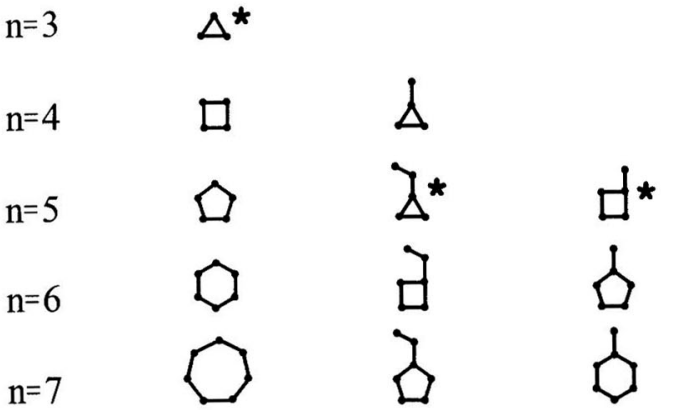

Fig. 2

Acknowledgement. The authors thank Professor Pierre Hansen (Montreal, Canada) for helpful discussions.

References

[1] M. Randić, *J. Am. Chem. Soc.* **97** (1975) 6609.

[2] L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976.

[3] L. B. Kier and L. H. Hall, *Molecular Connectivity in Structure–Activity Analysis*, Wiley, New York, 1986.

[4] N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, 1992.

[5] Z. Mihalić and N. Trinajstić, *J. Chem. Educ. 69* (1992) 701.

[6] S. P. Gupta, *Chem. Rev.* **91** (1991) 1109.

[7] E. Estrada, *J. Chem. Inf. Comput. Sci.* **35** (1995) 31.

[8] E. Estrada, *J. Chem. Inf. Comput. Sci.* **35** (1995) 1022.

[9] S. C. Basak and B. D. Gute, *SAR & QSAR Environ. Res.*, **7** (1997) 1.

[10] S. Nikolić, N. Trinajstić and I. Baučić, *J. Chem. Inf. Comput. Sci.* **38** (1998) 42.

[11] E. Estrada, N. Guevara and I. Gutman, *J. Chem. Inf. Comput. Sci.* **38** (1998) 428.

[12] E. Estrada, N. Guevara, I. Gutman and L. Rodriguez, *SAR & QSAR Environ. Res.* **9** (1998) 229.

[13] J. D. Gough and L. H. Hall, *J. Chem. Inf. Comput. Sci.* **39** (1999) 356.

[14] I. Gutman and O. E. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer–Verlag, Berlin, 1986.

[15] S. Fajtlowicz, *Written on the wall, Version 05–1998; a regularly updated file accessible from:* `siemion@math.uh.edu` .

[16] O. Araujo and J. E. de la Peña, *Lin. Algebra Appl.*, **283** (1998) 171.

[17] B. Bollobás and P. Erdős, *Ars Combin.*, **50** (1998) 225.

[18] G. Caporossi, I. Gutman and P. Hansen, *Comput. Chem.*, in press.

[19] D. A. Morales and O. Araujo, *J. Mol. Struct. (Theochem)*, **417** (1997) 241.

[20] O. Araujo O and J. E. de la Peña, *J. Chem. Inf. Comput. Sci*, **38** (1998) 827.

[21] I. Gutman, O. Miljković, G. Caporossi and P. Hansen, *Chem. Phys. Lett.*, **306** (1999) 366.