

Ordering of alkane isomers by means of connectivity indices

IVAN GUTMAN[#], DUŠICA VIDOVIĆ and ANKA NEDIĆ

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

(Received 3 September 2001)

The connectivity index of an organic molecule whose molecular graph is G is defined as $C(\lambda) = \sum (\delta_u \delta_v)^\lambda$, where δ_u is the degree of the vertex u in G , where the summation goes over all pairs of adjacent vertices of G and where λ is a pertinently chosen exponent. The usual value of λ is $-1/2$, in which case $\chi = C(-1/2)$ is referred to as the Randić index. The ordering of isomeric alkanes according to χ follows the extent of branching of the carbon-atom skeleton. We now study the ordering of the constitutional isomers of alkanes with 6 through 10 carbon atoms with respect to $C(\lambda)$ for various values of the parameter λ . This ordering significantly depends on λ . The difference between the orderings with respect to χ and with respect to $C(\lambda)$ is measured by a function Δ and the λ -dependence of Δ was established.

Keywords: connectivity index, Randić index, alkanes, isomer ordering.

INTRODUCTION

The *connectivity index* is one of the most popular molecular-graph-based structure-descriptors¹ and is defined as^{2,3}

$$C(\lambda) = C(\lambda; G) = \sum_{u,v} (\delta_u \delta_v)^\lambda \quad (1)$$

where δ_u denotes the degree (= number of first neighbors) of the vertex u of the molecular graph G , where the summation goes over all pairs of adjacent vertices of G and where λ is a pertinently chosen exponent. This structure-descriptor was introduced a quarter of century ago by Randić,² who chose $\lambda = -1/2$. However, in the chemical literature other choices of λ were also considered, in particular $\lambda = +1$ (e.g., in the papers^{4–6}), $\lambda = -1$ (e.g., in the papers^{2,3,7,8}). In several works^{9–12} the exponent λ was treated as an adjustable parameter, chosen so as to optimize the correlation between $C(\lambda)$ and some physico-chemical property of some selected class of organic compounds. More data and additional references on $C(\lambda)$ can be found elsewhere.^{3,8}

Anyway, the choice $\lambda = -1/2$ is certainly the most popular one and the most often employed.^{1,13,14} The respective structure-descriptor $\chi = C(-1/2)$ is referred to as the *Randić in-*

[#] Serbian Chemical Society active member.

index or *Randić connectivity index* or *molecular connectivity index* or simply *connectivity index*, although Randić himself² named it the *branching index*. It is commonly accepted that $\chi(G)$ is a measure of branching of the respective carbon-atom skeleton and that, consequently, the ordering of isomeric alkanes with respect to decreasing χ -values basically represents their ordering according to the increasing extent of branching. (More details on the concept of branching and on ordering of alkanes with respect to other topological indices can be found in a recent study by Balaban, Mills and Basak.¹⁵)

C-EQUIVALENT ALKANES

When ordering isomeric alkanes with regard to their connectivity indices one needs to take into account that there exist pairs of isomers whose $C(\lambda)$ -values coincide for all λ . To see this rewrite Eq. (1) as

$$C(\lambda) = \sum_{i < j} m_{ij} (i \cdot j)^\lambda$$

where m_{ij} denotes the number of edges of the graph G connecting a vertex of degree i with a vertex of degree j . For molecular graphs the above formula becomes:

$$C(\lambda) = m_{12}(1 \cdot 2)^\lambda + m_{13}(1 \cdot 3)^\lambda + m_{14}(1 \cdot 4)^\lambda + m_{22}(2 \cdot 2)^\lambda + m_{23}(2 \cdot 3)^\lambda + m_{24}(2 \cdot 4)^\lambda + m_{33}(3 \cdot 3)^\lambda + m_{34}(3 \cdot 4)^\lambda + m_{44}(4 \cdot 4)^\lambda$$

i.e.,

$$C(\lambda) = m_{12} 2^\lambda + m_{13} 3^\lambda + m_{14} 4^\lambda + m_{22} 4^\lambda + m_{23} 6^\lambda + m_{24} 8^\lambda + m_{33} 9^\lambda + m_{34} 12^\lambda + m_{44} 16^\lambda \quad (2)$$

From Eq. (2) it is evident that molecules for which all the nine parameters m_{12} , m_{13} , m_{14} , m_{22} , m_{23} , m_{24} , m_{33} , m_{34} , m_{44} are equal will have equal connectivity indices, independent of the value of the exponent λ . Such molecules are referred to as C -equivalent. Among alkanes there are numerous families of C -equivalent species, one of these is depicted in Fig. 1.

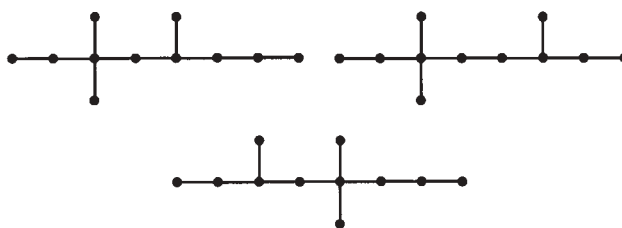


Fig. 1. Molecular graphs of three C -equivalent alkanes (3,3,5-trimethyloctane, 3,3,6-trimethyloctane and 3,5,5-trimethyloctane); their connectivity indices coincide for all values of the exponent λ because for all the three isomers $m_{12} = 2$, $m_{13} = 1$, $m_{14} = 2$, $m_{22} = 1$, $m_{23} = 2$, $m_{24} = 2$.

There are no C -equivalent alkanes with 7 or fewer carbon atoms, but beginning with octanes their number increases rapidly.

Within each family of C -equivalent alkanes the ordering with respect to $C(\lambda)$ is immaterial. In order to avoid this difficulty, we have considered sets of isomeric alkanes in which exactly one representative was taken from each C -equivalent family.

By the above described procedure, complete sets of isomeric alkanes with $n = 6, 7, 8, 9$ and 10 carbon atoms in which no two elements are C -equivalent were constructed. These sets consist of 5, 9, 16, 28 and 49 elements, respectively. (Recall that there are 5, 9, 18, 35 and 75 distinct constitutional isomers of C_nH_{2n+2} , $n = 6, 7, 8, 9, 10$.) In Fig. 2 are, for illustrative purposes, shown the molecular graphs of the nonane isomers examined in this work. They are numbered so that their order achieved for $\lambda = -1/2$ is 1, 2, 3, ..., 28.

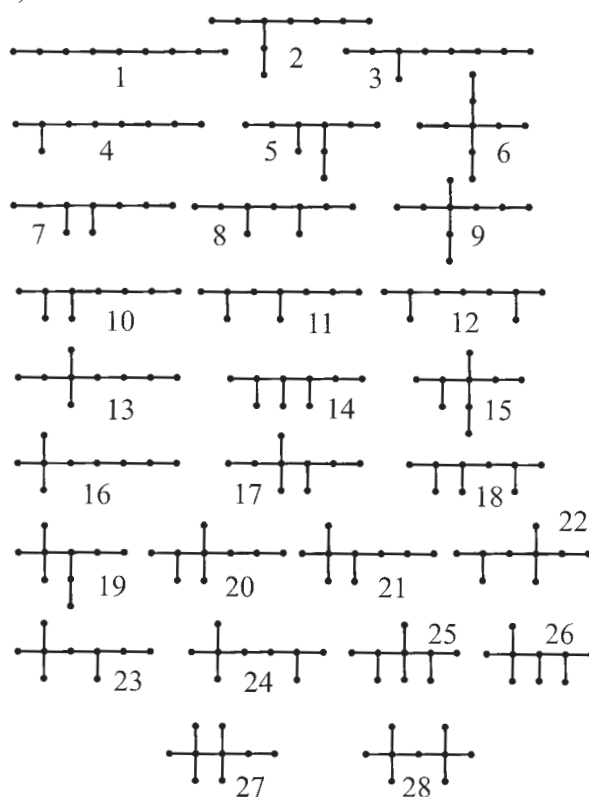


Fig. 2. A complete set of nonane isomers, without C -equivalent species; the isomers are ordered with respect to the Randić index $\chi = C(-1/2)$.

DEPENDENCE OF THE ORDERING OF ALKANE ISOMERS ON THE EXPONENT λ

The connectivity index $\chi = C(-1/2)$ was introduced² as a measure of branching. This means that if the alkanes are ordered so that their χ -values decrease then the extent of their branching should increase. This is indeed the case, especially if some “local” ambiguities are disregarded; *cf.* Fig. 2. Therefore the starting point for our analysis was

the ordering of isomeric alkanes according to the Randić χ -index, $C(\lambda)$, $\lambda = -0.5$. When λ assumes values other than -0.5 , then the ordering changes significantly.¹⁶ This is illustrated in Table I.

TABLE I. Isomeric decanes (without C-equivalent species) ordered according to $C(\lambda)$ for $\lambda = -0.3, -0.4, \dots, -1.2$; the order according to the Randić index χ (*i.e.*, according to $\lambda = -0.5$) is chosen to be 1, 2, 3, ..., 49

-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9	-1.0	-1.1	-1.2
1	1	1	1	1	1	1	4	4	4
2	2	2	2	2	2	4	1	7	7
3	3	3	4	4	4	2	2	2	2
5	5	4	3	3	7	7	7	1	1
4	4	5	6	7	3	6	6	6	6
6	6	6	7	6	6	3	3	8	8
8	8	7	5	8	8	8	8	3	16
9	7	8	8	5	5	11	11	16	3
10	9	9	9	9	9	9	16	11	11
7	10	10	10	11	11	16	9	9	18
12	12	11	11	10	10	5	5	18	9
13	13	12	12	16	16	10	10	10	10
15	11	13	13	12	18	18	18	5	20
11	15	14	16	14	12	14	14	14	14
17	17	15	14	13	14	12	20	20	5
14	14	16	18	18	13	20	12	12	22
19	16	17	17	17	20	13	13	22	12
23	19	18	15	20	17	17	17	17	17
21	18	19	20	19	22	22	22	13	13
16	21	20	19	15	19	19	19	19	25
18	23	21	22	22	15	24	24	25	24
26	20	22	21	21	21	25	25	24	19
20	22	23	24	24	24	21	21	21	21
27	26	24	23	25	25	15	15	29	29
22	24	25	25	23	28	28	28	28	28
24	27	26	26	28	29	29	29	15	36
25	25	27	28	29	23	23	23	36	15
28	28	28	29	26	26	30	30	31	31
29	29	29	27	30	30	31	31	30	30
30	30	30	30	27	31	26	36	23	23
31	31	31	31	31	27	36	26	26	37
32	32	32	32	32	36	27	27	37	26
33	33	33	33	36	32	32	32	34	41
35	34	34	34	33	33	33	33	33	39

TABLE I. Continued

-0.3	-0.4	-0.5	-0.6	-0.7	-0.8	-0.9	-1.0	-1.1	-1.2
34	35	35	36	34	34	34	34	32	34
38	36	36	35	37	37	37	37	41	33
36	38	37	37	35	39	39	39	39	32
37	37	38	39	39	35	41	41	27	27
39	39	39	38	41	41	35	35	35	40
40	40	40	40	40	40	40	40	40	35
42	42	41	41	38	38	38	38	42	42
43	41	42	42	42	42	42	42	38	45
44	43	43	43	43	45	45	45	45	38
41	44	44	44	45	43	43	43	43	43
45	45	45	45	44	44	44	44	44	44
46	46	46	46	46	46	46	46	46	46
47	47	47	48	48	48	48	48	48	48
48	48	48	47	47	47	47	47	49	49
49	49	49	49	49	49	49	49	47	47

As a kind of surprise, it was found that the ordering of alkanes with regard to $C(\lambda)$ is extremely sensitive to the choice of the exponent λ . Already minute deviations of λ from its usually accepted value -0.5 change the initial ordering, *cf.* Table II. This “instability” of the Randić index χ has obvious implication on its interpretation as a measure of branching. It shows that (in some cases) statements of the type “*isomer A is more branched than isomer B*” (based on χ) are not free of ambiguities.

TABLE II. Same data as in Table I for λ slightly deviating from its initial value -0.5

-0.45	-0.46	-0.47	-0.48	-0.49	-0.50	-0.51	-0.52	-0.53	-0.54	-0.55
1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6
8	8	8	8	7	7	7	7	7	7	7
7	7	7	7	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9
10	10	10	10	10	10	10	10	10	10	10
12	12	12	12	12	11	11	11	11	11	11
11	11	11	11	11	12	12	12	12	12	12
13	13	13	13	13	13	13	13	13	13	13
15	15	15	15	14	14	14	14	16	16	16
14	14	14	14	15	15	16	16	14	14	14

TABLE II. Continued

-0.45	-0.46	-0.47	-0.48	-0.49	-0.50	-0.51	-0.52	-0.53	-0.54	-0.55
17	17	17	17	16	16	15	15	17	17	17
16	16	16	16	17	17	17	17	15	15	15
18	18	18	18	18	18	18	18	18	18	18
19	19	19	19	19	19	19	19	19	19	19
21	21	20	20	20	20	20	20	20	20	20
20	20	21	21	21	21	21	21	21	21	21
23	23	23	23	22	22	22	22	22	22	22
22	22	22	22	23	23	23	23	23	23	23
26	26	24	24	24	24	24	24	24	24	24
24	24	26	26	26	25	25	25	25	25	25
25	25	25	25	25	26	26	26	26	26	26
27	27	27	27	27	27	28	28	28	28	28
28	28	28	28	28	28	27	27	29	29	29
29	29	29	29	29	29	29	29	27	27	27
30	30	30	30	30	30	30	30	30	30	30
31	31	31	31	31	31	31	31	31	31	31
32	32	32	32	32	32	32	32	32	32	32
33	33	33	33	33	33	33	33	33	33	33
34	34	34	34	34	34	34	34	34	34	34
35	35	35	35	35	35	35	35	35	36	36
36	36	36	36	36	36	36	36	36	35	35
38	38	37	37	37	37	37	37	37	37	37
37	37	38	38	38	38	39	39	39	39	39
39	39	39	39	39	39	38	38	38	38	38
40	40	40	40	40	40	40	40	40	40	40
41	41	41	41	41	41	41	41	41	41	41
42	42	42	42	42	42	42	42	42	42	42
43	43	43	43	43	43	43	43	43	43	43
44	44	44	44	44	44	44	44	44	44	44
45	45	45	45	45	45	45	45	45	45	45
46	46	46	46	46	46	46	46	46	46	46
47	47	47	47	47	47	47	47	47	47	47
48	48	48	48	48	48	48	48	48	48	48
49	49	49	49	49	49	49	49	49	49	49

In order to quantify the dependence of the ordering of the alkane isomers on the exponent λ , we introduced two functions Δ and Δ_w .

Suppose that the isomers considered are labeled by $1, 2, \dots, N$ and that their initial order (or more precisely: the initial order of their labels) is $a = (a_1, a_2, \dots, a_N)$. Let an-

other ordering of the same isomers be $b = (b_1, b_2, \dots, b_N)$. Hence, a and b are permutations of the numbers $1, 2, \dots, N$. The question is: *How to measure the difference between the permutations a and b ?*

A plausible and easy-to-guess measure of this difference is

$$\Delta = \Delta(a,b) = \frac{1}{N} \sum_{i=1}^N (a_i - b_i)^2. \quad (3)$$

The tacit assumption in the definition (3) is that all changes in the ordering b relative to a are equally important. However, in reality one is always more interested in the changes at the two ends of the orderings than at their middle. (For instance, knowing “*which is the most branched decane*” is certainly more important than “*which is the seventeenth-most branched decane*”.)

In view of this, we constructed another measure of the difference between two orderings:

$$\Delta_w = \Delta_w(a,b) = \frac{1}{N} \sum_{i=1}^N (a_i - b_i)^2 w(a_i) \quad (4)$$

where $w(a_i)$ is a pertinently chosen weight, constructed so as to give preference to the first and last members of the ordering relative to its middle members. The simplest such weight is the linear one defined as follows.

For N even:

$$\begin{aligned} w(a_i) &= 2 \frac{W-1}{N-2} \left(\frac{N}{2} - a_i \right) + 1 && \text{for } 1 \leq a_i \leq N/2 \\ &= 2 \frac{W-1}{N-2} \left(a_i - \frac{N+2}{2} \right) + 1 && \text{for } N/2 \leq a_i \leq N \end{aligned} \quad (5)$$

For N odd:

$$\begin{aligned} w(a_i) &= 2 \frac{W-1}{N-1} \left(\frac{N+1}{2} - a_i \right) + 1 && \text{for } 1 \leq a_i \leq (N+1)/2 \\ &= 2 \frac{W-1}{N-1} \left(a_i - \frac{N+1}{2} \right) + 1 && \text{for } (N+1)/2 \leq a_i \leq N \end{aligned} \quad (6)$$

In the above formulas, W is a weighting parameter, implying that changing place between the first and the second member (as well as between the last and the second-last member) is W times more important than the exchange of two middle members. In our calculations the value $W = 10$ was adopted, *i.e.*, that the first and last members are one order of magnitude more important than the middle members. Further, our labeling was such that $a_i = i$, which somewhat simplified the expressions (3) – (6).

DEPENDENCE OF Δ AND Δ_w ON THE EXPONENT λ

The quantities Δ and Δ_w , Eqs. (3) and (4) were calculated for $a = (1, 2, \dots, N)$ being the ordering of alkanes achieved by means of the Randić index χ and b being the analogous ordering according to $C(\lambda)$. By this Δ and Δ_w become functions of a continuous variable λ .

Both $\Delta(\lambda)$ and $\Delta_w(\lambda)$ are stepwise functions: when λ changes continuously, they remain constant and increase jumpwise at the value of λ for which a change (or, perhaps, several changes) in the ordering of the isomers occurs.

Calculations were performed for the above specified sets of alkanes with 6 through 10 carbon atoms (*cf.* Fig. 2). Two characteristic results are shown in Figs. 3 and 4. An example of the dependence of Δ on positive values of λ (which, from the point of view of chemical applications, is less interesting) is shown in Fig. 5.

What first needs to be noted – and what was by no means anticipated – is that $\Delta(\lambda)$ and $\Delta_w(\lambda)$ have almost identical forms (although their numerical values are quite differ-

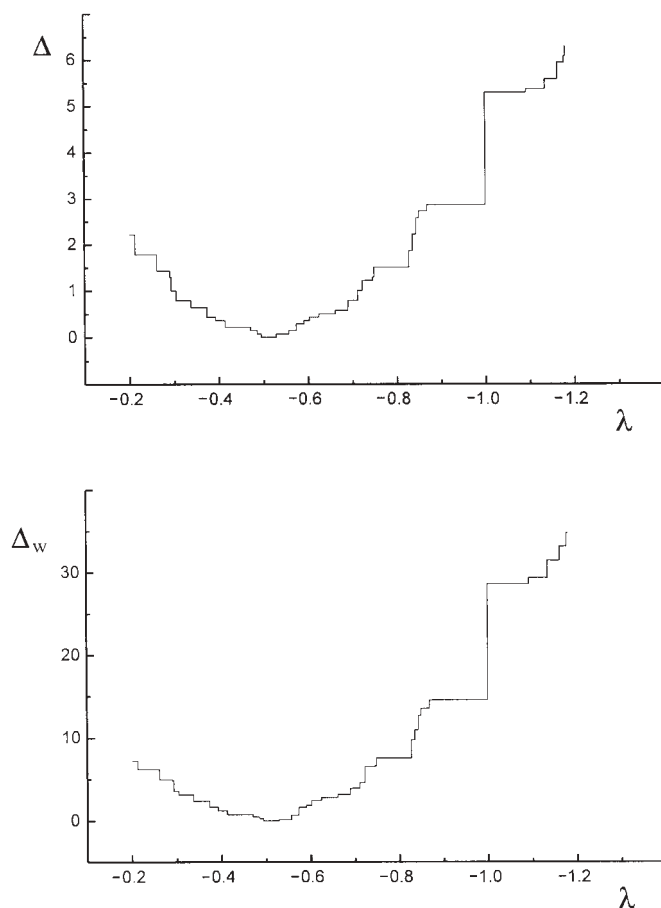


Fig. 3. The λ -dependence of the functions Δ and Δ_w for nonanes (*cf.* Fig. 2).

ent). Thus, at least for the alkane orderings studied in this work, the weighted Δ -function, Eqs. (4)–(6), yields essentially the same results as its unweighted variant, Eq. (3).

The most remarkable feature of $\Delta(\lambda)$ and $\Delta_w(\lambda)$ is that there are certain values (or narrow intervals) of λ at which the jumps of these functions are very large. The most obvious such jumps are at $\lambda = -1$ (seen in Figs. 3 and 4) and $\lambda = +1$ (seen in Fig. 5). These occur *exactly* at $\lambda = \pm 1$. Other pronounced jumps are *in the vicinity* of $\lambda = +2$ and $\lambda = +3$ (seen in Fig. 5) as well as of $\lambda = -2$ and $\lambda = -3$. Another sharp change of λ occurs around $\lambda = -0.85$ and consists of a large number of near-lying jumps (see Figs. 3 and 4). Each sudden change of Δ corresponds to a major rearrangement in the ordering of the alkane isomers. Two such rearrangements, at $\lambda \approx -0.85$ and at $\lambda = -1$ are especially significant for the interpretation of the connectivity index, and have been (from another point of view) examined in detail elsewhere.^{3,8}

Both $\Delta(\lambda)$ and $\Delta_w(\lambda)$ have a well-shaped minimum at $\lambda = -0.5$, a direct consequence of the definitions (3) and (4). In addition to this, there exist numerous other, shallow and difficult-to-notice minima: one such is seen in Fig. 5 at $\lambda \approx 1.95$.

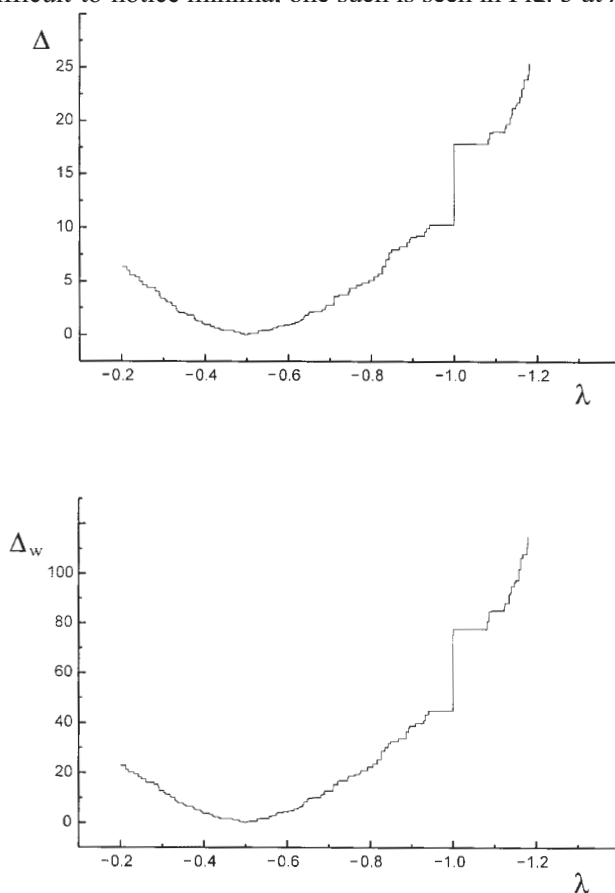


Fig. 4. Same as Fig. 3, for decanes.

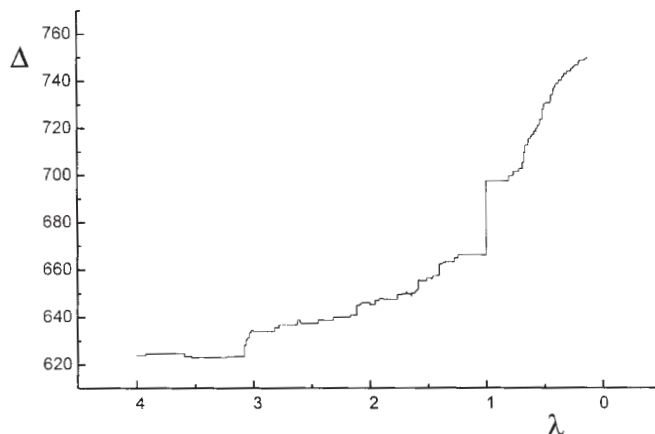


Fig. 5. The λ -dependence of the function D for decanes, for positive values of λ .

The maximal possible value of Δ is $(N^2 - 1)/3$, achieved for the permutation $b = (N, N - 1, \dots, 3, 2, 1)$. As demonstrated elsewhere,¹⁶ in the case of connectivity indices such an inverse ordering can never happen. On the other hand, when λ tends to zero (from the positive side) then $\Delta(\lambda)$ assumes values that are quite close to $(N^2 - 1)/3$. Thus for $n = 6, 7, 8, 9$ and 10 the limit value of $\Delta(\lambda)$ is $7.6, 25.3, 80.4, 244.8$ and 750.1 , respectively, which should be compared with $8, 26.7, 85, 261$ and 800 obtained for the hypothetical inverse ordering. It can be seen that the ordering of the alkane isomers according to $C(\lambda)$ maximally differs from the ordering according to the Randić index χ for near-zero positive values of the exponent λ .

ИЗВОД

РЕЂАЊЕ ИЗОМЕРНИХ АЛКАНА ПОМОЋУ ИНДЕКСА ПОВЕЗАНОСТИ

ИВАН ГУТМАН, ДУШИЦА ВИДОВИЋ и АНКА НЕДИЋ

Природно-математички факултет у Крајујевицу

Индекс повезаности органског молекула чији молекулски граф је G , дефинисан је као $C(\lambda) = \sum(\delta_u \delta_v)^\lambda$, где је δ_u степен чвора у графа G , где сумирање иде преко свих парова суседних чворова графа G и где је λ погодно одабрани експонент. Уобичајена вредност за λ је $-1/2$, када се $\chi = C(-1/2)$ назива Рандићев индекс. Поредак изомерних алкана према χ прати величину разгранатости угљеничног скелета. Ми смо проучили поредак изомера алкана са 6 до 10 угљеникових атома у односу на $C(\lambda)$ за разне вредности параметра λ . Поредак значајно зависи од λ . Разлика између поредака у односу на χ и у односу на $C(\lambda)$ мерена је функцијом Δ и установљено је како Δ зависи од λ .

(Примљено 3. септембра 2001)

REFERENCES

1. R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim, 2000
2. M. Randić, *J. Am. Chem. Soc.* **97** (1975) 6609

3. I. Gutman, M. Lepović, *J. Serb. Chem. Soc.* **66** (2001) 605
4. I. Gutman, N. Trinajstić, *Chem. Phys. Lett.* **17** (1972) 535
5. I. Gutman, B. Ruščić, N. Trinajstić, C. F. Wilcox, *J. Chem. Phys.* **62** (1975) 3399
6. B. Bollobás, P. Erdős, *Ars Combinatoria* **50** (1998) 225
7. L.H. Clark, J. W. Moon, *Ars Combinatoria* **54** (1999) 223
8. I. Gutman, M. Lepović, D. Vidović, L. H. Clark, *Indian J. Chem.* (in press)
9. M. Randić, P. J. Hansen, P. C. Jurs, *J. Chem. Inf. Comput. Sci.* **28** (1988) 60
10. E. Estrada, *J. Chem. Inf. Comput. Sci.* **35** (1995) 1022
11. D. Amić, D. Beslo, B. Lučić, S. Nikolić, N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **38** (1998) 819
12. E. Estrada, *Chem. Phys. Lett.* **312** (1999) 556
13. L. B. Kier, L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976
14. L. B. Kier, L. H. Hall, *Molecular Connectivity in Structure-Activity Analysis*, Research Studies Press, Letchworth, 1986
15. A. T. Balaban, D. Mills, S. Basak, *J. Chem. Inf. Comput. Sci.* **42** (2002), in press
16. As shown in the subsequent paper (I. Gutman, *J. Serb. Chem. Soc.* **67** (2002) 99), the order of certain pairs of alkanes (relative to each other) remains the same for all values of the exponent λ .