mathch no. 25 pp. 41-70 1990

DESIGN OF TOPOLOGICAL INDICES. PART 2.[1]

DISTANCE MEASURE CONNECTIVITY INDICES.

Alexandru T. Balaban,

Polytechnic Institute, Organic Chemistry Department,

Splaiul Independentei 313, 76206 Bucharest, Roumania

Dan Ciubotariu and Ovidiu Ivanciuc*

Polytechnic Institute, Organic Chemistry Department,

1900 Timisoara, Roumania

(Received: April 1989)

The use of distance measure in defining a highly selective, connectivity-based spectrum of topological indices, denoted by DM^k is presented. They are obtained by a distance measure method, using a spectrum of fourteen connectivity topological indices. The newly defined topological indices have no degenerate value for a set of 661 alkanes between butane and dodecane. The statistical characteristics and the ordering of structures induced by the values of the indices are also considered.

---

* correspondence address : Strada Grigorescu 18/2,
2750 Hunedoara, Roumania

A great number of methods from mathematics and computer science have been largely investigated for the elucidation of the causal relationships between chemical structure and physico-chemical properties. PAttern ReCognition (PARC)[2] methods are especially useful for the classification of chemical structures into distinct classes on the basis of characteristic features. A set of characteristic features of a chemical structure is considered as an abstract pattern which contains information about physico-chemical properties, chemical reactivity or biological activity of the investigated chemical structure.

We have to mention here that PARC methods try to find relationships between the pattern and a chemical property without using chemical knowledge.

A largely used classification method in PARC is the distance measure in pattern space, briefly described in what follows.

Let $X(x_1, x_2, \ldots, x_n)$, $Y(y_1, y_2, \ldots, y_n)$ and $Z(z_1, z_2, \ldots, z_n)$ be three points in an n-dimensional space. Any function $D(X,Y)$ which satisfies the following relationships:

$$\begin{aligned} &D(X,X)=0 \\ &D(X,Y)>0 \quad \text{for } X\neq Y \\ &D(X,Y)=D(Y,X) \\ &D(X,Z)+D(Z,Y)\geq D(X,Y) \end{aligned} \qquad (1)$$

has the quality of a distance.

The most frequently used distance measurements form the general Minkowski distance:

$$D_M = \sum_{i=1}^{n} \left[(x_i - y_i)^k \right]^{1/k} \qquad (2)$$

For k=1 we obtain the Manhattan distance("city block distance") in analogy with the shortest distance between two points in a city with rectangular streets:

$$D_{Mh} = \sum_{i=1}^{n} |x_i - y_i| \qquad (3)$$

and k=2 gives the most often used Euclidean distance:

$$D_E = \sum_{i=1}^{n} (x_i - y_i)^2 \quad {}^{1/2} \qquad (4)$$

Clasifications of chemical structures by distance measurements in pattern space were used for the recognition of the structure of molecules from mass spectra[3] and infrared spectra.[4] The distance measure method was also applied in studies concerning aromaticity, photoreactivity, thermal reactivity and intermolecular forces.[5]

Another method used in establishing correlations between structure and physico-chemical properties is to express the bonding topology of the molecular graph by a mathematical expression which may be a matrix, a polynomial or a numerical index.[6-11] Numerical indices developed in this way are callud topological indices (TI's).

Although in their origins TI's were developed for the purpose of obtaining correlations with a great variety of physicochemical

properties of chemical substances, namely QSPR (Quantitative Structure-Property Relationships), their range of applications has extended to bibliographical classification of chemical compounds and QSAR (Quantitative Structure-Activity Relationships).

A general problem of topological indices is that they are more or less degenerate, i.e. two or more nonisomorphic structures may lead to the same value for a given topological index. The discriminating ability or structural selectivity of a TI is inversely related to its degenaracy. A study on the structural selectivity of six TI's : Wiener's number W, the Zagreb group (Gutman, Ruscic, Trinajstic, Wilcox) TI denoted by M_1, Hosoya's index Z, Randic's connectivity index χ , the information on distance index I_D^E and Balaban's average distance sum connectivity index J, reveals that J has the greatest selectivity.[12] Of course, high selectivity, although strongly desirable, is not sufficient to obtain good results in QSPR and QSAR if the values of indices do not properly reflect structural information in conjunction with the property under investigation.

The Randic connectivity index[8] was defined as:

$$\chi = \sum_{m} (v_i v_j)^{-1/2} \qquad (5)$$

where v_i and v_j denote the degrees of the two endpoints of an edge in the molecular graph G, and the summation is extended over all m edges.

A generalized connectivity index was suggested by extending the summation in eq. (5) over all possible

connected subgraphs with m edges of four types : path, cluster, path/cluster and chain:[9]

$$ {}^{m}X_t = \sum_{j=1}^{n_m} \prod_{i=1}^{r} (v_i)_j^{-1/2} \qquad (6) $$

where n_m is the number of connected subgraphs of type t with m edges and r is the number of vertices of the subgraph. In the case of path, cluster and path/cluster subgraphs, r=m+1 and for a chain-type subgraph r=m.

A further extension of the connectivity index was suggested in order to take into account the presence in the molecule of atoms of different chemical nature as well as the presence of single, double, triple and aromatic bonds. The vertex degree v_i is replaced by the atom connectivity Δ_i^v:[10]

$$ \Delta_i^v = Z_i^v - H_i \qquad (7) $$

where Z_i^v and H_i are the number of valence electrons of atom i and the number of hydrogen atoms attached to this atom, respectively. In order to obtain valence delta-values (Δ_i^v) for second and third level atoms and halogens, a slightly different definition was proposed:[11]

$$ \Delta_i^v = \frac{Z_i^v - H_i}{Z_i - Z_i^v} \qquad (8) $$

where the signifiance of Z_i^v and H_i is the same as in equation (7) and Z_i is the count of all electrons of atom i. The connectivity

index is given by:

$$ {}^{m}\chi_t^{v} = \sum_{j=1}^{n_m} \prod_{i=1}^{r} (\Delta_i^{v})_j^{-1/2} \qquad (9) $$

The generalized connectivity indices have found numerous applications in establishing QSPR for alkanes. Good correlations were obtained for the heat of atomization, heat of formation, boiling point, molar refraction, molar magnetic susceptibilities, heat of vaporization, solubility.[9]

TOPOLOGICAL INDICES DM^k

The mechanism of constructing a TI, considered earlier,[13] was revealed to be a two-stage process, viz. the assignment and the operational stages, respectively. The assignment stage consists in finding a local vertex invariant (LOVI). In the operational stage the LOVI's are mathematically manipulated in order to produce a number which represents the TI. For the generalized connectivity index ${}^{m}\chi_t$, the LOVI is selected to be the vertex degree and the operational stage consists in a summation over all subgraphs of connectivity contributions of the subgraphs.

We introduce here a new spectrum of highly selective TI's as a result of another operational stage using as the basis of computations a spectrum of connectivity indices ${}^{m}\chi_t$, based on the method of distance measure; we denote these indices by DM^k:

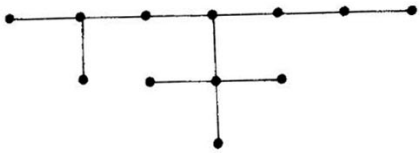
$$DM^k = \sum_{i=1}^{14} \left[({}^m\chi_{t,i} - {}^m\chi_{t,r})^k \right]^{1/k} \qquad (10)$$

where ${}^m\chi_{t,i}$ and ${}^m\chi_{t,r}$ represents the corresponding connectivity indices of the molecular and reference structure, respectively, and the summation is done over all fourteen connectivity terms up to the sixth order (m=6) and $k=\overline{1,5}$. The reference molecular graph is considered to be that of methane, in whose case all ${}^m\chi_{t,r}$ are equal to zero. By its definition, this type of TI is reminiscent of the superindex.[14] In FIGURE 1 is presented the computation of DM^k TI's for 4-tB-2-M-C_7.

We have investigated some basic properties of the new TI's DM^k for a population consisting of 661 alkane isomers between butane and dodecane. The alkane class of chemical compounds was chosen due to the rapidly increasing number of isomers which allows comparisons between isomers with subtle structural differences. The set was limited to dodecane due to the small interest in investigating higher alkane isomers, reflected in the small amount of data on various physico-chemical properties. For the set of 661 alkane isomers no degenerate value was detected.

In TABLE 1 are presented values of TI's DM^k for the 147 alkane isomers between butane and decane. The abbreviations used in the nomenclature of alkanes are: M=methyl, E=ethyl, P=propyl, iP=isopropyl, tB=tert-butyl and a linear chain of n carbon atoms is symbolized by C_n.

Some statistics are given in TABLES 2-6 for the whole set of 661 alkane isomers between butane and dodecane and for the sets of 35



$^{0}\chi = 9.77602$ $\quad$ $^{1}\chi = 5.37523$ $\quad$ $^{2}\chi = 5.48117$

$^{3}\chi_p = 2.61780$ $\quad$ $^{3}\chi_c = 1.91861$ $\quad$ $^{4}\chi_p = 1.82191$

$^{4}\chi_c = 0.28868$ $\quad$ $^{4}\chi_{pc} = 2.07886$ $\quad$ $^{5}\chi_p = 1.49367$

$^{5}\chi_c = 0.43301$ $\quad$ $^{5}\chi_{pc} = 2.36317$ $\quad$ $^{6}\chi_p = 0.23570$

$^{6}\chi_c = 0.14434$ $\quad$ $^{3}\chi_{pc} = 3.67297$

$$DM^{1} = \sum_{i=1}^{14} {}^{m}\chi_{t,i} = 37.70115 \qquad DM^{2} = \sum_{i=1}^{14} \left[{}^{m}\chi_{t,i}\right]^{1/2} = 13.94043$$

$$DM^{3} = \sum_{i=1}^{14} \left[{}^{m}\chi_{t,i}\right]^{1/3} = 11.08072 \qquad DM^{4} = \sum_{i=1}^{14} \left[{}^{m}\chi_{t,i}\right]^{1/4} = 10.28268$$

$$DM^{5} = \sum_{i=1}^{14} \left[{}^{m}\chi_{t,i}\right]^{1/5} = 9.99367$$

FIGURE 1. The computation of DM^{k} TI's for 4-tB-2-M-C_7.

TABLE 1. Values of the TI's DM^k for alkane isomers C_4 through C_{10}.

No.	Name	DM^1	DM^2	DM^3	DM^4	DM^5
	C_4H_{10}					
1	C_4	6.82843	4.07076	3.63268	3.50180	3.45268
2	$2\text{-}M\text{-}C_3$	7.61880	4.37387	3.83417	3.67240	3.61471
	C_5H_{12}					
3	C_5	8.94975	5.02701	4.43021	4.24947	4.17977
4	$2\text{-}M\text{-}C_4$	9.98960	5.26852	4.59278	4.39928	4.33074
5	$2,2\text{-}M_2\text{-}C_3$	12.00000	6.12372	5.12041	4.78246	4.64067
	C_6H_{14}					
6	C_6	11.15685	5.99576	5.23310	5.00010	4.90870
7	$2\text{-}M\text{-}C_5$	12.37312	6.22647	5.39444	5.14933	5.05843
8	$3\text{-}M\text{-}C_5$	12.55852	6.24659	5.38987	5.14463	5.05627
9	$2,2\text{-}M_2\text{-}C_4$	15.07107	6.86289	5.73880	5.41341	5.29451
10	$2,3\text{-}M_2\text{-}C_4$	13.95214	6.62242	5.60300	5.31770	5.21903
	C_7H_{16}					
11	C_7	13.42462	6.97257	6.03945	5.75273	5.63891
12	$2\text{-}M\text{-}C_6$	14.76562	7.19466	6.19899	5.90056	5.78707
13	$3\text{-}M\text{-}C_6$	15.08212	7.20870	6.19075	5.89492	5.78480
14	$3\text{-}E\text{-}C_5$	15.36658	7.25435	6.19705	5.89658	5.78615
15	$2,2\text{-}M_2\text{-}C_5$	17.94975	7.81372	6.53066	6.16006	6.02112
16	$2,3\text{-}M_2\text{-}C_5$	16.94921	7.57322	6.38401	6.05477	5.93981
17	$2,4\text{-}M_2\text{-}C_5$	16.36313	7.54215	6.40512	6.07560	5.95274
18	$3,3\text{-}M_2\text{-}C_5$	18.48528	7.88594	6.50882	6.12932	5.99923
19	$2,2,3\text{-}M_3\text{-}C_4$	20.54701	8.46943	6.85240	6.37878	6.20511
	C_8H_{18}					
20	C_8	15.61028	7.95382	6.84817	6.50676	6.37006
21	$2\text{-}M\text{-}C_7$	17.02015	8.16964	7.00646	6.65346	6.51690
22	$3\text{-}M\text{-}C_7$	17.42935	8.18587	6.99937	6.64828	6.51478
23	$4\text{-}M\text{-}C_7$	17.56044	8.19717	6.99968	6.64827	6.51494
24	$3\text{-}E\text{-}C_6$	17.91494	8.22740	7.00082	6.64815	6.51565
25	$2,2\text{-}M_2\text{-}C_6$	20.51561	8.75210	7.32066	6.90480	6.74622
26	$2,3\text{-}M_2\text{-}C_6$	19.60890	8.52990	7.18120	6.80400	6.66787
27	$2,4\text{-}M_2\text{-}C_6$	19.20822	8.49687	7.18043	6.80789	6.67080
28	$2,5\text{-}M_2\text{-}C_6$	18.60840	8.47418	7.19967	6.82303	6.67933
29	$3,3\text{-}M_2\text{-}C_6$	21.42983	8.83181	7.29276	6.87327	6.72557
30	$3,4\text{-}M_2\text{-}C_6$	20.00839	8.60561	7.19309	6.80405	6.66684
31	$3\text{-}E\text{-}2\text{-}M\text{-}C_5$	20.10744	8.61035	7.18685	6.80099	6.66587
32	$3\text{-}E\text{-}3\text{-}M\text{-}C_5$	22.13693	9.03071	7.34684	6.88412	6.72587
33	$2,2,3\text{-}M_3\text{-}C_5$	24.14856	9.39424	7.57885	7.07839	6.90390
34	$2,2,4\text{-}M_3\text{-}C_5$	22.80578	9.34954	7.63329	7.13473	6.94471
35	$2,3,3\text{-}M_3\text{-}C_5$	24.49869	9.54702	7.64187	7.10049	6.90990
36	$2,3,4\text{-}M_3\text{-}C_5$	22.07279	9.03695	7.41989	6.98234	6.83206
37	$2,2,3,3\text{-}M_4\text{-}C_4$	29.75000	11.00852	8.45704	7.65464	7.32502

TABLE 1. (Continued).

No.	Name	^{1}DM	^{2}DM	^{3}DM	^{4}DM	^{5}DM
	C_9H_{20}					
38	C_9	17.81111	8.93865	7.65860	7.26185	7.10190
39	2-M-C_8	19.15960	9.14907	7.81589	7.40758	7.24763
40	3-M-C_8	19.63434	9.16570	7.80947	7.40269	7.24561
41	4-M-C_8	19.85814	9.17512	7.80862	7.40243	7.24574
42	3-E-C_7	20.26217	9.20291	7.81016	7.40250	7.24652
43	4-E-C_7	20.46331	9.21844	7.80990	7.40194	7.24640
44	2,2-M_2-C_7	22.71384	9.69154	8.11706	7.65272	7.47306
45	3,3-M_2-C_7	23.89581	9.77848	8.09048	7.62313	7.45389
46	4,4-M_2-C_7	24.27449	9.85039	8.10692	7.62796	7.45594
47	2,3-M_2-C_7	21.89196	9.49158	7.98721	7.55667	7.39730
48	2,4-M_2-C_7	21.62238	9.48662	7.99251	7.56244	7.40113
49	2,5-M_2-C_7	21.20796	9.43959	7.98444	7.55983	7.39974
50	2,6-M_2-C_7	20.59842	9.42199	8.00267	7.57418	7.40796
51	3,4-M_2-C_7	22.57392	9.56926	7.99165	7.55354	7.39531
52	3,5-M_2-C_7	21.95915	9.50164	7.98150	7.55301	7.39575
53	3-E-2-M-C_6	22.74302	9.57620	7.98845	7.55237	7.39524
54	3-E-3-M-C_6	25.10292	9.99240	8.12868	7.62461	7.45050
55	4-E-2-M-C_6	22.12826	9.51664	7.98174	7.55340	7.39640
56	3-E-4-M-C_6	23.18683	9.67321	8.00961	7.55653	7.39627
57	2,2,3-M_3-C_6	26.88478	10.32016	8.35562	7.81829	7.62804
58	2,2,4-M_3-C_6	25.72740	10.25209	8.36344	7.83605	7.64225
59	2,2,5-M_3-C_6	24.59251	10.14575	8.37688	7.86131	7.66134
60	2,3,3-M_3-C_6	27.48250	10.45963	8.39217	7.82371	7.62582
61	3,3,4-M_3-C_6	28.03870	10.62413	8.43764	7.83174	7.62448
62	2,4,4-M_3-C_6	26.32512	10.34058	8.36583	7.82400	7.63170
63	2,3,4-M_3-C_6	25.16898	10.08645	8.22887	7.72698	7.55497
64	2,3,5-M_3-C_6	23.90616	9.88340	8.19818	7.73202	7.56312
65	3,3-E_2-C_5	26.01713	10.26930	8.22328	7.65772	7.46271
66	3-E-2,4-M_2-C_5	25.48212	10.16756	8.24897	7.73310	7.55742
67	3-E-2,3-M_2-C_5	28.56041	10.78609	8.50079	7.85603	7.63327
68	3-E-2,2-M_2-C_5	27.78762	10.52467	8.39863	7.82197	7.62442
69	2,2,3,3-M_4-C_5	34.31891	12.04613	9.17576	8.31011	7.97684
70	2,2,4,4-M_4-C_5	30.44023	11.65191	9.11569	8.33002	8.01423
71	2,2,3,4-M_4-C_5	30.24957	11.17576	8.75092	8.06560	7.82374
72	2,3,3,4-M_4-C_5	31.35641	11.47089	8.88594	8.12645	7.84904

TABLE 1. (Continued).

No.	Name	DM^1	DM^2	DM^3	DM^4	DM^5
	$C_{10}H_{22}$					
73	C_{10}	20.01193	9.92572	8.47029	8.01771	7.83427
74	2-M-C_9	21.36714	10.13338	8.62679	8.16264	7.97905
75	3-M-C_9	21.78617	10.14846	8.62093	8.15800	7.97714
76	4-M-C_9	22.07552	10.15850	8.62018	8.15775	7.97726
77	5-M-C_9	22.16822	10.16382	8.62032	8.15774	7.97728
78	2,2-M_2-C_8	24.81999	10.64846	8.91848	8.40280	8.20118
79	2,3-M_2-C_8	24.06698	10.46091	8.79527	8.31037	8.12753
80	2,4-M_2-C_8	23.89010	10.45508	8.79845	8.31531	8.13106
81	2,5-M_2-C_8	23.60678	10.42583	8.79376	8.31416	8.13057
82	2,6-M_2-C_8	23.18264	10.40213	8.79213	8.31331	8.12982
83	2,7-M_2-C_8	22.68408	10.39315	8.80978	8.32685	8.13756
84	3,3-M_2-C_8	26.06631	10.72254	8.89190	8.37493	8.18329
85	3,4-M_2-C_8	24.83083	10.53084	8.79828	8.30735	8.12574
86	3,5-M_2-C_8	24.34717	10.48318	8.78966	8.30658	8.12618
87	3,6-M_2-C_8	23.78137	10.44390	8.78975	8.30780	8.12639
88	4,4-M_2-C_8	26.71276	10.79542	8.90243	8.37746	8.18464
89	4,5-M_2-C_8	25.11330	10.55749	8.79845	8.30621	8.12534
90	3-E-C_8	22.41251	10.18394	8.62161	8.15778	7.97804
91	4-E-C_8	22.75588	10.19922	8.62052	8.15708	7.97792
92	2,2,3-M_3-C_7	29.08342	11.22327	9.13998	8.56287	8.35391
93	2,2,4-M_3-C_7	28.05714	11.19871	9.16114	8.58528	8.37006
94	2,2,5-M_3-C_7	27.10765	11.04539	9.12346	8.57334	8.36494
95	2,2,6-M_3-C_7	26.26410	11.01103	9.15583	8.60301	8.38489
96	2,3,3-M_3-C_7	29.85621	11.37078	9.17286	8.56576	8.35066
97	3,3,4-M_3-C_7	30.66136	11.53969	9.20671	8.56713	8.34680
98	3,3,5-M_3-C_7	29.13318	11.31418	9.14407	8.55331	8.34607
99	3,3,6-M_3-C_7	27.88043	11.13462	9.12128	8.56002	8.35456
100	2,4,4-M_3-C_7	29.07751	11.37966	9.19395	8.58424	8.36382
101	3,4,4-M_3-C_7	30.90894	11.58714	9.21826	8.57001	8.34761
102	2,3,4-M_3-C_7	27.64751	11.02861	9.02053	8.47417	8.28256
103	2,3,5-M_3-C_7	26.57008	10.87985	8.99496	8.47216	8.28383
104	2,3,6-M_3-C_7	25.62781	10.77823	8.99238	8.48127	8.29077
105	2,4,5-M_3-C_7	26.78416	10.90358	8.99605	8.47180	8.28389
106	2,4,6-M_3-C_7	25.64155	10.82830	9.01621	8.49608	8.29987
107	3,4,5-M_3-C_7	28.15810	11.15391	9.05370	8.48142	8.28354
108	3-E-2-M-C_7	25.04946	10.53028	8.79357	8.30580	8.12557
109	3-E-3-M-C_7	27.50197	10.92189	8.92121	8.37376	8.17941

TABLE 1. (Continued).

No.	Name	DM^1	DM^2	DM^3	DM^4	DM^5
110	3-E-4-M-C_7	25.68925	10.64597	8.81351	8.30818	8.12585
111	3-E-5-M-C_7	24.81607	10.52855	8.79272	8.30514	8.12538
112	3-E-6-M-C_7	24.00000	10.43927	8.78441	8.30659	8.12654
113	4-E-2-M-C_7	24.63584	10.50349	8.79255	8.30824	8.12735
114	4-E-3-M-C_7	25.75930	10.64386	8.81150	8.30782	8.12587
115	4-E-4-M-C_7	28.00197	11.00060	8.93485	8.37558	8.17959
116	4-P-C_7	23.00655	10.21685	8.62231	8.15743	7.97812
117	4-iP-C_7	25.40198	10.57325	8.79896	8.30666	8.12594
118	2,2,3,3-M_4-C_6	37.26561	12.89461	9.86278	8.98100	8.65585
119	2,2,4,4-M_4-C_6	33.92246	12.57776	9.79473	8.97267	8.66343
120	2,2,5,5-M_4-C_6	30.82107	12.03627	9.67511	8.97033	8.68891
121	2,2,3,4-M_4-C_6	33.31078	12.19834	9.54246	8.79551	8.53532
122	2,2,3,5-M_4-C_6	31.34085	11.78067	9.43108	8.77884	8.54299
123	2,2,4,5-M_4-C_6	30.58415	11.73366	9.43281	8.78793	8.55085
124	2,3,3,4-M_4-C_6	34.82191	12.57600	9.69677	8.85677	8.55724
125	2,3,3,5-M_4-C_6	32.47364	12.00048	9.48326	8.78394	8.53749
126	3,3,4,4-M_4-C_6	38.69543	13.44154	10.12324	9.10048	8.70714
127	2,3,4,4-M_4-C_6	34.06521	12.46106	9.64954	8.83575	8.54843
128	2,3,4,5-M_4-C_6	30.41994	11.66699	9.31851	8.67458	8.45424
129	3-E-2,2-M_2-C_6	30.48033	11.45571	9.17952	8.56379	8.34965
130	3-E-2,3-M_2-C_6	31.55197	11.74018	9.26854	8.58533	8.35152
131	3-E-2,4-M_2-C_6	28.58562	11.22622	9.06648	8.48296	8.28352
132	3-E-2,5-M_2-C_6	27.10464	10.94854	9.00526	8.47534	8.28582
133	3-E-3,4-M_2-C_6	32.04063	11.95056	9.35677	8.62027	8.36471
134	3-E-4,4-M_2-C_6	31.61798	11.79031	9.28443	8.59014	8.35303
135	3-E-4,5-M_2-C_6	28.37154	11.16887	9.05033	8.47865	8.28235
136	3-E-5,5-M_2-C_6	28.70456	11.25413	9.14581	8.56573	8.35623
137	3-E-3,5-M_2-C_6	30.02379	11.50524	9.19385	8.56378	8.34677
138	3,3-E_2-C_6	28.93763	11.23770	9.00674	8.39712	8.18646
139	3,4-E_2-C_6	26.32239	10.76658	8.84050	8.31541	8.12876
140	3-iP-2-M-C_6	28.18207	11.11838	9.04128	8.48043	8.28513
141	2,2,3,3,4-M_5-C_5	42.21763	14.41131	10.69097	9.50037	9.01246
142	2,2,3,4,4-M_5-C_5	39.68173	13.99388	10.52846	9.42363	8.97431
143	3-E-2,2,3-M_3-C_5	38.86701	13.44801	10.11571	9.09464	8.70363
144	3-E-2,2,4-M_3-C_5	34.02821	12.48650	9.68147	8.86589	8.57032
145	3-E-2,3,4-M_3-C_5	35.71442	12.88210	9.83714	8.92329	8.58781
146	3,3-E_2-2-M-C_5	32.76944	12.17784	9.44863	8.66004	8.38220
147	3-iP-2,4-M_2-C_5	31.40982	11.99926	9.45440	8.73297	8.47956

nonanes, 75 decanes, 159 undecanes and 355 dodecanes, consisting in the lower, upper and average values, standard deviation, relative width of the interval of values, defined as in eq. (11) and the

$$R_w = \frac{V_{upper} - V_{lower}}{V_{average}} \qquad (11)$$

distribution of the values of the corresponding TI in ten equal intervals between the lower upper value of the TI. In order to express quantitatively the distribution of the values of a TI, the informational energy content[15] was computed:

$$E = \sum p_i^2 \qquad (12)$$

where

$$p_i = \frac{N_i}{N} \qquad (13)$$

whose N_i and N are the number of isomers in the i-th interval and the total number of isomers, respectively. For an even distribution of the values of a TI, E=0.1, and the more uneven the distribution, the closer to 1 is the information energy content.

The most apparent features in TABLES 2-6 are that the populations of DM^k indices are overlapping for consecutive numbers of carbon atoms, and the standard deviation and relative width of the interval of values increase in the same direction. This feature reveals that the values are spread over a larger field as the number of carbon atoms increases, thus even if the structural differences are less pronounced the indices discriminate well the set of structures.

TABLE 7 contains the values of information energy content of the distribution of DM^k in ten equal intervals for the classes of isomers

TABLE 2. Statistics for the set of 661 C_4-C_{12} alkanes of the values of TI's DM^k.

TI	Lower value	Upper value	Average value	Standard deviation	Relative width of the interval of values
DM^1	6.82843	59.38371	32.74990	58.32095	1.60475
DM^2	4.07076	20.29166	12.71322	4.40355	1.27591
DM^3	3.63268	14.94078	10.20309	1.98213	1.10830
DM^4	3.50180	13.07701	9.47444	1.45499	1.01064
DM^5	3.45268	12.18046	9.20150	1.27673	0.94852

Intervals

TI	1	2	3	4	5	6	7	8	9	10
DM^1	6	12	35	97	198	174	89	33	13	4
DM^2	4	10	17	55	132	277	119	33	11	3
DM^3	4	6	14	31	74	182	286	51	9	4
DM^4	4	6	8	19	41	136	282	148	13	4
DM^5	3	5	10	17	33	76	161	331	21	4

TABLE 3. Statistics for the set of 35 C_9 alkanes of the values of TI's DM^k.

TI	Lower value	Upper value	Average value	Standard deviation	Relative width of the interval of values
DM^1	17.81111	34.31891	24.38361	14.50162	0.67700
DM^2	8.93865	12.04613	10.00126	0.56307	0.31071
DM^3	7.65860	9.17576	8.20269	0.12876	0.18496
DM^4	7.26185	8.33002	7.69098	0.06073	0.13889
DM^5	7.10190	8.01423	7.50359	0.04279	0.12159

Intervals

TI	1	2	3	4	5	6	7	8	9	10
DM^1	2	5	8	4	6	3	3	2	1	1
DM^2	6	6	6	5	5	3	0	1	2	1
DM^3	5	1	11	6	6	2	0	1	1	2
DM^4	1	5	9	5	3	8	0	1	1	2
DM^5	1	5	0	13	3	8	1	1	1	2

TABLE 4. Statistics for the set of 75 C_{10} alkanes of the values of TI's DM^k.

TI	Lower value	Upper value	Average value	Standard deviation	Relative width of the interval of values
DM^1	20.01193	42.21763	28.29778	21.73949	0.78471
DM^2	9.92572	14.41131	11.26055	0.88844	0.39835
DM^3	8.47029	10.69097	9.13417	0.19371	0.24312
DM^4	8.01771	9.50037	8.52595	0.08152	0.17390
DM^5	7.83427	9.01246	8.30245	0.05220	0.14191

Intervals

TI	1	2	3	4	5	6	7	8	9	10
DM^1	5	11	15	14	12	8	4	2	3	1
DM^2	8	21	18	8	8	6	2	2	0	2
DM^3	8	19	16	15	6	5	2	2	0	2
DM^4	8	14	9	25	4	7	4	2	0	2
DM^5	1	7	22	11	17	4	7	4	0	2

TABLE 5. Statistics for the set of 159 C_{11} alkanes of the values of TI's DM^k.

TI	Lower value	Upper value	Average value	Standard deviation	Relative width of the interval of values
DM^1	22.21276	54.37500	32.36131	29.79951	0.99385
DM^2	10.91443	18.17751	12.57235	1.38487	0.57770
DM^3	9.28294	13.20660	10.09848	0.32516	0.38854
DM^4	8.77419	11.48336	9.38314	0.13309	0.28873
DM^5	8.56706	10.67394	9.11676	0.07728	0.23110

Intervals

TI	1	2	3	4	5	6	7	8	9	10
DM^1	11	32	42	33	20	9	8	3	0	1
DM^2	36	51	33	18	11	4	4	1	0	1
DM^3	36	61	32	16	5	5	3	0	0	1
DM^4	9	76	42	18	7	3	3	0	0	1
DM^5	9	36	62	36	8	4	3	0	0	1

TABLE 6. Statistics for the set of 335 C_{12} alkanes of the values of TI's DM^k.

TI	Lower value	Upper value	Average value	Standard deviation	Relative width of the interval of values
DM^1	24.41358	59.38371	36.32847	37.12236	0.96261
DM^2	11.90437	20.29166	13.87807	1.90189	0.60436
DM^3	10.09633	14.94078	11.06665	0.48889	0.43775
DM^4	9.53115	13.07701	10.24626	0.20890	0.34606
DM^5	9.30018	12.18046	9.93676	0.11987	0.28986

	Intervals									
TI	1	2	3	4	5	6	7	8	9	10
DM^1	14	66	87	75	56	25	16	12	2	2
DM^2	63	119	78	47	21	13	8	3	1	2
DM^3	63	166	66	30	14	9	2	2	2	1
DM^4	61	145	97	27	13	6	2	1	2	1
DM^5	22	155	116	40	11	5	2	1	2	1

TABLE 7. The informational energy content of the distribution of DM^k.

Isomers	DM^1	DM^2	DM^3	DM^4	DM^5
C_4-C_{12}	0.20482	0.25854	0.28450	0.27984	0.32781
C_9	0.13796	0.14122	0.18694	0.17224	0.22449
C_{10}	0.14311	0.17867	0.17404	0.18756	0.18293
C_{11}	0.18010	0.21617	0.25145	0.31696	0.26134
C_{12}	0.17389	0.21512	0.29416	0.27851	0.31518

examined in TABLES 2-6. As a rule, with increasing the number of carbon atoms the distribution becomes more and more uneven. The most evenly distributed is DM^1 and the most unevenly distributed are DM^4 and DM^5.

TABLE 8 presents the intercorrelation matrix of the TI's DM^k for the set of 661 isomers. It results that all TI's are strongly intercorrelated, especially DM^2, DM^3 and DM^3, DM^4, DM^5. The intercorrelation coefficients indicate that all TI's express approximately the same type of structural information and they are not worthy to be used as independent variables in a multiparametric regression equation.

As previously mentioned, the main goal in developing a TI is its capacity to be a reliable parameter in QSPR and QSAR. Because TI's $^0\chi$, $^1\chi$ and $^2\chi$ are largely used as parameters in correlational equations, we have computed the correlation matrix between the above mentioned connectivity indices and DM^k and the results are presented in TABLE 9. Taking into consideration that the correlation coefficients are rather high, especially for DM^3, DM^4 and DM^5, we expect that they could be used as good correlational parameters.

There is another important feature of graph invariants: they can be used to characterize and to order the molecular structures according to their branching. The rules of branching were recently defined and tested using informational graph invariants.[16]

An example of ordering of a structural population using DM^k is presented in TABLE 10. All 62 heptane, octane and nonane isomers are ordered in ascending order according to the corresponding DM^k values. The order induced is in the increasing grade of branching, with small

TABLE 8. Intercorelation matrix of TI's DM^k.

TI	DM^1	DM^2	DM^3	DM^4	DM^5
DM^1	1.000	0.976	0.921	0.874	0.844
DM^2		1.000	0.982	0.954	0.932
DM^3			1.000	0.993	0.982
DM^4				1.000	0.998
DM^5					1.000

TABLE 9. Correlation matrix between TI's

$^0\chi$, $^1\chi$, $^2\chi$ and DM^k.

TI	DM^1	DM^2	DM^3	DM^4	DM^5
$^0\chi$	0.799	0.893	0.955	0.982	0.992
$^1\chi$	0.514	0.664	0.778	0.838	0.867
$^2\chi$	0.807	0.833	0.843	0.837	0.829

TABLE 10. Ordering of all isomers of heptane, octane and decane by DM^k.

No.	DM^1	DM^2	DM^3	DM^4	DM^5
	C_7H_{16}				
1	C_7	C_7	C_7	C_7	C_7
2	2-M-C_6	2-M-C_6	3-M-C_6	3-M-C_6	3-M-C_6
3	3-M-C_6	3-M-C_6	3-E-C_5	3-E-C_5	3-E-C_5
4	3-E-C_5	3-E-C_5	2-M-C_6	2-M-C_6	2-M-C_6
5	2,4-M_2-C_5	2,4-M_2-C_5	2,3-M_2-C_5	2,3-M_2-C_5	2,3-M_2-C_5
6	2,3-M_2-C_5	2,3-M_2-C_5	2,4-M_2-C_5	2,4-M_2-C_5	2,4-M_2-C_5
7	2,2-M_2-C_5	2,2-M_2-C_5	3,3-M_2-C_5	3,3-M_2-C_5	3,3-M_2-C_5
8	3,3-M_2-C_5	3,3-M_2-C_5	2,2-M_2-C_5	2,2-M_2-C_5	2,2-M_2-C_5
9	2,2,3-M_3-C_4	2,2,3-M_3-C_4	2,2,3-M_3-C_4	2,2,3-M_3-C_4	2,2,3-M_3-C_4
	C_8H_{18}				
1	C_8	C_8	C_8	C_8	C_8
2	2-M-C_7	2-M-C_7	3-M-C_7	3-E-C_6	3-M-C_7
3	3-M-C_7	3-M-C_7	4-M-C_7	4-M-C_7	4-M-C_7
4	4-M-C_7	4-M-C_7	3-E-C_6	3-M-C_7	3-E-C_6
5	3-E-C_6	3-E-C_6	2-M-C_7	2-M-C_7	2-M-C_7
6	2,5-M_2-C_6	2,5-M_2-C_6	2,4-M_2-C_6	3-E-2-M-C_5	3-E-2-M-C_5
7	2,4-M_2-C_6	2,4-M_2-C_6	2,3-M_2-C_6	2,3-M_2-C_6	3,4-M_2-C_6
8	2,3-M_2-C_6	2,3-M_2-C_6	3-E-2-M-C_5	3,4-M_2-C_6	2,3-M_2-C_6
9	3,4-M_2-C_6	3,4-M_2-C_6	3,4-M_2-C_6	2,4-M_2-C_6	2,4-M_2-C_6
10	3-E-2-M-C_5	3-E-2-M-C_5	2,5-M_2-C_6	2,5-M_2-C_6	2,5-M_2-C_6
11	2,2-M_2-C_6	2,2-M_2-C_6	3,3-M_2-C_6	3,3-M_2-C_6	3,3-M_2-C_6

TABLE 10. (Continued).

No.	DM[1]	DM[2]	DM[3]	DM[4]	DM[5]
12	3,3-M_2-C_6	3,3-M_2-C_6	2,2-M_2-C_6	3-E-3-M-C_5	3-E-3-M-C_5
13	2,3,4-M_3-C_5	3-E-3-M-C_5	3-E-3-M-C_5	2,2-M_2-C_6	2,2-M_2-C_6
14	3-E-3-M-C_5	2,3,4-M_3-C_5	2,3,4-M_3-C_5	2,3,4-M_3-C_5	2,3,4-M_3-C_5
15	2,2,4-M_3-C_5	2,2,4-M_3-C_5	2,2,3-M_3-C_5	2,2,3-M_3-C_5	2,2,3-M_3-C_5
16	2,2,3-M_3-C_5	2,2,3-M_3-C_5	2,2,4-M_3-C_5	2,3,3-M_3-C_5	2,3,3-M_3-C_5
17	2,3,3-M_3-C_5	2,3,3-M_3-C_5	2,3,3-M_3-C_5	2,2,4-M_3-C_5	2,2,4-M_3-C_5
18	2,2,3,3-M_4-C_4	2,2,3,3-M_4-C_4	2,2,3,3-M_4-C_4	2,2,3,3-M_4-C_4	2,2,3,3-M_4-C_4
	C_9H_{20}				
1	C_9	C_9	C_9	C_9	C_9
2	2-M-C_8	2-M-C_8	4-M-C_8	4-E-C_7	3-M-C_8
3	3-M-C_8	3-M-C_8	3-M-C_8	4-M-C_8	4-M-C_8
4	4-M-C_8	4-M-C_8	4-E-C_7	3-E-C_7	4-E-C_7
5	3-E-C_7	3-E-C_7	3-E-C_7	3-M-C_8	3-E-C_7
6	4-E-C_7	4-E-C_7	2-M-C_8	2-M-C_8	2-M-C_8
7	2,6-M_2-C_7	2,6-M_2-C_7	3,5-M_2-C_7	3-E-2-M-C_6	3-E-2-M-C_6
8	2,5-M_2-C_7	2,5-M_2-C_7	4-E-2-M-C_6	3,5-M_2-C_7	3,4-M_2-C_7
9	2,4-M_2-C_7	2,4-M_2-C_7	2,5-M_2-C_7	4-E-2-M-C_6	3,5-M_2-C_7
10	2,3-M_2-C_7	2,3-M_2-C_7	2,3-M_2-C_7	3,4-M_2-C_7	3-E-4-M-C_6
11	3,5-M_2-C_7	3,5-M_2-C_7	3-E-2-M-C_6	3-E-4-M-C_6	4-E-2-M-C_6
12	4-E-2-M-C_6	4-E-2-M-C_6	3,4-M_2-C_7	2,3-M_2-C_7	2,3-M_2-C_7
13	3,4-M_2-C_7	3,4-M_2-C_7	2,4-M_2-C_7	2,5-M_2-C_7	2,5-M_2-C_7
14	2,2-M_2-C_7	3-E-2-M-C_6	2,6-M_2-C_7	2,4-M_2-C_7	2,4-M_2-C_7

TABLE 10. (Continued).

No.	DM^1	DM^2	DM^3	DM^4	DM^5
15	3-E-2-M-C_6	3-E-4-M-C_6	3-E-4-M-C_6	2,6-M_2-C_7	2,6-M_2-C_7
16	3-E-4-M-C_6	2,2-M_2-C_7	3,3-M_2-C_7	3,3-M_2-C_7	3-E-3-M-C_6
17	3,3-M_2-C_7	3,3-M_2-C_7	4,4-M_2-C_7	3-E-3-M-C_6	3,3-M_2-C_7
18	2,3,5-M_3-C_6	4,4-M_2-C_7	2,2-M_2-C_7	4,4-M_2-C_7	4,4-M_2-C_7
19	4,4-M_2-C_7	2,3,5-M_3-C_6	3-E-3-M-C_6	2,2-M_2-C_7	3,3-E_2-C_5
20	2,2,5-M_3-C_6	3-E-3-M-C_6	2,3,5-M_3-C_6	3,3-E_2-C_5	2,2-M_2-C_7
21	3-E-3-M-C_6	2,3,4-M_3-C_6	3,3-E_2-C_5	2,3,4-M_3-C_6	2,3,4-M_3-C_6
22	2,3,4-M_3-C_6	2,2,5-M_3-C_6	2,3,4-M_3-C_6	2,3,5-M_3-C_6	3-E-2,4-M_2-C_5
23	3-E-2,4-M_2-C_5	3-E-2,4-M_2-C_5	3-E-2,4-M_2-C_5	3-E-2,4-M_2-C_5	2,3,5-M_3-C_6
24	2,2,4-M_3-C_6	2,2,4-M_3-C_6	2,2,3-M_3-C_6	2,2,3-M_3-C_6	3-E-2,2-M_2-C_5
25	3,3-E_2-C_5	3,3-E_2-C_5	2,2,4-M_3-C_6	3-E-2,2-M_2-C_5	3,3,4-M_3-C_6
26	2,4,4-M_3-C_6	2,2,3-M_3-C_6	2,4,4-M_3-C_6	2,3,3-M_3-C_6	2,3,3-M_3-C_6
27	2,2,3-M_3-C_6	2,4,4-M_3-C_6	2,2,5-M_3-C_6	2,4,4-M_3-C_6	2,2,3-M_3-C_6
28	2,3,3-M_3-C_6	2,3,3-M_3-C_6	2,3,3-M_3-C_6	3,3,4-M_3-C_6	2,4,4-M_3-C_6
29	3-E-2,2-M_2-C_5	3-E-2,2-M_2-C_5	3-E-2,2-M_2-C_5	2,2,4-M_3-C_6	3-E-2,3-M_2-C_5
30	3,3,4-M_3-C_6	3,3,4-M_3-C_6	3,3,4-M_3-C_6	3-E-2,3-M_2-C_5	2,2,4-M_3-C_6
31	3-E-2,3-M_2-C_5	3-E-2,3-M_2-C_5	3-E-2,3-M_2-C_5	2,2,5-M_3-C_6	2,2,5-M_3-C_6
32	2,2,3,4-M_4-C_5	2,2,3,4-M_4-C_5	2,2,3,4-M_4-C_5	2,2,3,4-M_4-C_5	2,2,3,4-M_4-C_5
33	2,2,4,4-M_4-C_5	2,3,3,4-M_4-C_5	2,3,3,4-M_4-C_5	2,3,3,4-M_4-C_5	2,3,3,4-M_4-C_5
34	2,3,3,4-M_4-C_5	2,2,4,4-M_4-C_5	2,2,4,4-M_4-C_5	2,2,3,3-M_4-C_5	2,2,3,3-M_4-C_5
35	2,2,3,3-M_4-C_5	2,2,3,3-M_4-C_5	2,2,3,3-M_4-C_5	2,2,4,4-M_4-C_5	2,2,4,4-M_4-C_5

variations between indices.

Following the rules of branching, we have determined that the ordering of structures induced by DM^k closely parallels these rules, with few exceptions concearning DM^3, DM^4 and DM^5.

The precision of only five decimal places of the reported data in TABLE 1 becomes insufficient in a comparison concerning the rules of branching for trees applied to dodecane isomers, so in TABLE 11 selected values of DM^k are presented with nine decimal figures.

After a careful inspection of the values of DM^k in TABLE 11 one can observe a clustering tendency for higher terms, i.e. for DM^4 and DM^5. For example, DM^4 for $i,i\text{-}M_2\text{-}C_{10}$ ranges between 9.88400 and 9.90723 and for $j,l\text{-}M_2\text{-}C_{10}$ ranges between 9.81621 and 9.83513. The following rules of branching are commented as defined in ref.16:

1. The influence of the total number of branches.

Rule 1.1. In a tree having a constant number of vertices, the branching increases when the number of branches attached to a given vertex increases owing to the chain decrease:

$$j\text{-}M\text{-}C_{11} < j,j\text{-}M_2\text{-}C_{10}$$

for $j=2,3,4,5$ and $k=1,2,3,4,5$

$$j,l\text{-}M_2\text{-}C_{11} < j,j,l\text{-}M_3\text{-}C_9$$

for $j=2$, $l=3,4,5,6,7,8$

$j=3$, $l=2,4,5,6,7,8$

TABLE 11. Selected values of DM^k for dodecane isomers.

No.	Name	DM^1	DM^2	DM^3	DM^4	DM^5
	$C_{12}H_{26}$					
1	C_{12}	24.413582088	11.904365753	10.096331102	9.531148172	9.300178393
2	2-M-C_{11}	25.768791693	12.107421704	10.251531692	9.674752701	9.443413486
3	3-M-C_{11}	26.194542605	12.123035169	10.246736580	9.670580166	9.441698937
4	4-M-C_{11}	26.434894580	12.131560130	10.246055632	9.670302868	9.4418023/9
5	5-M-C_{11}	26.537430900	12.134280336	10.245801669	9.670251139	9.441815838
6	6-M-C_{11}	26.602979976	12.136213821	10.245808630	9.670248135	9.441817747
7	2,2-M_2-C_{10}	29.232363935	12.585379671	10.529313742	9.907232937	9.660193088
8	2,3-M_2-C_{10}	28.431844671	12.415697643	10.416377077	9.820485520	9.589962340
9	2,4-M_2-C_{10}	28.264798019	12.409103692	10.418991162	9.824846969	9.593142854
10	2,5-M_2-C_{10}	28.139722858	12.391813750	10.415662889	9.823945838	9.592733489
11	2,6-M_2-C_{10}	27.939382289	12.378269610	10.414378471	9.823802877	9.592707833
12	2,7-M_2-C_{10}	27.757323180	12.373601331	10.414648653	9.823863691	9.592697624
13	2,8-M_2-C_{10}	27.552728657	12.363937556	10.414022692	9.823210585	9.592033890
14	2,9-M_2-C_{10}	27.124001418	12.355432373	10.429085149	9.835132363	9.598884134
15	3,3-M_2-C_{10}	30.394736852	12.651010044	10.505699872	9.882652128	9.644479305
16	3,4-M_2-C_{10}	29.173279841	12.475651533	10.418076343	9.817498633	9.58834897/5
17	3,5-M_2-C_{10}	28.847864542	12.439068478	10.410793072	9.816909249	9.588786903
18	3,6-M_2-C_{10}	28.505863045	12.416006390	10.409606336	9.817180993	9.588860082
19	3,7-M_2-C_{10}	28.223633860	12.400478583	10.408220068	9.816859812	9.588726464
20	3,8-M_2-C_{10}	27.948208127	12.393522572	10.411205778	9.818109681	9.588943001
21	4,4-M_2-C_{10}	31.010853733	12.704048140	10.511596292	9.884003864	9.6454888/0
22	4,5-M_2-C_{10}	29.565461148	12.495561235	10.416511803	9.816211914	9.587996499
23	4,6-M_2-C_{10}	29.023119662	12.458963349	10.413576407	9.817625806	9.589240348
24	4,7-M_2-C_{10}	28.599228386	12.423371108	10.409081823	9.816852190	9.588961631
25	5,5-M_2-C_{10}	31.342960323	12.718806114	10.511739150	9.883844275	9.645473137
26	5,6-M_2-C_{10}	29.647349943	12.499121086	10.416791578	9.816370659	9.588089473
27	3-E-C_{10}	26.784266639	12.156002110	10.247393821	9.670277829	9.442579796
28	4-E-C_{10}	27.082815814	12.167809019	10.246180751	9.669611822	9.442476547
29	5-E-C_{10}	27.223553407	12.171470417	10.245961958	9.669597243	9.442510732

TABLE 11. (Continued).

No.	Name	DM^1	DM^2	DM^3	DM^4	DM^5
30	2,2,3-M_3-C_9	33.396475297	13.104032602	10.732962695	10.061076682	9.809919424
31	2,2,4-M_3-C_9	32.528439086	13.082063346	10.749911517	10.080286281	9.823993533
32	2,2,5-M_3-C_9	31.802746277	12.985625553	10.729425084	10.074911661	9.821990552
33	2,2,6-M_3-C_9	31.275693160	12.933846054	10.722471025	10.073907098	9.821821920
34	2,2,7-M_3-C_9	30.962461721	12.914910540	10.718796595	10.071461915	9.820014475
35	2,2,8-M_3-C_9	30.592323821	12.919728384	10.751452432	10.097749289	9.837230678
36	2,3,3-M_3-C_9	34.129092556	13.222728239	10.754098548	10.059976925	9.805616967
37	3,3,4-M_3-C_9	35.010962915	13.357748441	10.773686818	10.057386416	9.801162876
38	3,3,5-M_3-C_9	33.706578981	13.197040092	10.734075217	10.051917864	9.803436633
39	3,3,6-M_3-C_9	32.770329487	13.063702457	10.707366288	10.046308198	9.801972837
40	3,3,7-M_3-C_9	32.167752218	13.019161629	10.704963158	10.046455015	9.801674756
41	3,3,8-M_3-C_9	31.695078325	12.978042227	10.711841456	10.058623823	9.810725396
42	2,4,4-M_3-C_9	33.682497095	13.228697481	10.765553650	10.073392065	9.816337834
43	3,4,4-M_3-C_9	35.432403963	13.405070545	10.780706881	10.057924509	9.801042718
44	4,4,5-M_3-C_9	35.718484353	13.436550823	10.784939588	10.058457112	9.801292494
45	4,4,6-M_3-C_9	34.203543615	13.297063593	10.765001656	10.062209873	9.807349661
46	4,4,7-M_3-C_9	33.191769879	13.106014405	10.713330703	10.047306813	9.802398576
47	4,4,8-M_3-C_9	32.429750931	13.035017234	10.720183884	10.061816451	9.812914947
48	2,5,5-M_3-C_9	33.131870847	13.133281931	10.739080772	10.065662796	9.813739895
49	3,5,5-M_3-C_9	34.303086203	13.298275527	10.760440006	10.059329607	9.805746172
50	4,5,5-M_3-C_9	35.893549960	13.472867418	10.795849533	10.061902386	9.802366701
51	2,3,4-M_3-C_9	32.075893452	12.922702123	10.621811975	9.977746374	9.741913333
52	2,3,5-M_3-C_9	31.222264056	12.817359416	10.604398336	9.977255424	9.743626111
53	2,3,6-M_3-C_9	30.596491163	12.753696033	10.596480427	9.976302457	9.743386964
54	2,3,7-M_3-C_9	30.197254246	12.741519503	10.600230124	9.977798171	9.743544084
55	2,3,8-M_3-C_9	29.756703054	12.700051337	10.603676545	9.986228177	9.749652398
56	3,4,5-M_3-C_9	32.873985817	13.049380325	10.645040528	9.980032739	9.741148692
57	3,4,6-M_3-C_9	31.816540544	12.891244650	10.609216278	9.972440321	9.739861016
58	3,4,7-M_3-C_9	31.112065201	12.824747070	10.605409227	9.974312979	9.740828682
59	3,4,8-M_3-C_9	30.551902538	12.756060470	10.597034988	9.976311677	9.743304924
60	4,5,6-M_3-C_9	33.104065289	13.075883879	10.647366361	9.979266216	9.740616234

TABLE 11. (Continued).

No.	Name	DM^1	DM^2	DM^3	DM^4	DM^5
61	4,5,7-M_3-C_9	31.967917984	12.913023335	10.613602390	9.973470185	9.740135867
62	2,5,6-M_3-C_9	31.102515851	12.796312145	10.599858079	9.976560849	9.743686202
63	2,4,5-M_3-C_9	31.576911782	12.855443972	10.609638194	9.978271906	9.744136321
64	2,4,6-M_3-C_9	30.750798663	12.802736023	10.612687958	9.985312955	9.748894985
65	2,4,7-M_3-C_9	30.209900220	12.727016939	10.595814474	9.980304173	9.746774509
66	2,4,8-M_3-C_9	29.669179281	12.700334408	10.612845702	9.996037420	9.756594925
67	2,5,7-M_3-C_9	30.361278673	12.744598034	10.598208602	9.980675252	9.746853536
68	2,5,8-M_3-C_9	29.678895553	12.703474930	10.612098567	9.995162309	9.755971527
69	3,5,7-M_3-C_9	31.063102921	12.813194152	10.597608504	9.972575036	9.741164965
70	3-E-2-M-C_9	29.337252397	12.476180120	10.414110293	9.816148577	9.588231621
71	3-E-3-M-C_9	31.767896722	12.815810061	10.524565261	9.879159454	9.640413674
72	3-E-4-M-C_9	30.111079749	12.571006600	10.427005576	9.817003956	9.588305182
73	3-E-5-M-C_9	29.461697314	12.499450214	10.413585410	9.814689468	9.587839885
74	3-E-6-M-C_9	28.962118242	12.452175282	10.408898910	9.814492434	9.587866488
75	3-E-7-M-C_9	28.533002649	12.432910790	10.410952444	9.816045602	9.588373727
76	3-E-8-M-C_9	28.112181042	12.390652545	10.407032063	9.817153870	9.589143471
77	4-E-2-M-C_9	29.064370427	12.452312381	10.412287924	9.818138603	9.589836823
78	4-E-3-M-C_9	30.163613375	12.565270292	10.425194195	9.816859946	9.588408776
79	4-E-4-M-C_9	32.517896901	12.879209091	10.529409591	9.878231405	9.639870981
80	4-E-5-M-C_9	30.468077805	12.597862171	10.427745020	9.816341575	9.588101712
81	4-E-6-M-C_9	29.661117111	12.511538548	10.414532319	9.814814251	9.588005197
82	4-E-7-M-C_9	29.014650915	12.450782071	10.408487996	9.814685650	9.588032079
83	4-E-8-M-C_9	28.440139387	12.403876907	10.407044219	9.817378542	9.589558192
84	5-E-2-M-C_9	28.771897827	12.421538239	10.408247924	9.817520301	9.589615311
85	5-E-3-M-C_9	29.563759897	12.495672762	10.412736330	9.814824943	9.587987775
86	5-E-4-M-C_9	30.517606612	12.591516838	10.426363303	9.816157587	9.588086589
87	5-E-5-M-C_9	32.746450136	12.901063749	10.533023042	9.879080291	9.640138044
88	4-P-C_9	27.419565576	12.184931559	10.246752837	9.669631684	9.442629570
89	5-P-C_9	27.494754749	12.188250782	10.246556347	9.669523573	9.442615318
90	4,4-E_2-C_8	34.156221653	13.160586619	10.595686725	9.893349444	9.643746065
91	4-M-4-P-C_8	33.164343606	12.973747912	10.550153374	9.883175721	9.641332701

$j=4,\ l=2,3,5,6,7$

and $k=1,2,3,4,5$

$j\text{-}E\text{-}C_{10} < j\text{-}E\text{-}j\text{-}M\text{-}C_9$

for $j=3,4,5$ and $k=1,2,3,4,5$

Rule 1.2. A tree is always more branched than a chain having the same number of vertices:

this is evident after an inspection of TABLE 11.

Rule 1.3. In a tree having a constant number of vertices, the branching increases when the number of branches connected to a given vertex increases owing to the decrease in length of a long branch attached to the same vertex:

$j\text{-}E\text{-}C_{10} < j,j\text{-}M\text{-}C_{10}$

for $j=3,4,5$ and $k=1,2,3,4,5$

$j\text{-}P\text{-}C_9 < j\text{-}E\text{-}j\text{-}M\text{-}C_9$

for $j=4,5$ and $k=1,3,4,5$

Rule 1.4. Among the trees having the same number of vertices and different number of branches of the same size attached to different vertices, the greater the number of branches, the greater the branching:

$i\text{-}M\text{-}C_{11} < i,j\text{-}M_2\text{-}C_{10} < i,j,l\text{-}M_3\text{-}C_9$

for all $i \neq j \neq l$ and $k=1,2,3,4,5$

2. The influence of the chain length.

Rule 2.1. The branching decreases when at a constant number of

vertices the chain length increases at the cost of some branch length decrease:

$i\text{-}E\text{-}C_{10} > i\text{-}M\text{-}C_{11}$

for i=3,4,5 and k=1,2,3,5

$i\text{-}P\text{-}C_9 > i\text{-}E\text{-}C_{10}$

for i=4,5 and k=1,2,3,5

3. The influence of the branch length.

Rule 3.1. In graphs having a constant number of vertices, the branching decreases when some branch increases its length at the cost of the decrease of the length of another branch having the same or smaller size than the first branch:

$4,4\text{-}E\text{-}C_8 > 4\text{-}M\text{-}4\text{-}P\text{-}C_8$

for k=1,2,3,4,5

4. The influence of the branch position.

Rule 4.1. In graphs having a constant number of vertices as well as a constant number and length of branches, the branching increases with the displacement of the branch from the terminal to the middle vertex of the chain:

$2\text{-}M\text{-}C_{11} < 3\text{-}M\text{-}C_{11} < 4\text{-}M\text{-}C_{11} < 5\text{-}M\text{-}C_{11} < 6\text{-}M\text{-}C_{11}$

for k=1,2

$3\text{-}E\text{-}C_{10} < 4\text{-}E\text{-}C_{10} < 5\text{-}E\text{-}C_{10}$

for k=1,2

$4\text{-}P\text{-}C_9 < 5\text{-}P\text{-}C_9$

for k=1,2

5. The influence of the number of branches attached to a given vertex.

Rule 5.1. In a tree having a constant number of vertices and branches, the branching increases when a branch is displaced from one vertex to another branched one, having equivalent position in the chain (thereby instead of two vertices with degree v we obtain two vertices with degrees v+1 and v-1, respectively):

$$2,9\text{-}M_2\text{-}C_{10} < 2,2\text{-}M_2\text{-}C_{10}$$
$$3,8\text{-}M_2\text{-}C_{10} < 3,3\text{-}M_2\text{-}C_{10}$$
$$4,7\text{-}M_2\text{-}C_{10} < 4,4\text{-}M_2\text{-}C_{10}$$
$$5,6\text{-}M_2\text{-}C_{10} < 5,5\text{-}M_2\text{-}C_{10}$$

for k=1,2,3,4,5

CONCLUDING REMARKS

The analysis of the spectrum of indices DM^k is favourable, being the basis of further studies. The calculation of the indices, based on a spectrum of connectivity indices, is straightforward and the computer time needed is very small. The new indices have no degenerate value for the set of 661 alkane isomers between butane and dodecane, thus eliminating one of the main drawbacks of connectivity indices. Moreover, they express in a proper way the structural information contained in the molecular graph and we expect that they could be used as good correlational parameters. Work is in progress in this direction.

DM^1 has the largest range of variation for a given set of isomeric alkanes, and the largest overlap with other sets.

DM^1 and DM^2 order alkanes according to generally accepted increasing branching (2-Me < 3-Me < 4-Me... alkanes), unlike DM^3-DM^5.

REFERENCES

1. O. Ivanciuc, Rev. Roum. Chim., 34, 1361 (1989).
2. K. Varmuza, "Pattern Recognition in Chemistry", Lecture Notes in Chemistry 21, Springer, Berlin, 1980.
3. L.R. Crawford and J.D. Morrison, Anal. Chem. 40, 1469 (1968).
4. J.B. Justice, Jr., and T.L. Isenhour, Anal. Chem. 46, 223, (1974).
5. A. Mehlhorn, F. Fratev, O.E. Polansky and V. Monev, Math. Chem. 15, 3 (1984).
6. A.T. Balaban, I. Motoc, D. Bonchev and O. Mekenyan, Top. Curr. Chem. 114, 21 (1983), Eds. M. Charton and I. Motoc.
7. N. Trinajstic, "Chemical Graph Theory", CRC Press, Boca Raton, 1983, vol. 2, chapter 4.
8. M. Randic, J. Am. Chem. Soc. 97, 6609 (1975).
9. L.B. Kier, L.H. Hall, W.J. Murray and M. Randic, J. Pharm. Sci. 64, 1971 (1975).
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, J. Pharm. Sci. 72, 1170 (1983).
12. M. Razinger, J.R. Chretien and J.E. Dubois, J. Chem. Inf. Comput. Sci. 25, 23 (1985).
13. P.A. Filip, T.-S. Balaban and A.T. Balaban, J. Math. Chem. 1, 61 (1987).
14. D. Bonchev and N. Trinajstic, J. Comput. Chem. 2, 127 (1981).
15. O. Onicescu, C.R. Acad. Sci. Paris, Serie A, 263, 841 (1966).
16. D. Bonchev and N. Trinajstic, J. Chem. Phys. 67, 4517 (1977).