

Isomer Discrimination by Topological Information Approach

D. Bonchev, Ov. Mekenyan, and N. Trinajstić*

Higher School of Chemical Technology, 8010 Burgas, Bulgaria

Received September 11, 1980

A comparative analysis of ten topological indices is made. No index is found to discriminate isomers uniquely. A combined topological index, named the superindex, consisting of a number of topological indices is proposed. Information theory is applied to express all components of the superindex on a common quantitative scale. The superindex is tested on the sample of 427 graphs consisting of all acyclic, monocyclic, and bicyclic graphs with 4–8 vertices.

INTRODUCTION

The interest in the characterization of chemical structures by means of algebraic graph theory¹ has greatly increased during the last decade.^{2–7} A graph theoretical characterization of molecules is usually realized via the topological indices.^{8,9} A topological index is a numerical characteristic of a molecule based on a certain topological feature of the corresponding graph. In other words, a topological index attempts to express numerically topological information for a given chemical compound. Thus, for example, the index introduced by Gutman et al.¹⁰ and the connectivity index of Randić¹¹ are related to the topological valency (degree) of the graph vertices. Lovasz and Pelikan¹² regard as an index the largest eigenvalue in the graph spectrum. Hosoya and co-workers^{13,14} have introduced a topological index Z on the basis of a class of graph decompositions. Several indices originate from the distance matrix of a graph: the Wiener number¹⁵ W , the information indices for the distribution of distances,¹⁶ etc. The center of the graph is also taken into consideration for topological characterization of molecules.^{17–20}

Besides their application to various structure–property and structure–activity correlations,^{9,11,15,21–25} the topological indices are often used for the discrimination of isomers.^{8,11,13,16,17} The latter is of importance for the coding and the computer processing of chemical structures. For the computer storing and retrieval of molecular systems an index is needed which uniquely characterizes a given structure. However, as will be shown in what follows the known topological indices fail to specify molecules uniquely. Starting from this point a new practical approach to isomer discrimination is proposed in the present work. The gist of this approach is in using simultaneously a number of indices in order to provide a more complete topological characterization of molecule. Information theory^{26,27} is also applied as a means for expressing all topological indices under the consideration on a unique quantitative scale. The combined index constructed in such a way is called a *topological information superindex*. Clearly, the more complete the topological characterization of isomers, the better is their discrimination. In this article we give a simple version of the superindex and discuss the possibilities for a further increase of its discriminating power.

* Present address: The Rugjer Bošković Institute, P.O.B. 1016, 41001 Zagreb, Croatia, Yugoslavia.

DISCUSSION OF SOME TOPOLOGICAL INDICES

All the graphs discussed in this article are connected graphs.

The adjacency matrix of a graph G , $A(G)$, is the most commonly used graph theoretical matrix.²⁸ It describes the connectivity in a given molecular graph (structure). Its entries $A_{ij} = A_{ji}$ equal unity if vertices i and j are adjacent (i.e., linked by an edge), and zero otherwise. Thus, the set of adjacency matrix elements consists of two subsets: one containing the number of zeros, N_0 , and the other containing ones, N_1 , respectively. The normalized information index for the adjacency relationship in the molecular graph, \bar{I}_{adj} , can be defined on the basis of the Shannon equation.²⁶

The information content of a system, I , having N elements is defined by²⁷

$$I = N \log_2 N - \sum_{i=1}^n N_i \log_2 N_i \quad (1)$$

where n is the number of different sets of elements and N_i is the number of the elements in the i th set of elements. The summation is over all sets of elements. The logarithm is taken at base 2 for measuring the information content in bits. The mean information content of one element of the system, \bar{I} , is defined by means of the total information content or by the Shannon relation,²⁶

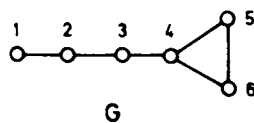
$$\bar{I} = \frac{I}{N} = - \sum_{i=1}^n p_i \log_2 p_i, \text{ bits per element} \quad (2)$$

where $p_i = N_i/N^2$. If we introduce in eq. (2) $p_1 = N_1/N^2$ and $p_0 = N_0/N^2$ as the probabilities of a randomly selected entry to signify or not signify, respectively, adjacency, the expression for \bar{I}_{adj} is obtained:

$$\bar{I}_{\text{adj}} = - \frac{N_0}{N^2} \log_2 \left(\frac{N_0}{N^2} \right) - \frac{N_1}{N^2} \log_2 \left(\frac{N_1}{N^2} \right), \text{ bits} \quad (3)$$

Example:

$$A(G) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$



$$\begin{aligned} N &= 6 \\ N^2 &= 36, \\ N_1 &= 12, \\ N_0 &= 24 \end{aligned}$$

Adjacency distribution: $P_A = N^2\{N_1, N_0\} = 36\{12, 24\}$, $\bar{I}_{\text{adj}} = 0.9813$ bits

The distance matrix of a graph G , $D(G)$, is a symmetric matrix²⁸ and its entries $D_{ij} = D_{ji}$ are integers called distances. D_{ij} is equal to the number of edges joining by the shortest path the vertices i and j . Clearly, $D_{ij} = 1$ if the vertices i and j are adjacent while $D_{ij} > 1$ for i and j being non-adjacent vertices. The half-sum of all entries in the distance matrix [or the sum of all entries in the triangular off-diagonal submatrix of $D(G)$] represents the total distance in the graph. It is usually called the Wiener number $W(G)$, since Wiener¹⁵ was the first to apply this number (though defining it empirically as the sum of all bonds connecting any pair of atoms in a given acyclic molecule) to various structure-property correlations. The Wiener number is defined as follows:

$$W(G) = \frac{1}{2} \sum_{i,j=1}^N D_{ij} \quad (4)$$

Regarding distances D_{ij} in the triangular D submatrix as being distributed into n subsets according to their magnitude, a finite probability scheme can be constructed to serve as a basis for the application of the Shannon equation (2),²⁹

Distances	1	2	3	...	n
Number of					
distances	N_1	N_2	N_3	...	N_n
Probabilities	p_1	p_2	p_3	...	p_n

$$\sum_{i=1}^n p_i = 1 \quad (6)$$

$$\sum_{i=1}^n N_i = N(N-1)/2 \quad (7)$$

Another, rather convenient expression for the Wiener number is

$$W(G) = \sum_{i=1}^n i \cdot N_i \quad (8)$$

The normalized information index on the distribution of distances in the graph according to their equality or inequality, \bar{I}_D^E , can be now defined¹⁶ as

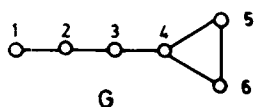
$$\bar{I}_D^E = - \sum_{i=1}^n \frac{2N}{N(N-1)} \log_2 \left(\frac{2N_i}{N(N-1)} \right), \quad \text{bits per distance} \quad (9)$$

The information index on the distribution of distances in the graph (molecule) according to their magnitude, \bar{I}_D^W , can also be given¹⁶ in bits per unit distance,

$$\bar{I}_D^W = - \sum_{i=1}^n n_i \frac{i}{W} \log_2 \left(\frac{i}{W} \right), \quad \text{bits} \quad (10)$$

Example:

$$\mathbf{D}(G) = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 4 \\ 1 & 0 & 1 & 2 & 3 & 3 \\ 2 & 1 & 0 & 1 & 2 & 2 \\ 3 & 2 & 1 & 0 & 1 & 1 \\ 4 & 3 & 2 & 1 & 0 & 1 \\ 4 & 3 & 2 & 1 & 1 & 0 \end{bmatrix}$$



$$W = 6 \times 1 + 4 \times 2 + 3 \times 3 + 2 \times 4 = 31$$

The distribution of distances:

$$(a) P_D^E = \frac{N(N-1)}{2} \{N_1, N_2, N_3, N_4\} = 15 \{6, 4, 3, 2\}$$

$$(b) P_D^W = W \{1 \times N_1, 2 \times N_2, 3 \times N_3, 4 \times N_4\} = 31 \{1 \times 6, 2 \times 4, 3 \times 3, 4 \times 2\}$$

$$\bar{I}_D^E = 1.8892 \text{ bits}$$

$$\bar{I}_D^W = 3.7198 \text{ bits}$$

The degrees (valencies) of the graph vertices v_i ($i = 1, 2, \dots, N$) equal the number of edges incident with the vertex i . The molecular connectivity index χ_R of Randić¹¹ results from the contributions of all M edges in the molecular graph,

$$\chi_R = \sum_{\substack{i,j=1 \\ i \neq j}}^M (v_i v_j)^{-1/2} = \sum_{i=1}^M \chi_i \quad (11)$$

where v_i and v_j denote the degrees of the two end points of an edge, and χ_i is the partial connectivity index of edge i . In the graph G , presented above as an example, the partial and molecular connectivity indices, respectively, are as follows:

$$\chi_{1-2} = 0.7071, \quad \chi_{2-3} = \chi_{5-6} = 0.5000, \\ \chi_{3-4} = \chi_{4-5} = \chi_{4-6} = 0.4083, \quad \text{and } \chi_R = 2.9320$$

A probability scheme can be constructed for the edges of the graph by partitioning them into n subsets depending on the equality of their partial connectivity indices χ_i . Making use of eq. (2) an information index on the edges distribution in the graph according to their equivalence, I_{edge}^E , can be defined as

$$\bar{I}_{\text{edge}}^E = - \sum_{i=1}^n \frac{M_i}{M} \log_2 \left(\frac{M_i}{M} \right), \quad \text{bits} \quad (12)$$

where M_i is the number of edges having the same partial connectivity index χ_i . Note that $M = \sum_{i=1}^n M_i$.

Example:

$$P_{\text{edge}}^E = M \{M_1, M_2, M_3\} = 6 \{1, 2, 3\}$$

$$\bar{I}_{\text{edge}}^E = 1.4592 \text{ bits}$$

The number of ways $p(G;k)$ in which k may be selected from all M edges of graph G so that no two of them are adjacent is a graph invariant. The Hosoya index¹³ Z is defined for any graph G as

$$Z = \sum_{k=0}^{[N/2]} p(G;k) \quad (13)$$

where $N/2$ in Gauss' square brackets represents the largest whole number in the quotient $N/2$. The distribution of this index into different numbers $p(G;k)$, $P_Z = Z \{p(G;0), p(G;1), p(G;2), \dots, p(G;[N/2])\}$, may be used to define the normalized information index on the Hosoya graph decompositions, \bar{I}_Z ,

$$\bar{I}_Z = - \sum_{k=0}^{[N/2]} \frac{p(G;k)}{Z} \log_2 \left(\frac{p(G;k)}{Z} \right), \quad \text{bits} \quad (14)$$

Since for an acyclic graph Z is alternatively defined as the sum of the absolute values of coefficients in its characteristic polynomial, \bar{I}_Z coincides in this case with the previously defined information¹⁶ for polynomial coefficients of the graph, \bar{I}_{PC} .

For instance, for the graph G discussed above,

$$P_Z = 16 \{1, 6, 8, 1\}$$

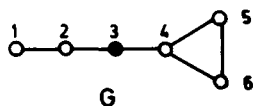
$$\bar{I}_Z = 1.5306 \text{ bits}$$

Balaban¹⁷ has introduced a topological *centric index* for trees. A *center* of a graph³⁰ is a vertex v_j with the property that the maximum possible distance between v_j and any other vertex in a graph is as small as possible. The maximum distance between v_j and the most distant vertex is called the radius of the graph and is denoted by r_j . All other vertices in the graph have their radii r_i larger than r_j . The set of graph vertices $\{N\}$ may be partitioned into subsets of vertices having the same radius (a radial partition). Here we discuss solely a normalized information centric index on the radial distribution of graph vertices, \bar{I}_C ,

$$\bar{I}_C = - \sum_{i=1}^n \frac{N_i}{N} \log_2 \left(\frac{N_i}{N} \right), \text{ bits per vertex} \quad (15)$$

where N_i is the number of vertices having the same radius r_i .

Example: The center of the graph G is vertex 3 whose distance is $r_3 = 2$, while the radii of the other vertices are, respectively, $r_2 = r_4 = 3$ and $r_1 = r_5 = r_6 = 4$.

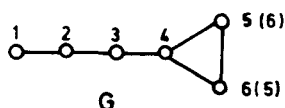


$$P_R = N\{N_1, N_2, N_3\} = 6\{3, 2, 1\}$$

$$\bar{I}_C = 1.4592 \text{ bits}$$

The orbits of the automorphism group give other important graph characteristics. Two vertices are in the same graph orbit if there is a permutation in the automorphism group (permutation preserving adjacency) that maps one onto the other. Denoting the cardinality of the i th orbit by N_i one can apply the Shannon equation in the form of relation (15) for the calculation of the *topological information*, or more precisely, *information on the graph orbits*, \bar{I}_{orb} , as it was done 25 years ago by Rashevsky,³¹ Trucco,³² and others.

Example: In the graph G one can see that (5, 6) is the only permutation which preserves adjacency.



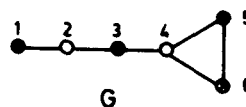
$$P_{\text{orb}} = N\{N_1, N_2, N_3, N_4, N_5\} = 6\{1, 1, 1, 1, 2\}$$

$$\bar{I}_{\text{orb}} = 2.2516 \text{ bits}$$

A *coloring* of a graph G is an assignment of m colors to the vertices of G in such a way that no two adjacent vertices have the same color. A chromatic decomposition into color classes can be also obtained for the graph vertices. However, this decomposition is not, in general, unique. Mowshovitz³³ has defined a unique information measure, i.e., *chromatic information content* of a graph, \bar{I}_{chr} , which reflects the chromatic structure of a graph minimizing the Shannon function over a certain class of chromatic decompositions V ,

$$\bar{I}_{\text{chr}} = \min_V \left\{ - \sum_{i=1}^n \frac{N_i}{N} \log_2 \left(\frac{N_i}{N} \right) \right\} \quad (16)$$

The coloring in three colors of the graph G discussed above, by minimizing \bar{I}_{chr} , is shown below:



$$P_{\text{chr}} = N\{N_1, N_2, N_3\} = 6\{1, 2, 3\}$$

$$\bar{I}_{\text{chr}} = 1.4592 \text{ bits}$$

A COMPARATIVE ANALYSIS OF THE DISCRIMINATION POWER OF TOPOLOGICAL INDICES

In order to study the applicability of the diverse topological indices for the isomers' discrimination, we made extensive calculations on 45 acyclic and 382 mono- and bicyclic graphs having 4–8 vertices (see Fig. 1).

The discrimination power of the information index on adjacency is very low. At a constant number of vertices it can distinguish only between the acyclic, monocyclic, and bicyclic graphs. Due to this low discrimination efficiency the information index \bar{I}_{adj} is excluded from the comparative analysis of topological indices. The comparison between the remaining ten topological or topological information indices is best specified by means of the *mean isomer degeneracy* \bar{i} ,

$$\bar{i} = N_{\text{isomers}}/k \quad (17)$$

where N_{isomers} is the number of isomers (graphs with the same number of vertices) and k is the number of distinct values of the topological index

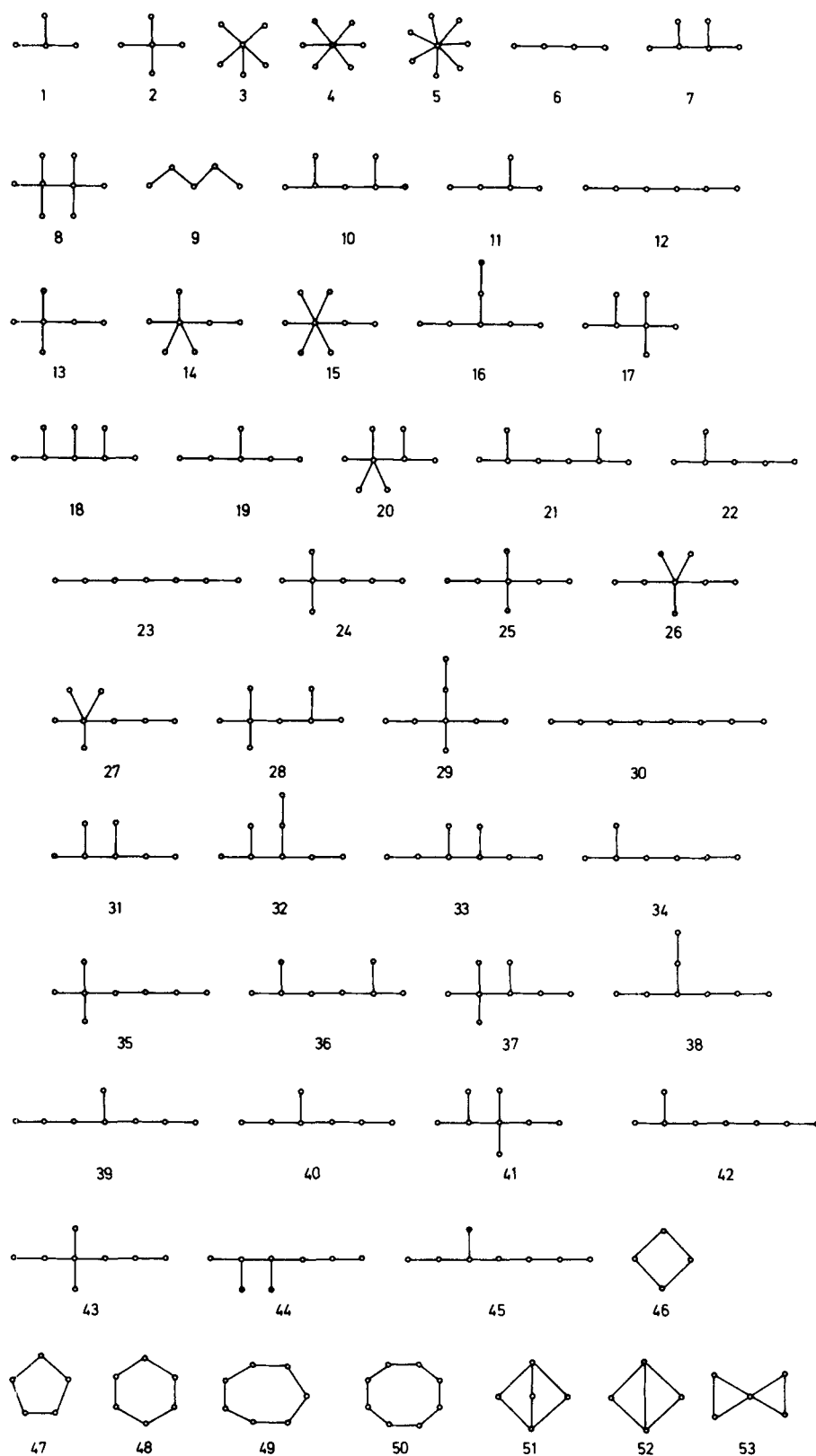


Figure 1. All acyclic, monocyclic, and bicyclic graphs with 4–8 vertices.

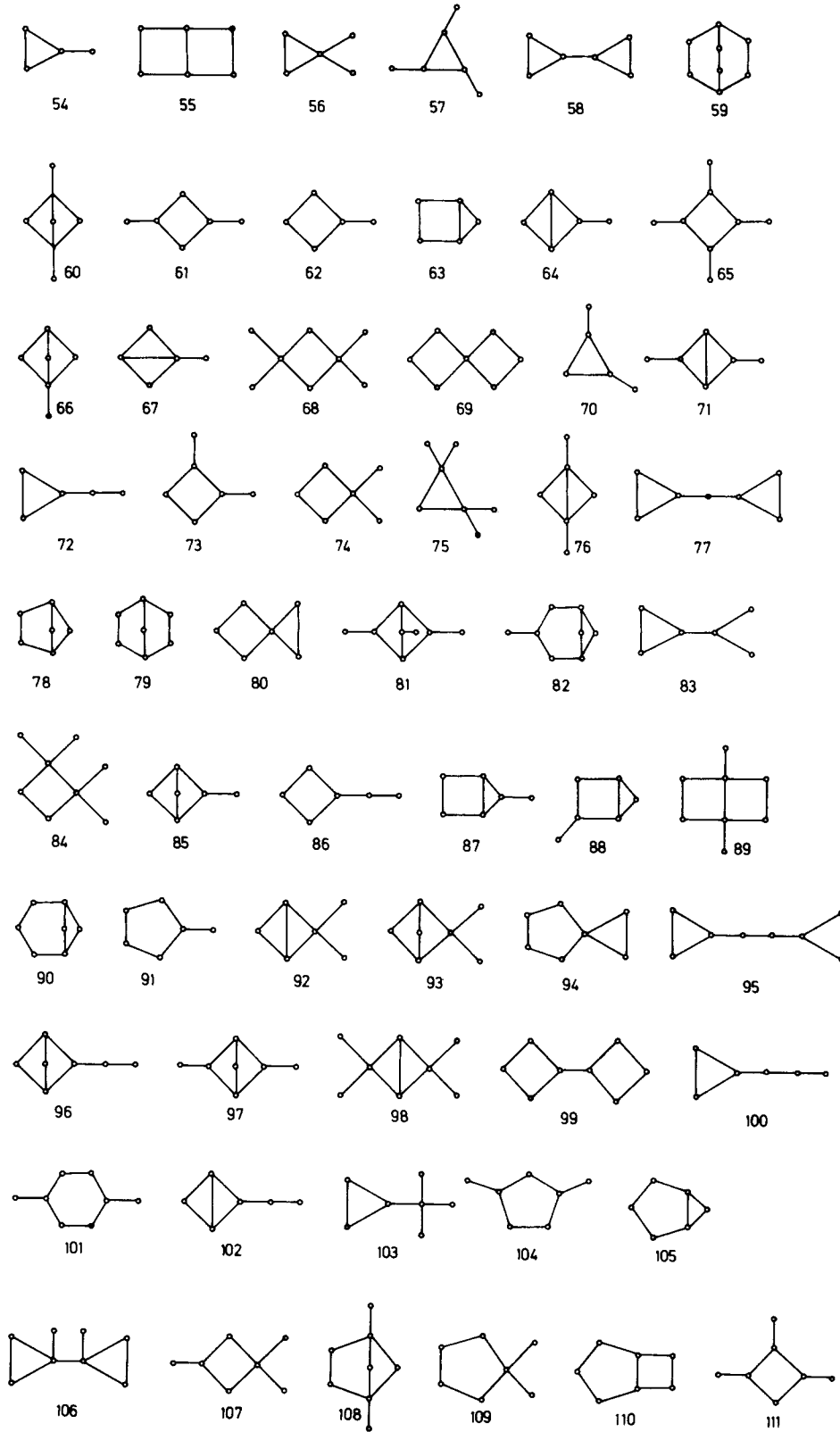
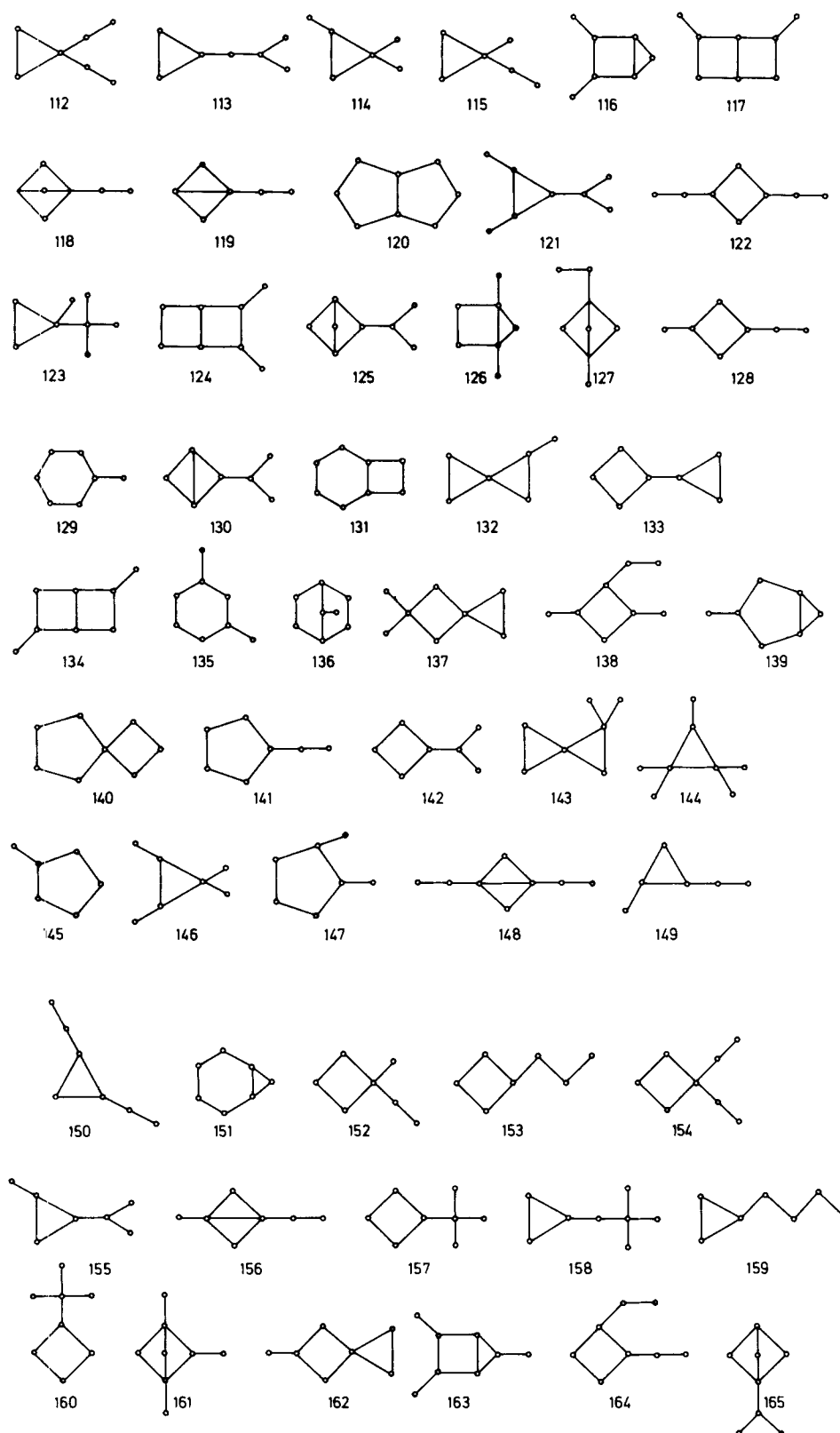


Figure 1 (continued from previous page)

**Figure 1** (continued from previous page)

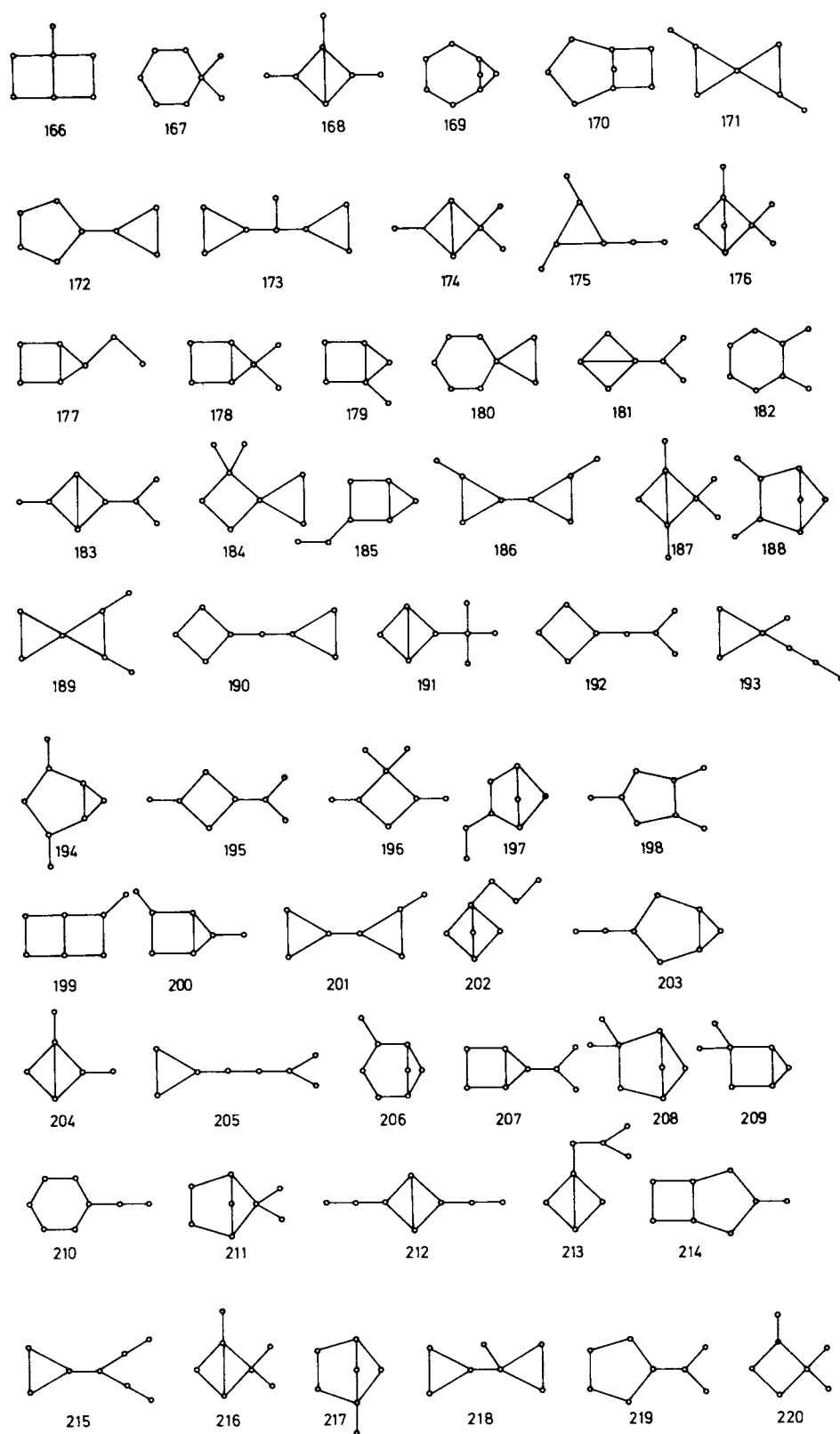
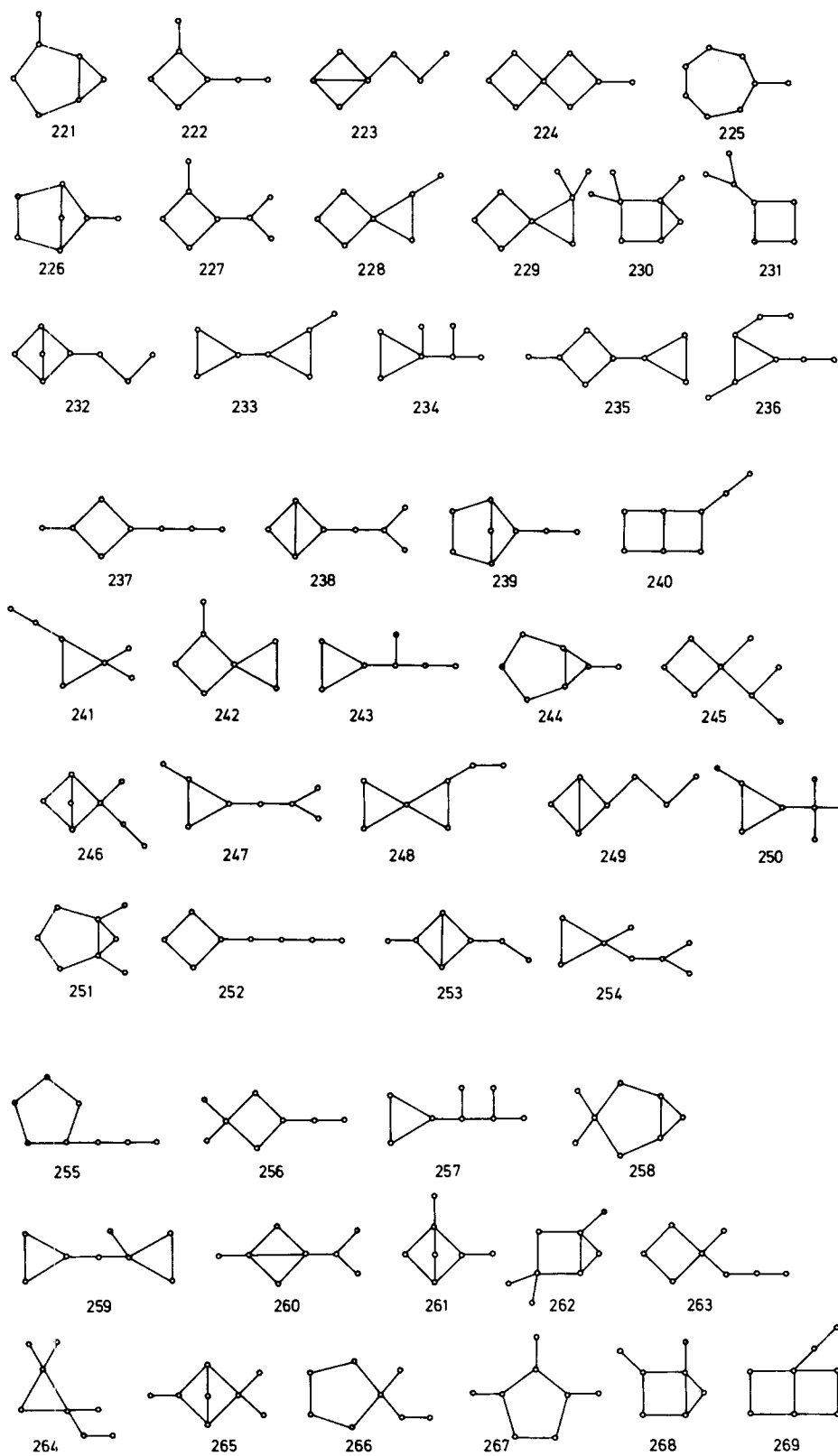


Figure 1 (continued from previous page)

**Figure 1** (continued from previous page)

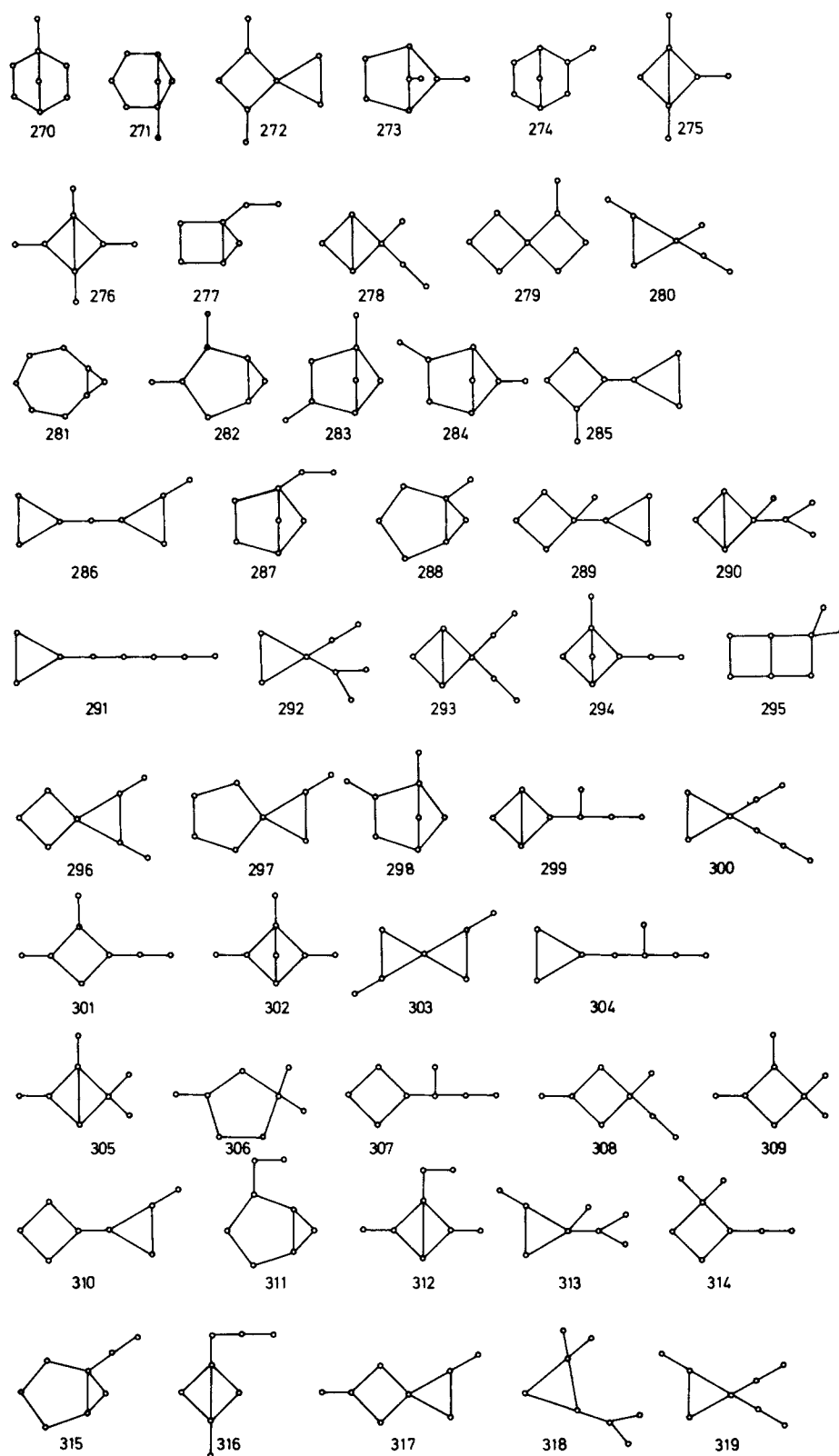
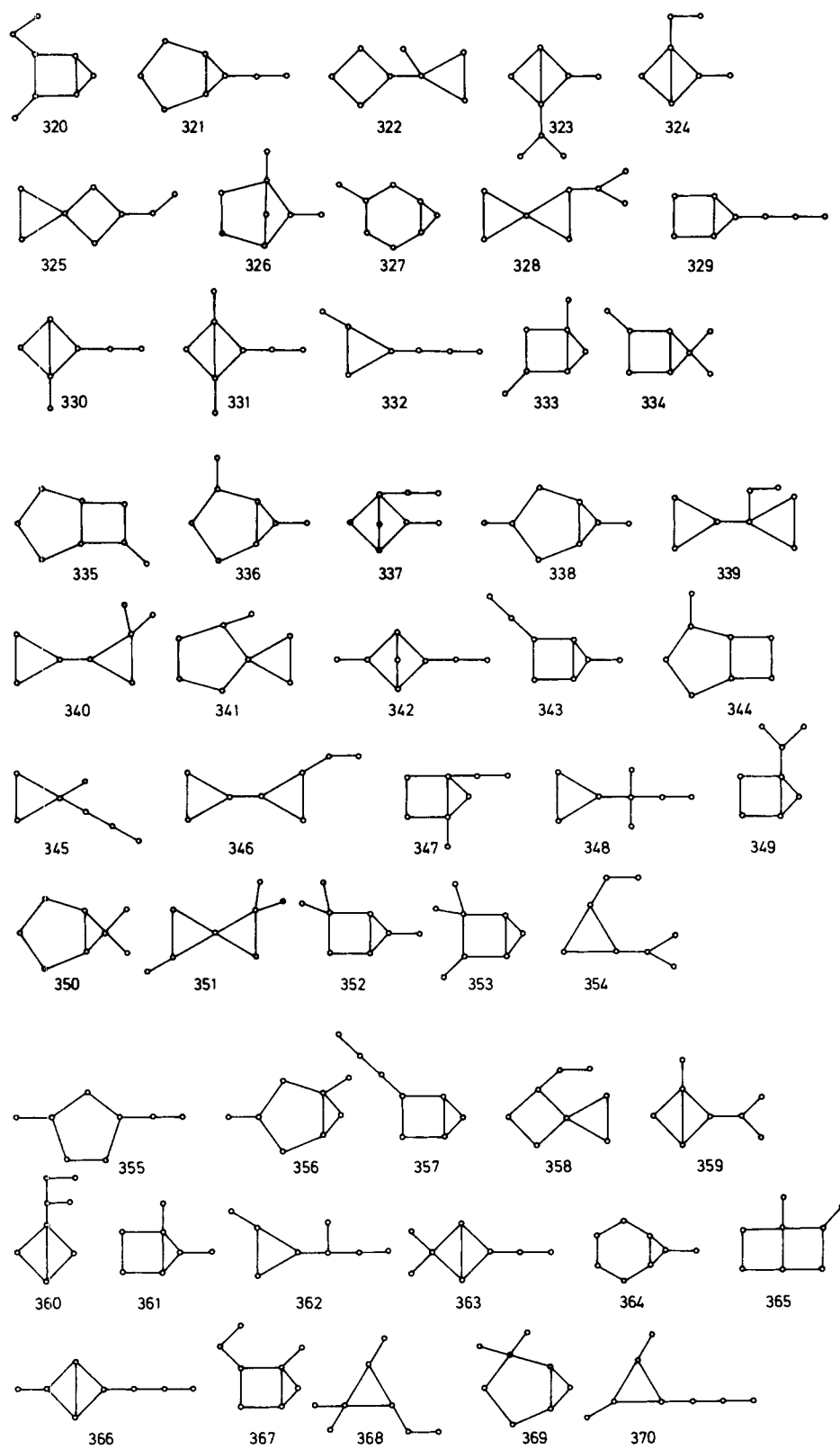


Figure 1 (continued from previous page)

**Figure 1** (continued from previous page)

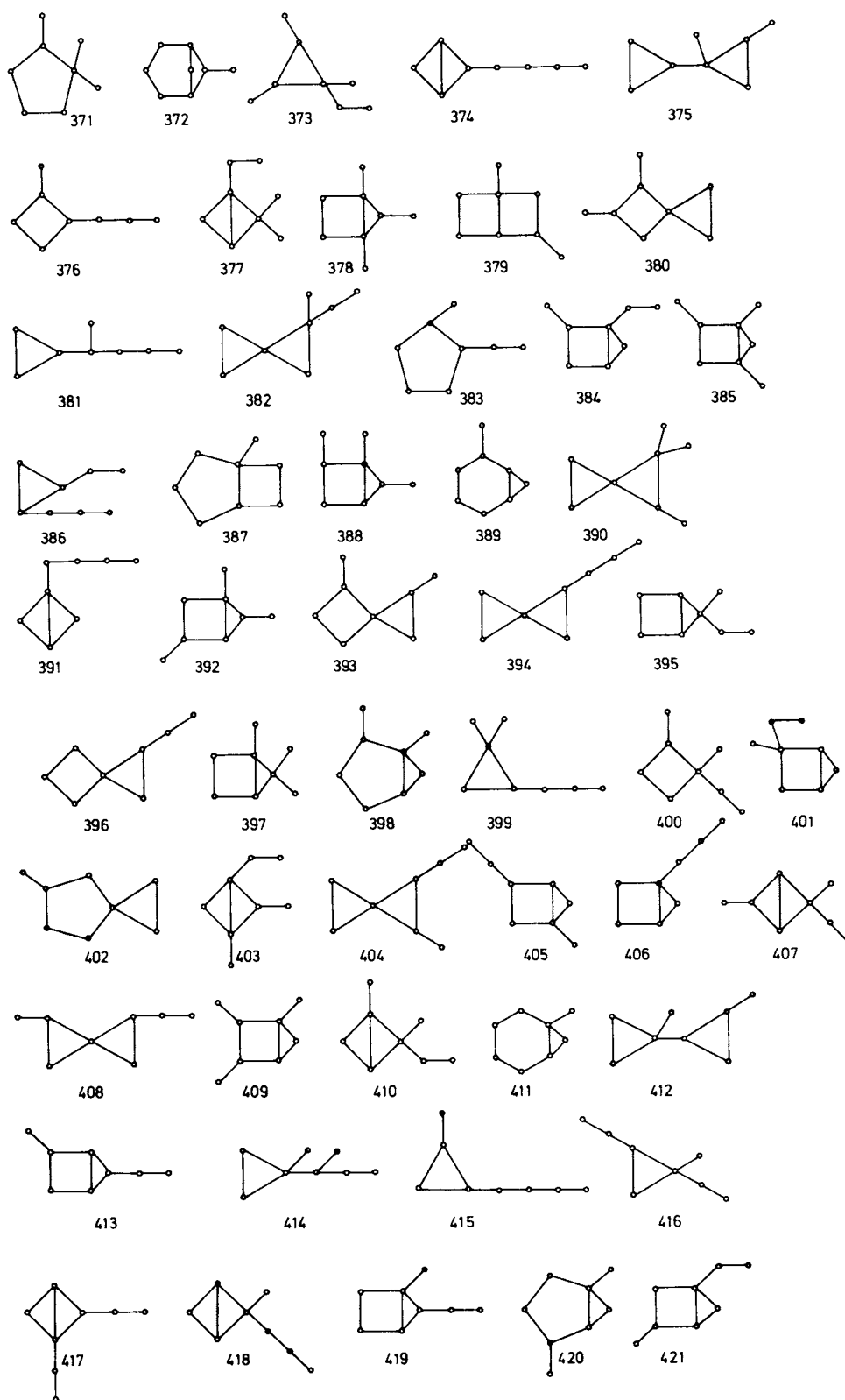


Figure 1 (continued from previous page)

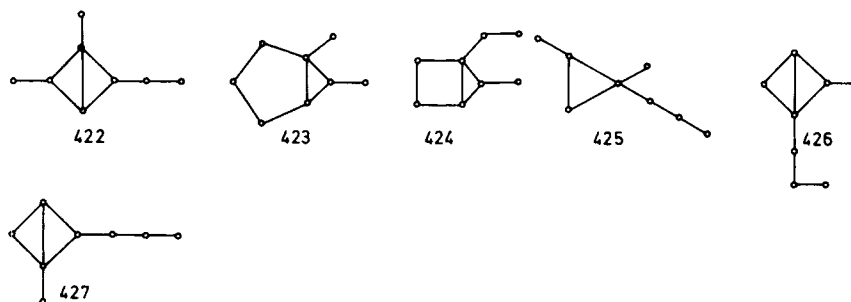


Figure 1 (continued from previous page)

under examination. The average degeneracies of the sets of isomers, evaluated for ten topological indices, are given in Table I.

The results obtained differ considerably for acyclic and cyclic graphs. In the case of acyclic graphs even for 23 graphs with 8 vertices there are four information indices (\bar{I}_Z , \bar{I}_{edge}^E , \bar{I}_D^E , and \bar{I}_D^W) that uniquely characterize the individual graphs. The Randić index ($\bar{i} = 1.09$) and the Wiener number ($\bar{i} = 1.15$) in this case also indicate a very high discriminating power. The lowest discriminating power is exhibited by the chromatic information index ($\bar{i} = 5.75$) having only four different values for 23 isomers structures with $N = 8$. The examination of the mean degeneracies of cyclic graphs demonstrates that *all* indices considered fail to

discriminate uniquely sets of isomers having 6, 7, and 8 atoms (their cardinalities are 29, 85, and 255, respectively). Here again the lowest discrimination power is shown by the chromatic information index. Similarly, a poor result is obtained by the centric and orbit information indices. The best discrimination power is exhibited by the connectivity index of Randić and then by the topological information index for the magnitudes of the topological distances \bar{I}_D^W .

Some general conclusion can be drawn from this comparative analysis of the topological indices. Evidently, the indices based on the vertex, edge, or distance equality have lesser discrimination power than the corresponding indices based on their magnitude (\bar{I}_C , \bar{I}_{chr} , \bar{I}_{orb} ; \bar{I}_{edge}^E as compared

Table I. Mean degeneracy of 10 topological or topological information indices calculated for 427 acyclic, monocyclic, and bicyclic graphs having 4–8 vertices.

Number of vertices in the isomer set	W	\bar{I}_D^E	\bar{I}_D^W	Z	\bar{I}_Z	χ_R	\bar{I}_{EDGE}^E	\bar{I}_C	\bar{I}_{CHR}	\bar{I}_{ORB}
Acyclic Graphs										
4	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1.5	1
6	1	1	1	1	1	1	1	1	2	1.2
7	1.22	1	1	1	1	1	1	1.8	3.7	1.2
8	1.15	1	1	1.28	1	1.09	1	2.3	5.8	1.5
Cyclic Graphs										
4	1.5	1.5	1.5	1	1	1	1	1	1.5	1
5	2.5	2.0	2.0	1.7	1.4	1	1.4	2.0	3.3	2
6	3.6	2.4	2.2	2.6	1.5	1.2	1.6	4.1	5.8	3.6
7	6.0	2.8	2.4	4.5	2.2	1.6	3.4	12.1	12.1	9.4
8	10.6	3.5	2.8	7.7	2.2	1.6	6.0	16.0	36.4	18.2

with χ_R ; \bar{I}_D^E as compared with \bar{I}_D^W). In addition, an individual topological index, based on summation of magnitudes, has in general a smaller number of distinct values than its information analog. Such representative pairs of topological information indices are, for example, Z/\bar{I}_Z and W/\bar{I}_D^W . However, the major conclusion which can be made from the analysis of results given in Table I is that *no topological, or topological information, index of a graph is able to discriminate uniquely the isomeric structures*. Neither such a carefully selected index as the Randić connectivity index χ_R nor the index \bar{I}_D^W based on a very detailed scrutiny for the optimal distance distribution can solve the problem. Admittedly the prime target of the most topological indices introduced up to now was the search for structure–property correlations. However, now it is evident that the development of other ways of approaching the problem are necessary for achieving unique isomer discrimination. One of the possible ways is developed in the present work by constructing a topological superindex which is simultaneously based on a number of topological indices.

THE TOPOLOGICAL INFORMATION SUPERINDEX FOR DISCRIMINATION OF ISOMERS

It should be stated at this point once again that the goal of this article is not to solve the isomorphism problem. Here we aim at a good *practical* solution of the problem of isomer discrimination, a solution that could be efficient for the majority of chemical structures though insufficient for the solution of the isomorphism problem. Due to this, we deal with acyclic, monocyclic, and bicyclic graphs as good examples from chemical practice though in general the isomorphism problem for these graphs is not so difficult. In addition, the numerical values of the topological indices for these graphs could be used in various structure–property correlations.

The following topological information indices are selected as possible elements for the construction of the superindex SI: \bar{I}_D^W (representing the distance matrix of a graph decomposition), \bar{I}_{edge}^E (representing the degree of graph vertices),

\bar{I}_{orb} (representing the graph orbits), \bar{I}_{chr} (representing the chromatic number of a graph), \bar{I}_Z (representing the Hosoya graph decomposition), and \bar{I}_C (representing the radial distribution of graph vertices). The numerical values of the above six topological information indices for the 427 acyclic, monocyclic, and bicyclic graphs, having 4–8 vertices, are given in Table II.

These indices may be combined in several ways.¹⁷ For example, the addition of the indices

$$\text{SI} = \sum_i I_i \quad (19)$$

would lose a great deal of information that these indices contain. Therefore, an attempt to design a practical superindex for discrimination of isomers must be carried out in such a way that all information is retained. One way of doing that is to represent SI as set of topological information indices,

$$\text{SI} = \{\bar{I}_D^W, \bar{I}_{\text{edge}}^E, \bar{I}_Z, \bar{I}_C, \bar{I}_{\text{orb}}, \bar{I}_{\text{chr}}\} \quad (20)$$

The discriminatory power of SI does not depend on the order of the elements in the set. However, a certain order once accepted must be kept for a meaningful comparison of a given set of isomers. The inspection of Table II clearly shows that the topological information index, as defined in eq. (20), discriminates completely all 427 graphs studied. Furthermore, several subsets of SI may be constructed with the power to discriminate the isomers studied. Therefore, one should regard the superindex (20) as the first approximation rather than a final solution of the problem of isomer discrimination.

The SI index as discussed above may be directly used in the comparative analysis of chemical structures. Obviously, if two chemical structures differ in any one index then they are distinct. One should, however, remember that even if the structures agree in all the indices they are not thereby known to be the same, i.e., the coincidence in the superindices of two graphs is still not sufficient for their isomorphism.

Estimating the efficiency of the superindex specified by eq. (20), we anticipate that though a counterexample could not be found in the limited sample of 427 graphs, such a case may occur at

Table II. The elements of topological information superindex, $SI = \{\bar{I}_D^W, \bar{I}_{EDGE}^E, \bar{I}_Z, \bar{I}_C, \bar{I}_{CHR}, \bar{I}_{ORB}\}$ calculated for 427 acyclic, monocyclic, and bicyclic graphs having 4–8 vertices.

Structure ^a	\bar{I}_D^W	\bar{I}_{EDGE}^E	\bar{I}_Z	\bar{I}_C	\bar{I}_{CHR}	\bar{I}_{ORB}	Structure ^a	\bar{I}_D^W	\bar{I}_{EDGE}^E	\bar{I}_Z	\bar{I}_C	\bar{I}_{CHR}	\bar{I}_{ORB}
1	2.5033	0.0000	0.8113	0.8113	0.8113	0.8113	63	3.2359	1.2516	1.3143	1.5219	1.5219	1.5219
2	3.2500	0.0000	0.7219	0.7219	0.7219	0.7219	64	3.1899	1.4592	1.3250	0.9709	1.5219	1.9219
3	3.8439	0.0000	0.6500	0.6500	0.6500	0.6500	65	4.6729	1.0000	1.7934	1.0000	1.0000	1.0000
4	4.3366	0.0000	0.5917	0.5917	0.5917	0.5917	66	3.8035	1.4488	1.2533	1.2516	0.9183	1.7925
5	4.7576	0.0000	0.5436	0.5436	0.5436	0.5436	67	3.2359	1.9183	1.3222	0.7219	1.3709	1.9219
6	2.4464	0.9183	1.3709	1.0000	1.0000	1.0000	68	4.6566	1.0000	1.1626	1.5000	0.8113	1.5000
7	3.7883	0.7219	1.3610	0.9183	1.0000	0.9183	69	4.2464	1.0000	1.6504	1.3788	0.9852	1.3788
8	4.7064	0.5917	1.2533	0.8113	1.0000	0.8113	70	3.2028	1.5219	1.3516	1.5219	1.5219	1.5219
9	3.1464	1.0000	1.4056	1.5219	0.9709	1.5219	71	3.7360	0.8631	1.4932	1.5850	1.4592	1.5850
10	4.2304	0.9183	1.2729	1.3788	0.8631	1.3788	72	3.1751	1.3709	1.3610	1.3709	1.5219	1.9219
11	3.1972	1.5000	1.3788	1.3709	0.9709	1.9219	73	3.7821	1.9183	1.5628	0.9183	1.0000	1.5850
12	3.7043	0.9709	1.6143	1.5850	1.0000	1.5850	74	3.7962	1.5850	1.3143	1.2516	0.9183	1.9183
13	3.7979	1.3709	1.3516	1.2516	0.9183	1.7925	75	4.2849	1.3788	1.2718	1.3788	1.1488	1.3788
14	4.2964	1.2516	1.3222	1.1488	0.8631	1.6645	76	3.8035	1.3788	1.2867	1.5850	1.2516	1.5850
15	4.7220	1.1488	1.2957	1.0613	1.8113	1.5480	77	4.1970	0.8113	1.5938	1.3788	1.5524	1.3788
16	4.2406	1.0000	1.7200	1.4488	0.9852	1.4488	78	3.8279	0.5917	1.5492	1.5850	1.4592	1.9183
17	4.2845	1.4592	1.3143	1.3788	0.9852	1.8424	79	4.2946	0.8113	1.6320	1.3788	1.4488	1.3788
18	4.6679	0.8631	1.6403	1.4056	0.9544	1.7500	80	3.7835	0.9852	1.5784	1.2516	1.4592	1.9183
19	3.7600	1.5219	1.6500	1.4592	1.0000	1.9183	81	4.6533	0.9183	1.6214	1.5613	0.9544	1.5613
20	4.7094	1.3788	1.2718	1.2988	0.9544	1.7500	82	4.6657	0.5033	1.6022	1.2988	0.9544	2.2500
21	4.6120	1.3788	1.6136	1.5000	1.0000	1.5000	83	3.7710	1.9183	1.2958	0.9183	1.4592	1.9183
22	3.7417	1.9219	1.3486	1.4592	0.9183	2.2516	84	4.6995	1.7500	1.5955	0.8113	1.0000	1.5000
23	4.1690	0.9183	1.6909	1.9502	0.9852	1.9502	85	3.7835	1.3788	1.5784	1.4592	1.0000	2.2516
24	4.2405	1.7925	1.2958	1.3788	0.8631	2.1281	86	3.7454	1.4592	1.6457	1.4592	1.0000	2.2516
25	4.2656	1.5850	1.6774	1.3788	0.9852	1.9502	87	3.7835	1.8424	1.5784	1.0000	1.4592	1.9183
26	4.6955	1.5567	1.6751	1.2988	0.9544	1.9056	88	3.7835	1.3788	1.4652	0.9183	1.4592	2.5850
27	4.6675	1.6645	1.2533	1.5488	0.8113	2.0000	89	4.6941	1.8366	1.7490	0.8113	1.0000	1.5000
28	4.6539	1.8424	1.2108	1.2988	0.8113	2.1556	90	4.2737	0.8113	1.6320	1.9502	0.9852	1.9502
29	4.6751	1.4488	1.8673	1.4056	1.0000	1.8113	91	3.7962	1.4592	1.5306	1.4592	1.4592	1.9183
30	4.8016	0.8631	1.8088	2.0000	1.0000	2.0000	92	3.7835	1.9502	1.2718	1.2516	1.4592	1.9183
31	4.2512	1.7925	1.6457	1.4488	0.9852	2.5216	93	4.2737	1.5000	1.5955	1.3788	0.9852	1.9502
32	4.6625	1.9502	1.6856	1.4056	0.9544	2.2500	94	4.2737	1.0000	1.6359	1.3788	1.4488	1.9502
33	4.6515	1.9502	1.8439	1.5000	1.0000	2.0000	95	4.5836	0.9183	1.6994	1.5000	1.5613	1.5000
34	4.1958	1.9183	1.6122	1.8424	0.9852	2.5216	96	4.2252	1.2988	1.6504	1.4488	0.9852	2.2359
35	4.6201	1.8424	1.5920	1.7500	0.9544	2.4056	97	4.2202	1.5000	1.6364	1.5566	0.9852	1.9502
36	4.6322	1.4488	1.6479	1.7500	0.9544	2.7500	98	4.6450	1.3921	1.5129	1.5000	1.2988	1.5000
37	4.6751	2.1281	1.6203	1.4056	0.9544	2.4056	99	4.6228	1.3921	1.8406	1.5000	1.0000	1.5000
38	4.6352	1.4488	1.8286	1.8113	1.0000	2.5000	100	3.7198	1.4592	1.5306	1.4592	1.4592	2.2516
39	4.6120	1.9502	1.6751	1.9056	0.9544	2.2500	101	4.6494	1.5000	1.7669	1.5000	1.0000	1.5000
40	4.2102	1.9183	1.6798	1.8424	0.9852	2.8074	102	3.7265	1.4488	1.5784	1.4592	1.4592	2.2516
41	4.6833	2.2359	1.6668	1.4056	1.0000	2.5000	103	4.2714	1.8424	1.2326	1.3788	1.3788	1.8424
42	4.5870	1.8424	1.6647	1.9056	0.9544	2.7500	104	4.2595	1.3788	1.5920	1.3788	1.4488	1.9502
43	4.6498	1.9502	1.6726	1.7500	0.9544	2.7500	105	3.8035	1.3788	1.5492	1.9183	1.4592	1.9183
44	4.6364	2.1281	1.6709	1.8113	1.0000	2.7500	106	4.6860	1.8366	1.7345	0.8113	1.5000	1.5000
45	4.6049	1.9502	1.8266	1.9056	1.0000	3.0000	107	4.2432	1.9502	1.2326	1.5566	0.8631	2.2359
46	2.5000	0.0000	1.3788	0.0000	1.0000	0.0000	108	4.6898	1.2244	1.5711	1.2988	1.2988	2.0000
47	3.2402	0.0000	1.3486	0.0000	1.5219	0.0000	109	4.2849	1.5566	1.5492	1.3788	1.3788	1.9502
48	3.7821	0.0000	1.6122	0.0000	1.0000	0.0000	110	4.2737	1.4056	1.6431	1.3788	1.4488	1.9502
49	4.2665	0.0000	1.6647	0.0000	1.4488	0.0000	111	4.2548	1.5566	1.6203	1.4488	0.9852	2.2359
50	4.6336	0.0000	1.8006	0.0000	1.0000	0.0000	112	4.2512	1.3788	1.6968	1.3788	1.4488	1.9502
51	3.2359	0.0000	1.3143	1.5219	0.9709	0.9709	113	4.2138	1.3788	1.5198	1.3788	1.3788	2.2359
52	2.5216	0.7219	1.2988	1.0000	1.5000	1.0000	114	3.7962	2.2516	1.3250	1.2516	1.2516	1.2516
53	3.2359	0.9183	1.3250	0.7219	1.5219	0.7219	115	3.7821	1.7925	1.5917	1.2516	1.4592	2.2516
54	2.5000	1.5000	1.2516	0.8113	1.5000	1.5000	116	4.2655	1.5000	1.6093	1.3788	1.4488	1.9502
55	3.7835	1.3788	1.6203	0.9183	1.0000	0.9183	117	4.6533	1.5305	1.6118	1.4056	0.9544	2.0000
56	3.2402	1.5219	1.2988	0.7219	1.3709	1.5219	118	4.2464	1.4056	1.6364	1.3788	1.3788	2.1281
57	3.7821	1.0000	1.5917	1.0000	1.4592	1.0000	119	3.7672	1.8424	1.6068	1.2516	1.4592	2.2516
58	3.7542	1.3788	1.4367	0.9183	1.5950	0.9183	120	4.6808	1.3921	1.7537	1.5000	1.4056	1.5000
59	4.6940	0.9183	1.8068	0.8113	1.0000	0.8113	121	4.6553	1.0000	1.6359	1.4056	1.2988	2.2500
60	4.2608	0.3113	1.1819	1.3788	0.8631	1.5566	122	4.5970	0.8113	1.8434	2.0000	1.0000	2.0000
61	3.7534	0.9183	1.2958	1.5850	0.9183	1.5850	123	4.6961	1.8500	1.5444	0.8113	1.2988	2.1556
62	3.2028	1.5219	1.3610	1.3709	0.9709	1.9219	124	4.6634	1.8366	1.7935	1.0000	1.0000	2.0000

Table II (continued from previous page)

Structure ^a	\bar{I}_D^W	\bar{I}_{EDGE}^E	\bar{I}_Z	\bar{I}_C	\bar{I}_{CHR}	\bar{I}_{ORB}	Structure ^a	\bar{I}_D^W	\bar{I}_{EDGE}^E	\bar{I}_Z	\bar{I}_C	\bar{I}_{CHR}	\bar{I}_{ORB}
125	4.6434	1.4355	1.6118	1.4056	0.9544	2.2500	188	4.6941	1.5305	1.7946	1.5000	1.4056	2.0000
126	4.2834	1.7500	1.5694	1.3788	1.3788	1.9502	189	4.2655	2.2500	1.6319	1.3788	1.4488	1.9502
127	4.6388	0.9864	1.6060	1.7500	0.9544	2.4056	190	4.6023	0.9183	1.7618	1.7500	1.4056	2.5000
128	4.2055	1.1488	1.6403	1.8424	0.9852	2.5216	191	4.6522	1.8911	1.5455	1.4056	1.2988	2.1556
129	4.2548	1.3788	1.6479	1.8424	0.9852	2.2359	192	4.6159	1.5000	1.6344	1.7500	0.9544	2.5000
130	4.2222	1.5000	1.5715	1.4488	1.3788	2.2359	193	4.2247	1.8424	1.6203	1.3788	1.3788	2.5216
131	4.6634	1.3921	1.8020	1.5000	1.0000	2.0000	194	4.6773	1.5305	1.6087	1.5000	1.4056	2.2500
132	3.7835	2.1281	1.4932	1.2516	1.4592	2.2516	195	4.6240	1.4050	1.6320	1.8113	1.0000	2.5000
133	4.2222	1.4056	1.6320	1.4488	1.4488	2.2359	196	4.6849	2.0000	1.6364	1.4056	1.0000	2.2500
134	4.6379	1.5305	1.7261	1.5000	1.0000	2.0000	197	4.6550	0.9864	1.7762	1.4056	1.4056	2.7500
135	4.6638	1.5000	1.6344	1.4056	0.9544	2.2500	198	4.6744	1.4056	1.6320	1.7500	1.2988	2.2500
136	4.6991	1.8366	1.7762	0.9544	1.4056	1.7500	199	4.2464	1.7500	1.6359	1.4488	1.1281	2.8074
137	4.6450	1.2244	1.5455	1.5000	1.2988	2.2500	200	4.2464	1.5000	1.6093	1.4488	1.4488	2.8074
138	4.5553	1.2988	1.6503	1.4056	0.9544	2.5000	201	4.2222	1.7500	1.5866	1.4488	1.5524	2.5216
139	4.2530	1.0613	1.5866	1.9502	1.3788	2.2359	202	4.6220	1.7527	1.6133	1.7500	0.9544	2.4056
140	4.6657	0.9911	1.7727	1.4056	1.4056	2.2500	203	4.6294	0.9864	1.7789	1.8113	1.4056	2.5000
141	4.2398	1.4488	1.6709	1.4488	1.4488	2.2359	204	3.7835	2.5216	1.5197	1.2516	1.4592	2.5850
142	4.2373	1.9502	1.6403	1.4488	0.9852	2.2359	205	4.5979	1.5000	1.6258	1.5000	1.4056	2.5000
143	4.2737	1.7500	1.4972	1.3788	1.3788	2.2359	206	4.6773	1.4466	1.7789	1.5000	1.0000	2.7500
144	4.6995	1.7500	1.6194	1.4056	1.2988	1.7500	207	4.6434	1.8366	1.6118	1.4056	1.4056	2.2500
145	4.2834	1.0613	1.6152	1.3788	1.3788	2.8074	208	4.6991	1.6577	1.6018	1.4056	1.2988	2.5000
146	4.2776	1.9502	1.6068	1.3788	1.3788	1.9502	209	4.2737	2.1556	1.4717	1.3788	1.3788	2.5216
147	4.2776	1.9502	1.6403	1.3788	1.3788	1.9502	210	4.6423	1.4056	1.8228	1.8113	1.0000	2.5000
148	4.6374	1.2244	1.8235	1.5000	1.4056	2.0000	211	4.6991	1.8366	1.6173	1.4056	1.4056	2.2500
149	3.7454	1.7925	1.5628	1.4592	1.4592	2.5850	212	4.5830	1.4355	1.7911	2.0000	1.4056	2.0000
150	4.2060	1.3788	1.6726	1.9502	1.4488	1.9502	213	4.6285	1.8911	1.6018	1.2988	1.2988	2.5000
151	4.2655	1.4056	1.6320	1.9502	1.4488	1.9502	214	4.6657	1.4466	1.7124	1.7500	1.4056	2.2500
152	4.2548	1.8424	1.6668	1.3788	0.9852	2.5216	215	4.6498	1.7500	1.7700	1.4056	1.4056	2.2500
153	4.1968	1.4488	1.6709	1.8424	0.9852	2.5216	216	4.2737	2.1556	1.5227	1.3788	1.3788	2.5216
154	4.6632	1.5000	1.8593	1.4056	1.0000	2.2500	217	4.2946	1.8113	1.5715	1.3788	1.3788	2.8074
155	4.2373	1.5566	1.5492	1.4488	1.3788	2.5216	218	4.2587	2.2500	1.5715	1.3788	1.5524	2.2359
156	4.2464	1.5488	1.6194	1.3788	1.3788	2.5216	219	4.6553	1.9056	1.6397	1.4056	1.4056	2.2500
157	4.6638	1.9056	1.6152	1.4056	0.9544	2.1556	220	4.2776	2.5216	1.5784	1.3788	0.9852	2.5216
158	4.6412	1.8113	1.4947	1.2988	1.2988	2.1556	221	4.2737	1.7500	1.6152	1.3788	1.4488	2.8074
159	4.1775	1.4488	1.2632	1.8424	1.4488	2.5216	222	4.2373	2.1281	1.6726	1.4488	0.9852	2.8074
160	4.6860	2.1972	1.5650	0.8113	1.2988	2.1556	223	4.2090	2.1556	1.6364	1.3788	1.3788	2.5216
161	4.6808	1.8366	1.5896	1.4056	0.9544	2.2500	224	4.6294	1.8366	1.6118	1.7500	0.9544	2.5000
162	4.2291	1.7500	1.5715	1.5566	1.3788	2.2359	225	4.6718	1.2988	1.7538	1.9056	1.4056	2.2500
163	4.6634	0.9183	1.7370	1.7500	1.4056	2.2500	226	4.2834	1.7500	1.6359	1.9502	1.4488	2.2359
164	4.6369	1.7500	1.8408	1.5000	1.0000	2.0000	227	4.6553	1.9056	1.6431	1.4056	0.9544	2.7500
165	4.6634	1.8911	1.6173	1.4056	1.0000	2.1556	228	4.2464	2.1556	1.6364	1.3788	1.3788	2.5216
166	4.2737	2.2500	1.6364	1.3788	0.9852	2.2359	229	4.6657	1.8366	1.6172	1.4056	1.2988	2.5000
167	4.6744	1.5000	1.6320	1.7500	0.9544	2.2500	230	4.6941	2.1972	1.5896	0.8113	1.2988	2.7500
168	4.2464	1.9056	1.6194	1.3788	1.3788	2.2359	231	4.6434	1.5305	1.6087	1.4056	1.4056	2.7500
169	4.6860	0.9181	1.7581	2.0000	1.4056	2.0000	232	4.6051	1.6577	1.7727	1.8113	1.0000	2.5000
170	4.6940	0.9183	1.7581	1.7500	1.4056	2.2500	233	4.6432	1.8366	1.7248	1.4056	1.5000	2.5000
171	4.4817	2.0000	1.5955	1.3788	1.3788	1.9502	234	4.2714	2.2359	1.5784	1.3788	1.3788	2.5216
172	4.6434	1.3921	1.6942	1.4056	1.4056	2.2500	235	4.6108	1.4466	1.6022	1.8113	1.4056	2.5000
173	4.6433	1.8366	1.6022	1.4056	1.5613	1.7500	236	4.6369	1.9056	1.7863	1.4056	1.4056	2.2500
174	4.2291	1.9056	1.4972	1.5566	1.3788	2.2359	237	4.5934	1.5488	1.6397	1.9056	0.9544	2.7500
175	4.2373	1.8424	1.6668	1.4488	1.3788	2.2359	238	4.6023	1.5305	1.6036	1.7500	1.4056	2.5000
176	4.6941	2.1972	1.6084	0.8113	1.0000	2.5000	239	4.6550	1.6577	1.7727	1.4056	1.4056	2.5000
177	4.2252	1.8113	1.6504	1.4488	1.4488	2.2359	240	4.6294	1.4466	1.8218	1.5000	1.0000	3.0000
178	4.2737	2.2500	1.5955	1.3788	1.3788	1.9502	241	4.2398	2.2359	1.5784	1.4488	1.3788	2.5216
179	3.8035	2.2359	1.4932	1.2516	1.4592	2.5850	242	4.2655	2.1556	1.6364	1.3788	1.4488	2.5216
180	4.6634	0.9911	1.7762	1.7500	1.4056	2.2500	243	4.2356	2.1281	1.6403	1.4488	1.4488	2.5216
181	4.2587	2.0000	1.5955	1.3788	1.3788	2.2359	244	4.2530	1.9056	1.6359	1.9502	1.4488	2.2359
182	4.6729	1.9056	1.7697	1.5000	1.0000	2.0000	245	4.6729	2.2500	1.6504	1.4056	0.9544	2.5000
183	4.6108	0.9183	1.6108	1.8113	1.4056	2.5000	246	4.6634	2.0588	1.8084	1.4056	1.0000	2.5000
184	4.6897	1.8366	1.6172	0.8113	1.4056	2.5000	247	4.6159	1.4056	1.6159	1.7500	1.2988	2.7500
185	4.2252	1.2988	1.6359	1.4488	1.4488	2.8074	248	4.2252	2.1556	1.6364	1.4488	1.4488	2.5216
186	4.6228	1.5305	1.7217	1.5000	1.5000	2.0000	249	4.1795	1.8113	1.6359	1.8424	1.4488	2.5216
187	4.6941	1.5305	1.6042	1.4056	1.4056	2.2500	250	4.6638	2.1556	1.5194	1.4056	1.2988	2.4056

Table II (continued from previous page)

Structure ^a	I ^W _D	I ^E _{EDGE}	I _Z	I _C	I _{CHR}	I _{ORB}	Structure ^a	I ^W _D	I ^E _{EDGE}	I _Z	I _C	I _{CHR}	I _{ORB}
251	4.6991	1.8366	1.6172	1.7500	1.2988	2.2500	314	4.6638	2.5000	1.6504	1.4056	0.9544	2.7500
252	4.5863	1.4056	1.8228	1.9056	1.0000	2.7500	315	4.6657	1.6577	1.7911	1.4056	1.4056	3.0000
253	4.1890	1.5488	1.6364	2.1281	1.4488	2.5216	316	4.6220	1.8800	1.6253	1.7500	1.2988	2.7500
254	4.6412	2.1556	1.5866	1.2988	1.2988	2.5000	317	4.6294	1.8911	1.6108	1.7500	1.2988	2.7500
255	4.6184	1.4056	1.7515	1.8113	1.4056	2.5000	318	4.6553	2.5000	1.5715	1.4056	1.2988	2.5000
256	4.6225	1.9056	1.6152	1.9056	0.9544	2.5000	319	1.9044	4.6632	2.1556	1.8027	1.4056	2.5000
257	4.6556	1.9056	1.6320	1.4056	1.4056	2.5000	320	4.6522	1.7527	1.7258	1.4056	1.4056	3.0000
258	4.6685	1.8366	1.5734	1.9056	1.2988	2.2500	321	4.5710	1.8910	1.7727	1.8113	1.4056	2.5000
259	4.6285	1.8911	1.7165	1.2988	1.5000	2.5000	322	4.6522	2.1972	1.7946	1.4056	1.4056	2.5000
260	4.6634	2.0588	1.6084	1.4056	1.2988	2.5000	323	4.6616	2.1133	1.6253	1.4056	1.4056	2.7500
261	4.2737	2.5000	1.5955	1.3788	0.9852	2.8074	324	4.2411	2.5000	1.6575	1.3788	1.3788	2.8074
262	4.6685	1.8911	1.5331	1.4056	1.2988	2.7500	325	4.6095	1.7527	1.7946	1.9056	1.4056	2.5000
263	4.6356	1.8113	1.6503	1.7500	0.9544	2.7500	326	4.6991	1.7255	1.7370	1.4056	1.4056	3.0000
264	4.6744	1.8113	1.6319	1.4056	1.2988	2.7500	327	4.6634	1.4466	1.7132	1.7500	1.4056	3.0000
265	4.6657	2.2810	1.6172	1.5613	0.9544	2.5000	328	4.6434	2.4194	1.6108	1.4056	1.4056	2.5000
266	4.6744	1.8113	1.7850	1.4056	1.4056	2.5000	329	4.6051	1.8911	1.7727	1.8113	1.4056	2.5000
267	4.6849	1.9056	1.7806	1.5613	1.4056	2.2500	330	4.2252	2.5000	1.6319	1.4488	1.3788	2.8074
268	4.2737	2.1556	1.5955	1.3788	1.3788	2.8074	331	4.6533	2.5033	1.6326	1.4056	1.2988	2.5000
269	4.6634	2.1972	1.8374	1.4056	1.0000	2.5000	332	4.1968	2.1281	1.6403	1.8424	1.3788	2.8074
270	4.7048	1.8911	1.7248	1.4056	1.4056	2.5000	333	4.2530	2.5000	1.5444	1.4488	1.4488	2.8074
271	4.6808	1.8910	1.6087	1.7500	0.9544	2.7500	334	4.6657	2.2810	1.6060	1.4056	1.2988	2.7500
272	4.6743	2.2810	1.6214	1.4056	1.4056	2.2500	335	4.6657	1.8366	1.7105	1.4056	1.4056	3.0000
273	4.6808	1.8366	1.7258	2.0000	1.4056	2.0000	336	4.6657	1.8366	1.7261	1.4056	1.4056	3.0000
274	4.6850	1.4466	1.7124	1.4056	1.4056	3.0000	337	4.6634	2.4194	1.8084	1.4056	1.0000	2.7500
275	4.2737	2.2500	1.5944	1.9502	1.3788	2.2359	338	4.6388	1.5305	1.7248	2.2500	1.4056	2.5000
276	4.6773	1.8366	1.7712	2.0000	1.4056	2.0000	339	4.6603	2.1972	1.7258	1.4056	1.5613	2.5000
277	4.2464	2.1556	1.6556	1.3788	1.4488	2.8074	340	4.6434	2.4194	1.5895	1.4056	1.5000	2.5000
278	4.2411	2.5000	1.6319	1.3788	1.4488	2.5216	341	4.6773	2.1133	1.7258	1.4056	1.4056	2.7500
279	4.6634	2.1133	1.7911	1.4056	1.0000	2.7500	342	4.6294	1.7527	1.7911	2.1556	1.0000	2.7500
280	4.2548	2.5216	1.3845	1.3788	1.3788	2.8074	343	4.6294	1.7527	1.7935	1.5000	1.4056	3.0000
281	4.6719	1.3921	1.7581	2.2500	1.4056	2.2500	344	4.6773	1.8366	1.7762	1.4056	1.4056	3.0000
282	4.6773	1.5305	1.7248	1.4056	1.4056	3.0000	345	4.6064	1.8113	1.7806	1.7500	1.4056	2.7500
283	4.6850	2.1133	1.6018	1.2988	1.2988	2.7500	346	4.6125	1.6977	1.7124	1.8113	1.5613	2.7500
284	4.6808	1.5305	1.7261	1.4056	1.4056	3.0000	347	4.6657	1.8800	1.7521	1.4056	1.4056	3.0000
285	4.6432	1.8366	1.7124	1.4056	1.4056	2.7500	348	4.6718	2.5000	1.6359	1.4056	1.4056	2.5000
286	4.6023	1.4466	1.7126	1.7500	1.5000	2.7500	349	4.6634	2.2810	1.6214	1.4056	1.4056	2.7500
287	4.6685	1.7527	1.7935	1.4056	1.4056	2.7500	350	4.6685	2.2810	1.6173	1.9056	1.4056	2.2500
288	4.2834	2.2500	1.6093	1.3788	1.4488	2.8074	351	4.6657	2.4194	1.5896	1.4056	1.2988	2.7500
289	4.6616	2.1972	1.6118	1.4056	1.4056	2.5000	352	4.6657	2.4194	1.5909	1.4056	1.2988	2.7500
290	4.6616	2.1972	1.6172	1.4056	1.4056	2.5000	353	4.6897	2.2810	1.6060	1.4056	1.4056	2.7500
291	4.5723	1.4056	1.7538	1.9056	1.4056	2.7500	354	4.6257	1.9056	1.6431	1.8113	1.4056	2.7500
292	4.6718	2.1556	1.6601	1.4056	1.4056	2.5000	355	4.6417	1.5488	1.7697	1.8113	1.4056	3.0000
293	4.6514	2.2810	1.8064	1.4056	1.4056	2.2500	356	4.6685	2.0588	1.6018	1.5613	1.2988	3.0000
294	4.6533	2.4194	1.6214	1.4056	1.4056	2.2500	357	4.6051	1.6577	1.7105	1.8113	1.4056	3.0000
295	4.6657	2.4194	1.6108	1.4056	0.9544	2.7500	358	4.6522	2.1972	1.7911	1.4056	1.4056	2.7500
296	4.6634	2.2810	1.8084	1.4056	1.4056	2.2500	359	4.6434	2.5033	1.6060	1.4056	1.2988	2.7500
297	4.6657	2.1133	1.7258	1.4056	1.4056	2.5000	360	4.6514	2.1972	1.8089	1.4056	1.4056	2.7500
298	4.6991	2.0588	1.6108	1.4056	1.2988	2.7500	361	4.2737	2.7500	1.6319	1.3788	1.3788	2.8074
299	4.6240	1.7527	1.7946	1.5000	1.4056	2.7500	362	4.6369	1.9056	1.7806	1.5000	1.4056	3.0000
300	4.6369	1.5000	1.7863	1.7500	1.4056	2.7500	363	4.6095	2.1972	1.6172	1.9056	1.4056	2.5000
301	4.6423	1.9056	1.7850	1.5000	1.0000	3.0000	364	4.6634	1.8911	1.7762	2.2500	1.4056	2.2500
302	4.6773	2.5033	1.7490	1.4056	1.0000	2.5000	365	4.6773	2.4194	1.7391	1.4056	1.0000	3.0000
303	4.6634	2.1133	1.7490	1.4056	1.4056	2.5000	366	4.5792	1.8800	1.7258	1.9056	1.4056	2.7500
304	4.6186	1.5488	1.7669	1.7500	1.4056	2.7500	367	4.6533	2.0588	1.7391	1.4056	1.4056	3.0000
305	4.6657	2.1133	1.6082	1.4056	1.2988	2.7500	368	4.6638	2.5000	1.6556	1.4056	1.2988	2.7500
306	4.6764	2.2500	1.5866	1.2988	1.2988	2.7500	369	4.6941	2.4194	1.6018	1.4056	1.4056	2.7500
307	4.6369	2.1556	1.8434	1.5000	1.0000	2.7500	370	4.6206	2.1556	1.7850	1.8113	1.4056	2.5000
308	4.6417	2.1556	1.6359	1.7500	0.9544	2.7500	371	4.6995	2.5000	1.6359	1.4056	1.2988	2.7500
309	4.6744	2.5000	1.6093	1.4056	0.9544	2.7500	372	4.6773	1.8366	1.7762	2.5000	1.0000	2.5000
310	4.6228	1.8366	1.7789	1.5000	1.4056	2.7500	373	4.6729	2.5000	1.8140	1.4056	1.4056	2.5000
311	4.6533	1.6577	1.7762	1.4056	1.4056	3.0000	374	4.5718	1.8911	1.7762	1.9056	1.4056	2.7500
312	4.6522	2.1972	1.7532	1.4056	1.4056	2.5000	375	4.6616	2.4194	1.6108	1.4056	1.5000	2.7500
313	4.6729	2.1556	1.6364	1.4056	1.2988	2.7500	376	4.6206	2.1556	1.7700	1.8113	1.0000	3.0000

Table II (continued from previous page)

Structure ^a	\bar{T}_D^W	\bar{T}_{EDGE}^E	\bar{T}_Z	\bar{T}_C	\bar{T}_{CHR}	\bar{T}_{ORB}	Structure ^a	\bar{T}_D^W	\bar{T}_{EDGE}^E	\bar{T}_Z	\bar{T}_C	\bar{T}_{CHR}	\bar{T}_{ORB}
377	4.6634	2.5033	1.6326	1.4056	1.4056	2.7500	403	4.6634	2.4194	1.7650	1.4056	1.4056	3.0000
378	4.6941	2.5033	1.7605	1.7500	1.4056	2.2500	404	4.6522	2.7255	1.7394	1.4056	1.4056	2.7500
379	4.6657	2.7255	1.6173	1.4056	0.9544	3.0000	405	4.6294	2.4194	1.7370	1.5000	1.4056	3.0000
380	4.6634	2.4194	1.7370	1.4056	1.4056	2.7500	406	4.6220	2.1972	1.7242	1.7500	1.4056	3.0000
381	4.6230	2.1556	1.6397	1.8113	1.4056	2.7500	407	4.6294	2.4194	1.7490	1.7500	1.4056	2.7500
382	4.6634	2.4194	1.7490	1.4056	1.4056	2.7500	408	4.6294	2.1972	1.7391	1.7500	1.4056	3.0000
383	4.6638	2.1556	1.7700	1.4056	1.4056	3.0000	409	4.6773	2.5033	1.7432	1.4056	1.4056	3.0000
384	4.6634	2.1972	1.7394	1.4056	1.4056	3.0000	410	4.6634	2.5033	1.7605	1.4056	1.4056	3.0000
385	4.6808	2.4194	1.6084	1.4056	1.2988	3.0000	411	4.6773	2.1972	1.7261	1.7500	1.4056	3.0000
386	4.5969	1.7500	1.7700	1.9056	1.4056	3.0000	412	4.6522	2.7255	1.7260	1.4056	1.5000	2.7500
387	4.6941	2.1972	1.7258	1.4056	1.4056	3.0000	413	4.6294	1.7527	1.7258	2.2500	1.4056	3.0000
388	4.6773	2.1972	1.7490	1.4056	1.4056	3.0000	414	4.6632	2.7500	1.7934	1.4056	1.4056	2.7500
389	4.6743	1.8366	1.7124	1.8113	1.4056	3.0000	415	4.5863	2.1556	1.7697	1.9056	1.4056	3.0000
390	4.6897	2.5033	1.6224	1.7500	1.4056	2.5000	416	4.6423	2.2500	1.7999	1.7500	1.4056	3.0000
391	4.5925	2.1972	1.7935	1.7500	1.4056	2.7500	417	4.6298	2.2810	1.8064	1.7500	1.4056	3.0000
392	4.6657	2.5033	1.6173	1.4056	1.2988	3.0000	418	4.6228	2.7255	1.6214	1.7500	1.4056	2.7500
393	4.6634	2.2810	1.7391	1.4056	1.4056	3.0000	419	4.6533	2.7255	1.7394	1.4056	1.4056	3.0000
394	4.6051	2.1972	1.7258	1.8113	1.4056	2.7500	420	4.6808	2.7255	1.7327	1.4056	1.4056	3.0000
395	4.6634	2.7255	1.8084	1.4056	1.4056	2.5000	421	4.6379	2.4194	1.7391	1.7500	1.4056	3.0000
396	4.6294	2.1972	1.7911	1.7500	1.4056	2.7500	422	4.6294	2.4194	1.7521	1.7500	1.4056	3.0000
397	4.6941	2.4194	1.6224	1.7500	1.2988	2.7500	423	4.6808	2.7255	1.7391	1.4056	1.4056	3.0000
398	4.6897	2.4194	1.7370	1.4056	1.2988	3.0000	424	4.6634	2.7255	1.8064	1.4056	1.4056	3.0000
399	4.6184	2.5000	1.6359	1.8113	1.2988	2.7500	425	4.6356	2.7500	1.6504	1.7500	1.2988	3.0000
400	4.6729	2.7500	1.7999	1.4056	1.0000	3.0000	426	4.6228	2.7255	1.7394	1.7500	1.4056	3.0000
401	4.6634	2.4194	1.7370	1.4056	1.4056	3.0000	427	4.6051	2.7255	1.7391	1.8113	1.4056	3.0000
402	4.6657	1.8366	1.7248	2.2500	1.4056	2.7500							

^a Structures are given in Fig. 1.

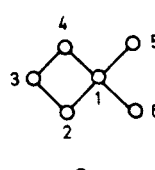
structures having a large number of atoms. Different refinements can be used to increase the discrimination power of the superindex. First of all one can reexamine the necessity of using real numbers in the construction of the superindex. The real numbers are very appropriate when the topological index they express is applied to structure–property correlations (due to this the data presented in Table II may be of interest for such purposes). They may, however, cause some difficulties when dealing with isomer discrimination. Thus, when two indices differ only in the last decimal place this could be due either to their distinctness or to a round-off error. Indeed, to avoid the ambiguity one can write the indices using more figures after the decimal point. The indices written in such a way would, however, lose some of their advantages—the concise expression and simplicity in use.

Proceeding from the above arguments we can redefine the superindex in a better approximation expressing each of the indices in terms of sets of integers instead of real numbers. These sets are the distributions already used in the determination of

the information indices and express the cardinality of the classes of equivalence of different graph elements: vertices, edges, distances, etc. In order to provide a practical superindex the sets used in its construction should be ordered according to the increasing complexity of their determination. This would help in finding a fast algorithm for isomer discrimination which would determine further sets only if the preceding ones failed to discriminate the isomers. For these reasons we have extended the basis for the superindex specification, including here some simple but important graph characteristics to start with: the number of ones in the adjacency and incidence matrix of the graph, N_1^A and N_1^I , respectively (the number of zeros in these matrices is not an independent variable), as well as the number of vertices with different degree (valencies), N_i^d . Simultaneously, we have excluded the orbit distribution of vertices since the determination of the orbits of the graph often requires a complicated calculation. As a compensation we have extended the classical radial (or centric) partition of the graph vertices to the so-called complete radial partition.¹⁹ Additional subsets are

formed in the latter, making subsequent use of additional criteria of vertex equivalence: the vertex distance rank (the sum of distances of the vertex to all other vertices in the graph), the vertex distance code (an increasing sequence of all distances of the vertex with their occurrence numbers written as superscripts), and finally, if a number of vertices still remain in the central set of vertices (graph kernel) a further discrimination between them may result from the application of the first three criteria to the kernel subgraph.

Example:

	vertex	distance code	distance radius	distance rank
 <p>G</p>	1	1 ⁴ 2 ¹	2	6
	2	1 ² 2 ³	2	8
	3	1 ² 2 ³	2	8
	4	1 ² 2 ¹ 3 ²	3	10
	5	1 ¹ 2 ³ 3 ¹	3	10
	6	1 ¹ 2 ³ 3 ¹	3	10

The vertex radius divides the six vertices into two subsets, {1, 2, 3} and {4, 5, 6}. The distance rank additionally splits the first set into two subsets, {1} and {2, 3}, and the distance code does the same with the second set, {4} and {5, 6}. The larger member of classes of vertex equivalence which results increases the discriminating ability of the superindex as will be shown later.

After making all these alterations we arrive at a refined definition of the superindex (or rather "supercode" or "combined code").

$$\begin{aligned}
 \text{SI} = \{ & N_1^A; N_1^I; N_1^d, N_2^d, \dots, N_{\max}^d; \\
 & N_1^E, N_2^E, \dots, N_{\max}^E; N_1^D, N_2^D, \dots, N_{\max}^D; \\
 & N_1^C, N_2^C, \dots, N_{\max}^C; \\
 & N_1^Z, N_2^Z, \dots, N_{\max}^Z; N_1^{\text{chr}}, N_2^{\text{chr}}, \dots, N_{\max}^{\text{chr}} \} \quad (21)
 \end{aligned}$$

Here, N_i^E , N_i^D , N_i^C , N_i^Z , and N_i^{chr} stand for the number of elements in the i th class of equivalence within the five distributions discussed above (the edge and distance distributions, centric or complete radial vertex partition, the Hosoya decomposition, and the chromatic vertex partition, respectively).

An important advantage of eq. (21) is that most of the sequences of integers in it are ordered. Thus we have N_1^d , N_2^d , N_3^d , and N_4^d for the vertices with

degrees (valencies) 1–4, respectively. Similarly, we denote N_1^E , N_2^E , \dots , N_9^E for carbon-carbon bonds 1–2, 1–3, 1–4, 2–2, 2–3, 2–4, 3–3, 3–4, and 4–4, respectively, where 1, 2, 3, and 4 stand for primary, secondary, tertiary, and quaternary carbon atoms.¹¹ N_i^D are ordered according to the increased magnitude of the distances: 1, 2, 3, etc. The Hosoya partition $p(G, k)$ is also ordered according to the increasing values of $k = 0, 1, 2, \dots$, $[N/2]$. The remaining two distributions can be ordered by convention. Thus, the complete radial (centric) distribution, N_1^C , N_2^C , N_3^C , etc., is ordered by convention as the number of vertices in the center (polycenter¹⁹ or oligocenter²⁰), the first, second, etc., neighboring spheres, respectively. In the cases of chromatic distribution of graph vertices, the cardinalities of the classes of equivalence, N_i^{chr} , should be ordered to form a nondecreasing sequence.

The presentation of the superindex as a *completely* ordered sequence of integers is essential since any ambiguity in comparing the superindices of two isomers is thus avoided. Moreover, such a superindex has a greater discriminating ability as compared with superindices composed of a set of real numbers (topological indices) which are obtained mainly on the basis of unordered distributions of the graph elements (for instance, the same centric index I_c results for the radial distributions 1, 2, 3; 3, 2, 1; 2, 1, 3, etc.). The two kinds of superindex defined by eqs. (21) and (20) will, however, have the same discriminating power when all indices in eq. (20) are calculated after a preliminary ordering of the corresponding distribution, using the procedure suggested by Muirhead.³⁴ Still, in this case the superindex defined by eq. (21) should be preferred since the calculation is simpler.

Comparing two isomeric structures by means of the superindex it is seen that only one case does not differentiate them among the 2^n combinations derived from eq. (21), where n is the sum of the number of equivalence classes N_{\max}^j :

$$\begin{aligned}
 n = 2 + N_{\max}^d + N_{\max}^E + N_{\max}^D + N_{\max}^C \\
 + N_{\max}^Z + N_{\max}^{\text{chr}} \quad (22)
 \end{aligned}$$

In estimating 2^n one may attempt to answer the question: "What is the smallest pair of graphs that

the superindex fails to distinguish?" The answer to this question is not straightforward. We have attempted to approach this problem by taking two extreme estimates of n , an upper and a lower one. The pair of graphs which the superindex fails to distinguish will appear when the following inequality holds:

$$N_I > 2^n - 1 \quad (23)$$

i.e., when the number of isomers N_I becomes greater than the number of distinct superindices ($2^n - 1$). The two extreme estimates of n could be given thus answering two questions: "When does the pair of graphs which the superindex fails to distinguish (a) *obligatorily* occur and (b) when does it have the least possible chance to appear *randomly*?" Case (a) determines the upper limit of n and occurs when the number of isomers is greater than the maximal number of the superindices for a given series of isomers. Case (b) refers to the lower estimate of n . It deals with the first two isomers that are not obviously distinct and therefore could have the same superindex. Examples are given below to elucidate the point.

Alkanes were selected as a suitable model since the number of their isomers is known up to 57 carbon atoms (unfortunately such data are still missing for polycyclic compounds). For this class of compounds the upper and lower estimates for N_{\max}^j in eq. (22) can, respectively, be obtained:

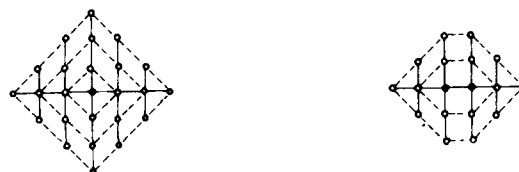
(a) $N_{\max}^d = 4; 3$ (carbon atoms of valencies 1–4, and 1–3, respectively);

(b) $N_{\max}^E = 9; 4$ (all kinds of bonds between primary, secondary, tertiary, and quaternary carbon atoms; bonds between primary and secondary, primary and tertiary, secondary and secondary, and secondary and tertiary carbon atoms).

In the above two cases the lower value for N_{\max} is not 2, since this situation occurs in the *single* linear alkane in each isomer series. The second isomer to be compared will have at least one tertiary carbon atom.

(c) $N_{\max}^D = N - 3; [-1 + (2N - a)^{1/2}]$, where $a = 0$ for N even, and $a = 1$ for N odd, and the square brackets mean the least integer exceeding the real number in them. Making the first estimate

we have excluded the single linear isomer ($N_{\max}^D = N - 1$), as well as the branched isomers having one branch of length one ($N_{\max}^D = N - 2$) since the latter are all distinct. The second estimate is obtained dealing with the close-packed odd and even isomers depicted below (the first, second, and third neighbor spheres around the center are also given in order to facilitate the analysis in the next paragraph):



The following series are obtained for N and N_{\max}^D :

N (even)	2	8	$18 \dots N$
N_{\max}^D	1	3	$5 \dots -1 + \sqrt{2N}$
N (odd)	5	13	$25 \dots N$
N_{\max}^D	2	4	$6 \dots -1 + (2N - 1)^{1/2}$

(d) $N_{\max}^C = (N - 2)/2$ for N even, and $(N - 1)/2$ for N odd; $[\sqrt{2N}/2]$ for N even, and $[(1 + (2N - 1)^{1/2})/2]$ for N odd. The first two estimates are made taking the linear part of alkane having $N - 2$ atoms, while the second two estimates are again obtained in an inductive way from the series:

N (even)	2	8	18	$\dots N$
N_{\max}^C	1	2	3	$\dots \sqrt{2N}/2$
N (odd)	1	5	13	$25 \dots N$
N_{\max}^C	1	2	3	$4 \dots (1 + (2N - 1)^{1/2})/2$

(e) $N_{\max}^Z = [N/2] + 1; 2 + N/5$. Here the square brackets mean the largest integer *not* exceeding the real number in them. The first estimate follows from the definition of the classes of equivalence in the Hosoya graph decomposition ($k = 0, 1, 2, \dots, [N/2]$). The second estimate is a rough extrapolation of the series¹⁴

N	5	6	7	8	9	10	15	20	25
N_{\max}^Z	2	3	3	3	3	4	to 5	6	7

(f) $N_{\max}^{\text{chr}} = 2; 2$. The estimate is one and the same since all the acyclic graphs are able to be bicolored.

By substituting the two estimates for N_{\max}^j into eq. (22), the corresponding estimates of n , and

Table III. The upper and lower estimates for the discrimination power of the topological superindex for alkane series.

N	10	15	20	25
n_{\max}	34	44	54	64
n_{\min}	21	24	26	28
$2^{n_{\max}}$	1.7×10^{10}	1.8×10^{13}	1.8×10^{16}	1.8×10^{19}
$2^{n_{\min}}$	2.1×10^6	1.7×10^7	6.7×10^7	2.7×10^8
N_I^a	75	4347	366,319	3.7×10^7

^a Reference 3.

hence of 2^n , are obtained [eqs. (24) and (25) and Table III]:

$$n_{\max} = 15 + [N/2] + 1.5 - a \quad (24)$$

where $a = 1$ for N even, and $a = 0.5$ for N odd. The square brackets mean the *largest integer* not exceeding the real number in them.

$$n_{\min} = 12 + N/5 + [b + 1.5(2N - c)^{1/2}] \quad (25)$$

where $b = 0, c = 0$ for N even, and $b = 0.5, c = 1$ for N odd. The square brackets here mean the *least integer* exceeding the real number in them.

It is seen from Table III that the upper estimate of the discrimination power of the superindex (21) far exceeds the number of isomers in the alkane series and, moreover, it increases much faster than the number of isomers. The lower estimate is also greater than the number of isomers even for alkanes having 25 atoms. The number of isomers increases faster and one should expect $N_I > 2^{n_{\min}}$ for $N > 30$. Therefore, it seems plausible to suppose that the superindex, as defined by eq. (21), resolves the problem for the isomer differentiation in alkane series. Still the superindex seems to be insufficient for very large polycyclic systems where the number of isomers and their possible index degeneracy increase very fast (see Table I).

One can conclude that, though far from solving the graph isomorphism problem, the superindex approach presented in this article seems powerful. It presumes a further development, extending the sets of integers that form the superindex as much as necessary for the solution of a very large number of isomer discrimination problems of practical importance. Work in this direction is in progress.³⁵

Professors A. T. Balaban (Bucuresti) and M. Randić (Ames) have kindly allowed the authors to read their manuscripts prior to publication. Their kindness is gratefully acknowledged here. The authors would also like to thank Professors B. M. Gimarc (Columbia, SC) and M. Randić (Ames) for helpful discussions about the isomer discrimination. They also thank the referees for their valuable suggestions leading to the substantial improvement of the manuscript. N. T. was supported in part by the Self-Managing Authority for Scientific Research of Croatia and by the National Science Foundation (Grant No. F6F006-Y).

References

1. N. L. Biggs, *Algebraic Graph Theory*, Cambridge U.P., Oxford, 1974.
2. I. Gutman and N. Trinajstić, *Top. Curr. Chem.*, **42**, 49 (1973).
3. A. T. Balaban, Ed., *Chemical Applications of Graph Theory*, Academic, London, 1976.
4. R. J. Wilson, in *Colloq. Math. Soc. János Bolai*, **18**, *Combinatorics, Keszthely (Hungary)*, 1147 (1976).
5. M. Randić and C. L. Wilkins, *J. Chem. Inf. Comput. Sci.*, **19**, 31 (1979).
6. M. Randić, in *Proc. Bremen Conf., Mathematical Structures in Chemistry*, June 1978, Bremen, Germany, *Math. Chem.*, **7**, 5 (1979).
7. D. H. Rouvray and A. T. Balaban, in *Applications of Graph Theory*, R. J. Wilson and L. W. Beineke, Eds., Academic, London, 1979, p. 177.
8. D. Bonchev, J. V. Knop, and N. Trinajstić, in *Proc. Bremen Conf., Mathematical Structures in Chemistry*, June 1978, Bremen, Germany, *Math. Chem.*, **7**, 21 (1979).
9. A. T. Balaban, in "Steric Fit in QSAR," *Lecture Notes in Chemistry*, Springer-Verlag, Berlin, 1980.
10. I. Gutman, B. Rušćić, N. Trinajstić, and C. F. Wilcox, Jr., *J. Chem. Phys.*, **62**, 3399 (1975).
11. M. Randić, *J. Am. Chem. Soc.*, **97**, 6609 (1975).
12. L. Lovasz and J. Pelikan, *Period. Math. Hung.*, **3**, 175 (1973).
13. H. Hosoya, *Bull. Chem. Soc. Jpn.*, **44**, 2332 (1971).
14. K. Mizutani, K. Kawasaki, and H. Hosoya, *Nat. Sci. Rep., Ochanomizu Univ.*, **22**, 39, 181 (1971).
15. H. Wiener, *J. Am. Chem. Soc.*, **69**, 17, 2636 (1947); *J. Phys. Chem.*, **52**, 425 (1948).
16. D. Bonchev and N. Trinajstić, *J. Chem. Phys.*, **67**, 4517 (1977); *Int. J. Quantum Chem. Quantum Chem. Symp.*, **12**, 293 (1978).
17. A. T. Balaban, *Theor. Chim. Acta*, **53**, 355 (1979).
18. A. T. Balaban and I. Motoc, *Math. Chem.*, **5**, 197 (1979).
19. D. Bonchev, A. T. Balaban, and Ov. Mekenyan, *J. Chem. Inf. Comput. Sci.*, **20**, 106 (1980).
20. D. Bonchev, A. T. Balaban, and M. Randić, *Int. J. Quantum Chem.*, in press.
21. J. R. Platt, *J. Chem. Phys.*, **15**, 419 (1947); *J. Phys. Chem.*, **56**, 328 (1952).
22. E. A. Smolenskii, *Zh. Fiz. Khim.*, **38**, 1288 (1964).
23. L. B. Kier and L. H. Hall, *Molecular Connectivity in*

- Chemistry and Drug Research*, Academic, New York, 1976.
24. M. Randić, *J. Chromatogr.*, **161**, 1 (1978); *Int. J. Quantum Chem. Quantum Biol. Symp.*, **5**, 245 (1978); see also M. Randić and C. L. Wilkins, *J. Phys. Chem.*, **23**, 1525 (1979).
25. D. Bonchev, Ov. Mekenyan, G. Protić, and N. Trinajstić, *J. Chromatogr.*, **176**, 159 (1979).
26. C. Shannon and W. Weaver, *Mathematical Theory of Communication*, University of Illinois, Urbana, 1949.
27. L. Brillouin, *Science and Information Theory*, Academic, New York, 1956.
28. D. H. Rouvray, in *Chemical Applications of Graph Theory*, A. T. Balaban, Ed., Academic, London, 1976, p. 175.
29. D. Bonchev, in *Proc. Bremen Conf., Mathematical Structures in Chemistry*, June 1978, Bremen, Germany, *Math. Chem.*, **7**, 65 (1979).
30. F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1971, 2nd ed., p. 35.
31. N. Rashevsky, *Bull. Math. Biophys.*, **17**, 229 (1955).
32. E. Trucco, *Bull. Math. Biophys.*, **18**, 129, 237 (1956).
33. A. Mowshowitz, *Bull. Math. Biophys.*, **30**, 175, 225, 337, 533 (1968).
34. R. F. Muirhead, *Proc. Edinburgh Math. Soc.*, **19**, 36 (1901); **21**, 144 (1903); **24**, 45 (1906).
35. D. Bonchev and I. Grncharov, unpublished results.