Discrimination of Isomeric Structures Using Information Theoretic Topological Indices*

C. Raychaudhury, S. K. Ray, and J. J. Ghosh†
Department of Biochemistry, 35, Ballygunge Circular Road, Calcutta-700 019, India
A. B. Roy
Department of Mathematics, Jadavpur University, Calcutta-700 032, India
S. C. Basak
Department of Chemistry, University of Minnesota, Duluth, Minnesota 55812
Received 21 July 1983; accepted 1 March 1984

Three newly defined information theoretic topological indices, namely "degree complexity (I^d)," "graph vertex complexity (H^V)," and "graph distance complexity (H^D)" along with three other information indices have been used to study their discriminating power of 45 trees and 19 monocyclic graphs. It is found that the newly defined indices have satisfactory discriminating power while H^D has been found to be the only index to discriminate all the graphs studied.

INTRODUCTION

One of the recent trends in the evaluation of different topological features of a molecule like branching pattern, bonding type, connectivity, etc., is to derive a numerical descriptor of the structural formal of the molecule known as a topological index.[1-3] Topological indices are generally derived from graphs[4] by which the structural formula of a molecule can be represented. The use of information theory[5] in graphs has been found to be very useful in deriving efficient information theoretic topological indices.[3,6-18] An information theoretic topological index is basically a measure of "information content" or "complexity" of a graph, which reflects the diversity embedded in the graph with respect to any topological property (of the graph) like degrees or neighborhoods of the vertices, distances in a graph, etc., in the form of a single numerical value. Recently in his book[19] Bonchev has discussed in details about different information theoretic topological indices and their various uses.

There are two different important fields where topological indices have found applications, viz. structure-property, structure-activity relationship studies[1-3,19-22] and chemical documentation.[13,23] For structure-property and structure-activity relationship studies uniqueness in the values of a topological index for different molecular graphs is not a necessary criterion, while for good documentation the molecular structure of a compound should have an unique name or code.[3] The present work is addressed to the application of information theoretic topological indices in the latter mentioned field of study.

The usefulness of a topological index in chemical documentation is revealed from its potentiality of discriminating isomeric structures and hence discriminating the corresponding graphs. But it has been found that no single topological index can discriminate graphs uniquely,[19] though the distance connectivity index (J) proposed by Balaban[23] shows considerable amount of discriminating power. However, in this connection Bonchev, Mekenyan, and Trinajstic[13] have proposed a concept of topological "superindex" that allows more than one topological indices as its components. The concept of superindex emphasizes that two graphs are said to be discriminated by the superindex if the values of at least one of its components are different for the two graphs.

In this article we have used three newly defined[24] information theoretic topological indices called "degree complexity (I^d)," "graph vertex

*Original proofs of this paper were lost in the mail.
† To whom correspondence should be addressed.

Journal of Computational Chemistry, Vol. 5, No. 6, 581–588 (1984)

CCC 0192-8651/84/060581-08$04.00

complexity (H^V)," and "graph distance complexity (H^D)" to find their potentiality of discriminating 45 trees having 4–8 vertices and 19 monocyclic graphs having 4–6 vertices. A comparative study has also been made among I^d, H^V, H^D, the first information theoretic topological index I_{top} due to Rashevsky[6] and the distance information indices $\bar{I}_D^E$ and $\bar{I}_D^W$ proposed by Bonchev and Trinajstic[9] to investigate the relative efficacy of the indices in discriminating the above mentioned graph. Finally, the discriminating power of J index of Balaban[23] and that of H^V, H^D, $\bar{I}_D^E$ and $\bar{I}_D^W$, I_{top} and I^d has been tested on few cyclic graphs.

DERIVATION OF GRAPH COMPLEXITIES

In this article we will derive topological information content of undirected, connected graphs.

Degree complexity: Let there be n_1 vertices of degree d_1, n_2 vertices of degree $d_2, \ldots, n_k$ vertices of degree d_k in a graph G where

$$\sum_{i=1}^{k} n_i d_i = d \tag{1}$$

$$\sum_{i=1}^{k} n_i = n \tag{2}$$

n being thus the number of vertices in G.

If the n vertices of G are partitioned in such a manner that the vertices with same degree lie in one class, then one gets $A_1, A_2, \ldots, A_k, k$ disjoint subsets of the vertex set of G and can attach the following finite probability scheme with the partition of n:

$$\begin{pmatrix} A_1 & A_2 & \cdots & A_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}, \qquad p_i > 0,$$

$$\sum_{i=1}^{k} p_i = 1 \tag{3}$$

where $p_i = n_i/n$. n_i is thus number of vertices in A_i. One can now calculate the information content of G with the help of scheme (3) using Shannon's formula.[5] Rashevsky[6] first introduced the idea of such a measure called topological information content of graph and denoted it by I_{top}. Thus I_{top} is given by

$$I_{\text{top}} = -\sum_{i=1}^{k} \frac{n_i}{n} \log_2 \frac{n_i}{n} \tag{4}$$

One can also now make a partition of d, the sum of the degrees of the vertices of G, where the degrees $d_1, d_1, \ldots, n_1$ times, $d_2, d_2, \ldots, n_2$ times $\ldots, d_k, d_k, \ldots, n_k$ times may be considered to be the summands of the partition. The following finite probability scheme thus can be attached with the partition of d:

$$\begin{pmatrix} d_1 & d_2 & \cdots & d_k \\ n_1 & n_2 & \cdots & n_k \\ q_1 & q_2 & \cdots & q_k \end{pmatrix}, \qquad q_i > 0,$$

$$\sum_i n_i q_i = 1 \tag{5}$$

where $q_i = d_i/d$. Thus a measure of information content of G can be obtained with the help of scheme (5) using Shannon's formula.[5] We call this measure the degree complexity of G and denote it by $I^d(G)$. $I^d(G)$ is given by

$$I^d(G) = -\sum_{i=1}^{k} n_i \frac{d_i}{d} \log_2 \frac{d_i}{d} \tag{6}$$

$I^d(G)$ is measured in bits.

Distance complexities: let in a graph G, there are a_1 vertices at a distance 1, a_2 vertices at a distance $2, \ldots, a_e$ vertices at a distance e from a vertex v of G, where e is the eccentricity[4] of v. Then one can associate an expression with v of the following form:

$$v\colon 1^{a_1}, 2^{a_2}, \ldots, e^{a_e} \tag{7}$$

Expression (7) is analogous to "distance code" defined by Bonchev, Mekenyan, and Trinajstic[13] Considering v to be situated at a distance zero from itself, expression (7) can be written in a revised form:

$$v\colon 0^1, 1^{a_1}, 2^{a_2}, \ldots, e^{a_e} \tag{8}$$

One can now construct an ordered sequence $1, a_1, a_2, \ldots, a_e$ with the help of the indices of the components of expression (8). The sequence

$$1, a_1, a_2, \ldots, a_e \tag{9}$$

will be called the "distance frequency sequence" of v and will be denoted by $D(v)$. If $v_1, v_2, \ldots, v_n$ be the n vertices of G, then one gets $D(v_1)$, $D(v_2), \ldots, D(v_n)$, n distance frequency sequences. Evidently, $D(v)$ generates a partition $(1, a_1, a_2, \ldots, a_e)$ of the vertex set $V(G)$ of G.

It is to be noted that $D(v)$ is analogous to "distance partition" defined by Balaban.[3]

Now defining

$$R(v) = \sum_{i=1}^{e} ia_i$$

$$= \text{distance sum of vertex } v \qquad (10)$$

it follows immediately that expression (7) generates a partition ($1, 1, \ldots, a_1$ times, $2, 2, \ldots, a_2$ times, $\ldots, e, e, \ldots a_e$ times) of $R(v)$ (in this case the distance zero for the vertex v with itself need not be considered).

One can thus compute two information measures for each vertex v of G, one for the partition of $v(G)$ with respect to v and the other for the partition of $R(v)$. The information measures can be obtained as follows.

Scheme 1

Let

$$\begin{pmatrix} 1 & a_1 & a_2 & \cdots & a_e \\ p_0 & p_1 & p_2 & \cdots & p_e \end{pmatrix}, \qquad p_j > 0,$$

$$\sum_{j=0}^{e} p_j = 1 \qquad (11)$$

be a probability scheme associated with the partition of $V(G)$ with respect to v in G where the elements of the first row give the number of vertices at distances $0, 1, 2, \ldots, e$, from v and $p_0 = 1/n$, $p_i = a_i/n$, $i = 1, 2, \ldots, e$. One can now calculate the information content of the partition with the help of scheme (11) using Shannon's formula.[5] This measure of information will be called the vertex complexity of v and will be denoted by $V^c(v)$. Thus $V^c(v)$ is given by

$$V^c(v) = \frac{1}{n}\log_2 n - \sum_{i=1}^{e} p_i \log_2 p_i \qquad (12)$$

It is to be noted that $V^c(v)$ is analogous to $I_{c,r}$ defined by Balaban,[3] where such a measure was done for the centre(s) of a graph.

Scheme 2

Let

$$\begin{pmatrix} 1 & 2 & \cdots & e \\ a_1 & a_2 & \cdots & a_e \\ q_1 & q_2 & \cdots & q_e \end{pmatrix}, \qquad q_i > 0,$$

$$\sum_{i=1}^{e} a_i q_i = 1 \qquad (13)$$

be a probability scheme associated with the partition of $R(v)$ where the elements of first row give the distances $1, 2, \ldots, e$ occuring in G from v, the elements of second row give the number of times the distances $1, 2, \ldots, e$ have occured and $q_i = i/R(v)$.

Thus an information measure of the partition of $R(v)$ can be obtained using Shannon's formula[5] with the help of scheme (13). This measure of information will be called the vertex distance complexity of v and will be denoted by $V^d(v)$. Thus $V^d(v)$ is given by

$$V^d(v) = -\sum_{i=1}^{e} a_i \frac{i}{R(v)} \log_2 \frac{i}{R(v)} \qquad (14)$$

Finally, we compute two average information measures of G with the help of the V^c and V^d values of the vertices of G.

Let $V^c(v_1), V^c(v_2), \ldots, V^c(v_n)$ be the vertex complexities of $v_1, v_2, \ldots, v_n$, the n vertices of G.

We define

$$H^V = \frac{1}{n}\sum_{j=1}^{n} V^c(v_j) \qquad (15)$$

We call $H^V(G)$ the graph vertex complexity of G.

Again, let $V^d(v_1), V^d(v_2), \ldots, V^d(v_n)$ be the vertex distance complexities of $v_1, v_2, \ldots, v_n$ the n vertices of G.

Let us define

$$R(G) = \sum_{i=1}^{n} R(v_i) \qquad (16)$$

Considering $R(G)$ to be partitioned into $R(v_1)$, $R(v_2), \ldots, R(v_n)$, n parts, and since $V^d(v_i)$, $i = 1, 2, \ldots, n$, are the complexities obtained from the partitions of $R(v_i)$, an average information measure, called graph distance complexity (H^D) of G can be obtained by defining a probability

$$r_i = \frac{R(v_i)}{R(G)}, \qquad r_i > 0, \qquad \sum_{i=1}^{n} r_i = 1 \qquad (17)$$

where r_i is the probability of occurance of $R(v_i)$ in $R(G)$.

$$H^D(G) = \sum_{i=1}^{n} r_i V^d(v_i)$$

$$= \sum_{i=1}^{n} \frac{R(v_i)}{R(G)} V^d(v_i) \qquad (18)$$

G:

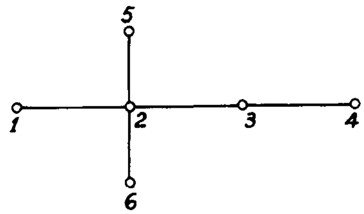

Figure 1.

It can be verified easily that the V^c and V^d values of the vertices of G can be calculated from the entries of the corresponding rows (or columns) in the distance matrix $D(G)$ and consequently $H^V(G)$ and $H^D(G)$ can be obtained.

Illustration:

Computation of I^d, H^V and H^D values of the graph shown in Figure 1 is illustrated below.

$I^d(G)$: There are four vertices of degree 1, one vertex of degree 2, and one vertex of degree 4 in G shown in Figure 1.

Thus

$$I^d(G) = 4\tfrac{1}{10}\log_2\tfrac{10}{1} + 1\tfrac{2}{10}\log_2\tfrac{10}{2} + 1\tfrac{4}{10}\log_2\tfrac{10}{4}$$
$$= 2.3219 \text{ bits}$$

$H^V(G)$ and $H^D(G)$: Let $D(G)$ be the distance matrix corresponding to the graph G shown in Figure 1. Then

$$D(G) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array}\begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{bmatrix} 0 & 1 & 2 & 3 & 2 & 2 \\ 1 & 0 & 1 & 2 & 1 & 1 \\ 2 & 1 & 0 & 1 & 2 & 2 \\ 3 & 2 & 1 & 0 & 3 & 3 \\ 2 & 1 & 2 & 3 & 0 & 2 \\ 2 & 1 & 2 & 3 & 2 & 0 \end{bmatrix} \end{array}$$

Therefore

$$D(v_i) = (1,1,3,1);\ R(v_1)$$
$$= (1 \times 1) + (2 \times 3) + (3 \times 1) = 10$$

$$V^c(v_1) = 3(1/6)\log_2(6/1) + (3/6)\log_2(6/3)$$
$$= 1.7925 \text{ bits}$$
$$V^d(v_1) = (1/10)\log_2(10/1) + 3(2/10)\log_2(10/2)$$
$$+ (3/10)\log_2(10/3)$$
$$= 2.2464 \text{ bits}$$

The values of different indices for the six vertices of the graph shown in Figure 1 are given in Table I.

Hence $H^V(G) = (1/6)(4 \times 1.7925) + 1.2516 + 1.4592 = 1.6468$ bits, and $H^D(G) = [3(10/56)(2.2464)] + [(6/56)(2.2516)] + [(8/56)(2.2500)] + [(12/56)(2.2296)] = 2.2439$ bits.

It may be noted that the V^c values of the vertices having degree frequency sequences $(1,1,3,1)$ and $(1,1,1,3)$ are same. This is due to the fact that the V^c value of a vertex depend only on the number of times any distance occurring in a graph with respect to that vertex but not on the magnitude of the distance. The magnitude is taken care of in the derivation of V^d and the vertices having the above two types of distance frequency sequences are distinguished by V^d.

A COMPARATIVE ANALYSIS ON DISCRIMINATION OF GRAPHS BY TOPOLOGICAL INDICES

To analyze the relative efficacy of different information theoretic topological indices in the discrimination of graphs, the values of the newly defined indices I^d, H^V, H^D and of the indices $I_{\text{top}}, \bar{I}_D^E, \bar{I}_D^W$ for the graphs given in Figures 2 and 3 have been furnished in Tables II and III, respectively. It emerges from the data that I^d is comparatively more powerful than I_{top} in dis-

Table I. Values of different indices for the vertices of G.

Indices / Vertices (v_i)	$D(v_i)$	$R(v_i)$	$V^c(v_i)$ (bits)	$V^d(v_i)$ (bits)
1	(1,1,3,1)	10	1.7925	2.2464
2	(1,4,1)	6	1.2516	2.2516
3	(1,2,3)	8	1.4592	2.2500
4	(1,1,1,3)	12	1.7925	2.2296
5	(1,1,3,1)	10	1.7925	2.2464
6	(1,1,3,1)	10	1.7925	2.2464

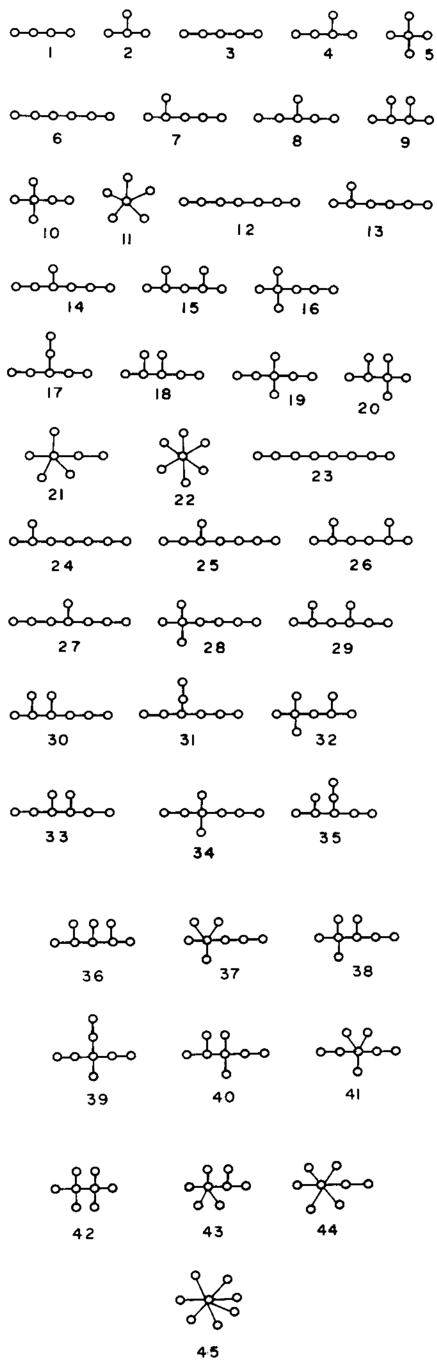

Figure 2.

criminating the graphs studied here. On the other hand, both H^V and H^D have discriminated those graphs well, while H^D seems to be more powerful than H^V or other indices in discriminating monocyclic graphs. Moreover, among the six topological indices, taken here for study, H^D has been found to be the only index to discriminate all the graphs given in Figures 2 and 3. The potentiality of H^V and H^D to discriminate graphs has been, perhaps, due to the fact that in the derivation of both the indices emphasis has been given on the measure of information

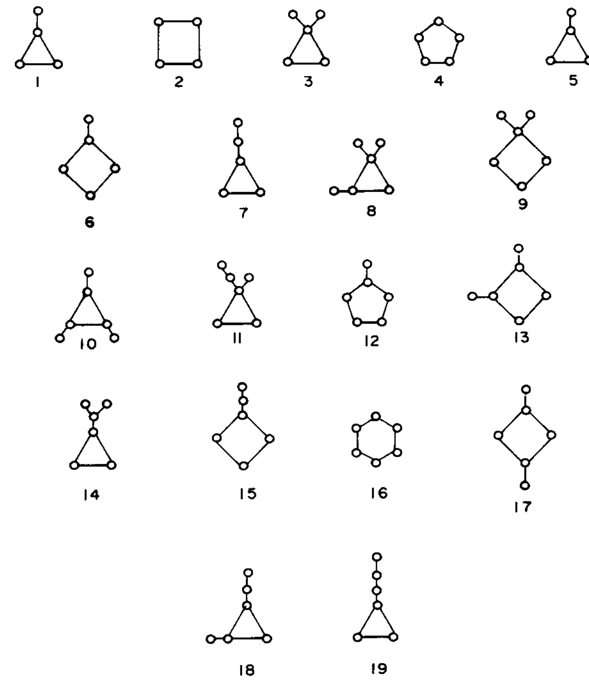

Figure 3.

content of each individual vertex of a graph. In the case of H^D, in particular, care has been taken to reflect the distribution of different distances occuring in a graph with respect to each of its vertices within the entire distribution of distances in the graph realized by the corresponding distance matrix. This seems to be an important factor, to be probed, to distinguish two nonisomorphic graphs. It is expected that H^D is capable of discriminating a large number of acyclic and cyclic graphs with more vertices and may find use in chemical documentation.

However, the degeneracy in the values of I_{top}, I^d, H^V and of $\bar{I}_D^E$ and $\bar{I}_D^W$ appears first for graphs of size 4 (i.e. graphs with four vertices). For graphs 1 and 2 shown in Figure 4 the value of I_{top} is 0.0000 and that of I^d is 2.0000. The index H^V has value 1.3278 for graph 2 shown in Figure 2 and graph 1 shown in Figure 3. While for graphs 1 and 2 shown in Figure 3 the value of $\bar{I}_D^E$ is 0.9183 and that of $\bar{I}_D^W$ is 2.5000. The degeneracy in the values of H^D appears first for graphs of size 5. As for example, the H^D value of graphs 3 and 4 shown in Figure 4 is 1.9196. It is worth mentioning that the J index of Balaban[23] can discriminate all the graphs shown in Figures 2 and 3 and the graphs 1, 2, 3 and 4 shown in Figure 4. But J index fails to discriminate the two pairs of graphs, viz. graphs 5, 6 and graphs 7, 8 shown in Figure 4 which are also the only two pairs of graphs among the 427 graphs con-

Table II. The values of information indices for the trees having four to eight vertices.

Graph*	I_{top}	I^d	H^V	H^D	$\bar{I}^{E**}$	$\bar{I}^{W***}$
n = 4						
1	1.0000	1.9183	1.7500	1.4755	1.4592	2.4464
2	0.8113	1.7925	1.3278	1.5324	1.0000	2.5033
n = 5						
3	0.9710	2.2500	2.0019	1.8558	1.3464	3.1464
4	1.3710	2.1556	1.7317	1.9064	1.5219	3.1972
5	0.7219	2.0000	1.2412	1.9564	0.9710	3.2500
n = 6						
6	0.9183	2.5219	2.2516	2.1513	2.1493	3.7043
7	1.4592	2.4464	1.9875	2.1380	1.9086	3.7417
8	1.4592	2.4464	1.9529	2.2087	1.8256	3.7600
9	0.9183	2.3710	1.7653	2.2324	1.5656	3.7883
10	1.2516	2.3219	1.6468	2.2439	1.5058	3.7979
11	0.6500	2.1610	1.1513	2.2851	0.9183	3.8439
n = 7						
12	0.8631	2.7516	2.4400	2.3959	2.3983	4.1690
13	1.4488	2.6887	2.2459	2.4195	2.2126	4.1953
14	1.4488	2.6887	2.2051	2.4429	2.1359	4.2102
15	1.3788	2.6258	2.00103	2.4509	1.9561	4.2304
16	1.3788	2.5850	1.9140	2.4650	1.9035	4.2405
17	1.4488	2.6887	2.00101	2.4686	1.9502	4.2406
18	1.3788	2.6258	1.9702	2.4751	1.8842	4.2512
19	1.3788	2.5850	1.8578	2.4916	1.7723	4.2656
20	1.1488	2.5221	1.7199	2.5047	1.5567	4.2845
21	1.1488	2.4508	1.5500	2.5189	1.4607	4.2964
22	0.5917	2.2925	1.0692	2.5532	0.8631	4.3366
n = 8						
23	0.8113	2.9502	2.6250	2.6011	2.6100	4.8016
24	1.4056	2.8963	2.4334	2.6180	2.4621	4.5870
25	1.4056	2.8963	2.4139	2.6335	2.3942	4.6049
26	1.5000	2.8424	2.2653	2.6392	2.2623	4.6120
27	1.4056	2.8963	2.3703	2.6475	2.3637	4.6120
28	1.4056	2.8074	2.1792	2.6517	2.2165	4.6201
29	1.5000	2.8424	2.2264	2.6622	2.2034	4.6322
30	1.5000	2.8424	2.1285	2.6690	2.1394	4.6364
31	1.4056	2.8963	2.1521	2.6710	2.2158	4.6352
32	1.5488	2.7534	1.8610	2.6806	1.9506	4.6539
33	1.5000	2.8424	1.99170	2.68437	2.1055	4.6515
34	1.4056	2.8074	2.1167	2.68443	2.0991	4.6498
35	1.5000	2.8424	1.99172	2.6954	1.9766	4.6625
36	0.9544	2.7884	1.9958	2.6963	1.9438	4.6679
37	1.2988	2.6924	1.8247	2.7000	1.8688	4.6675
38	1.5488	2.7534	1.9292	2.7049	1.8922	4.6751
39	1.4056	2.8074	1.8313	2.7098	1.8979	4.6751

Contd ....(ii).

Graph[a]	I_{top}	I^{d}	H^{V}	H^{D}	$\bar{I}^{E**}$	$\bar{I}^{W***}$
40	1.5488	2.7534	1.6202	2.7138	1.3238	4.6833
41	1.2988	2.6924	1.7561	2.7279	1.7120	4.6955
42	0.8113	2.6645	1.1474	2.7315	1.5502	4.7064
43	1.0613	2.6335	1.6506	2.7359	1.5303	4.7094
44	1.0613	2.5567	1.4566	2.7504	1.4052	4.7220
45	0.5436	2.4037	0.9966	2.7795	0.3113	4.7576

[a] Numbers correspond to the trees given in Figure 2.
[b] Data taken from Bonchev and Trinajstic (ref. 9).
[c] Data taken from Bonchev, Mekenyan, and Trinajstic (ref. 13).

Table III. The values of information indices for the monocyclic graphs having four to six vertices and edges.

Graph[a]	I_{top}	I^{d}	H^{V}	H^{D}	$\bar{I}^{E**}$	$\bar{I}^{W***}$
n = 4						
1	1.5000	1.9056	1.3278	1.5228	0.9183	2.5000
2	0.0000	2.0000	1.5000	1.5000	0.9183	2.5000
n = 5						
3	1.5219	2.1219	1.3015	1.9441	1.0000	3.2402
4	0.0000	2.3219	1.5219	1.9183	1.0000	3.2402
5	1.5219	2.1710	1.6215	1.9131	1.3610	3.2028
6	1.3710	2.2464	1.6517	1.8991	1.3610	3.2028
7	1.3710	2.2464	1.7317	1.8304	1.4855	3.1715
n = 6						
8	1.7925	2.3554	1.6122	2.2430	1.4295	3.7962
9	1.4592	2.4183	1.6122	2.2336	1.4295	3.7962
10	1.0000	2.3962	1.6888	2.2331	1.5219	3.7821
11	1.4592	2.4183	1.6387	2.2274	1.5219	3.7821
12	1.2516	2.5221	1.6888	2.2259	1.4295	3.7962
13	1.5850	2.4592	1.7653	2.2211	1.5219	3.7821
14	1.5850	2.4592	1.7653	2.2109	1.5656	3.7710
15	1.2516	2.5221	1.9529	2.1996	1.7819	3.7454
16	0.0000	3.5350	1.9183	2.1972	1.5219	3.7821
17	1.5850	2.4592	1.3344	2.1968	1.7056	3.7534
18	1.5850	2.4592	1.9319	2.1949	1.7819	3.7454
19	1.2516	2.5221	1.9875	2.1629	1.8392	3.7198

[a] Numbers correspond to the graphs given in Figure 3.
[b] Calculated following Bonchev and Trinajstic (ref. 9).
[c] Data taken from Bonchev, Mekenyan, and Trinajstic.

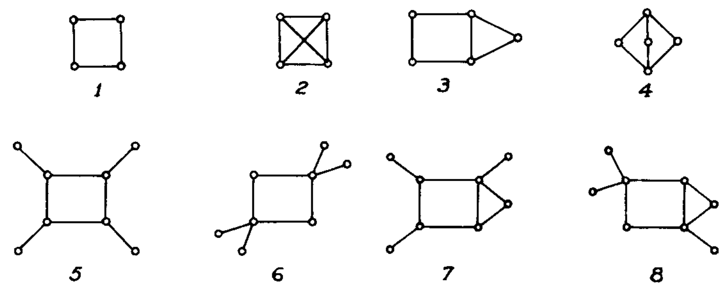

Figure 4.

Table IV. The values of seven topological indices for four cyclic graphs.

Graphs* \ Indices	I_{top}	I^d	$\bar{I}_D^E$	$\bar{I}_D^W$	H^V	H^D	J
5	1.0000	2.8113	1.8352	4.6729	1.9834	2.7020	2.4220
6	1.5000	2.7500	1.8424	4.6566	1.8400	2.6856	2.4220
7	1.8113	2.8219	1.7275	4.6773	1.8431	2.7046	2.1610
8	1.9056	2.7946	1.7660	4.6685	1.3221	2.6958	2.1610

[a] Numbers correspond to the graphs shown in fig. 4.

sidered by Bonchev, Mekenyan, and Trinajstic,[13] for which J has degeneracies. But it is interesting to note that each of the six information theoretic topological indices considered in this paper, can discriminate these two pairs of graphs. Hence a superindex composed of the index J and any one of the above mentioned six indices can discriminate all the 427 graphs considered in ref. 13. The values of six information theoretic topological indices, considered here, and those of J index for graphs 5, 6, 7, and 8 shown Figure 4 are given in Table IV.

References

1. L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic, New York, 1976.
2. A. Cammarata, *J. Pharm. Sci.*, **68**, 839 (1979).
3. A. T. Balaban, in *Steric Fit in QSAR*, Lecture Notes in Chemistry, Springer, Berlin, 1980.
4. F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1972.
5. C. Shannon and W. Weaver, *Mathematical Theory of Communication*, University of Illinois, Urbana, 1949.
6. N. Rashevsky, *Bull. Math. Biophys.*, **17**, 229, (1955).
7. E. Trucco, *Bull. Math. Biophys.*, **18**, 129, 237 (1956).
8. M. Valentinuzzi and M. E. Valentinuzzi, *Bull. Math. Biophys.*, **25**. 11 (1963).
9. D. Bonchev and N. Trinajstic, *J. Chem. Phys.*, **67**, 4517 (1977).
10. R. Sarkar, A. B. Roy, and P. K. Sarkar, *Math. Biosci.*, **39**, 299 (1978).
11. S. C. Basak, A. B. Roy, and J. J. Ghosh, in Proceedings of the 2nd International Conference on Mathematical Modeling, University of Missouri, Rolla, 1979, Vol. 2.
12. L. B. Kier, *J. Pharm. Sci.*, **69**, 807 (1980).
13. D. Bonchev, O. Mekenyan, and N. Trinajstic, *J. Comput. Chem.*, **2**, 127 (1981).
14. S. K. Ray, S. C. Basak, C. Raychaudhury, A. B. Roy, and J. J. Ghosh, Abst. Int. Conf. Communication Circuits and Systems, Jadavpur University, Calcutta, India, 1981, p. 12.
15. S. H. Bertz, *J. Am. Chem. Soc.*, **103**, 3599 (1981).
16. C. Raychaudhury, S. C. Basak, S. K. Ray, A. B. Roy, and J. J. Ghosh, Abst. 19th Annual Conference, Society of Engineering Sciences, University of Missouri, Rolla, 1982, p. 208.
17. A. B. Roy, C. Raychaudhury, J. J. Ghosh, S. K. Ray, and S. C. Basak, in Proceedings of the 4th European Symposium on Chemical Structure—Biological Activity: A Quantitative Approach, Elsevier, New York, 1982, p. 75.
18. S. K. Ray, S. C. Basak, C. Raychaudhury, A. B. Roy, and J. J. Ghosh, *IRCS Med. Sci.*, **10**, 933 (1982).
19. D. Bonchev, *Information Theoretic Indices for Characterization of Chemical Structures*, Wiley, Chichester, 1983.
20. S. C. Basak, D. P. Gieschen, V. R. Magnuson, and D. K. Harriss, *IRCS Med. Sci.*, **10**, 619 (1982).
21. S. C. Basak, D. P. Gieschen, D. K. Harriss, and V. R. Magnuson, *J. Pharm. Sci.*, **72**, 934 (1983).
22. S. K. Ray, S. C. Basak, C. Raychaudhury, A. B. Roy, and J. J. Ghosh, *Arzneim. Forsch. Drug Res.*, **33**(1), 352 (1983).
23. A. T. Balaban, *Chem. Phys. Lett.*, **89**, 399 (1982).
24. C. Raychaudhury, Ph.D. thesis, Jadavpur University, Calcutta, India, 1983.