 J. Chem. Inf. Comput. Sci. **2003,** *43,* 1860−1871

Discriminating Tests of Information and Topological Indices. Animals and Trees

Elena V. Konstantinova*

Sobolev Institute of Mathematics, Novosibirsk 630090, Russia

Maxim V. Vidyuk

Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia

Received December 20, 2002

In this paper we consider 13 information and topological indices based on the distance in a molecular graph with respect to their discrimination power. The numerical results of discriminating tests on 3 490 528 trees up to 21 vertices are given. The indices of the highest sensitivity are listed on the set of 1 528 775 alkane trees. The discrimination powers of indices are also examined on the classes of 849 285 hexagonal, 298 382 square, and 295 365 triangular simply connected animals. The first class of animals corresponds to the structural formulas of planar benzenoid hydrocarbons. The values of all indices were calculated for all classes of animals as well as for the united set of 1 443 032 animals. The inspection of the data indicates the great sensitivity of four information indices and one topological index.

1. INTRODUCTION

A single number representing a chemical structure is called a topological index. All structural formulas of chemical compounds are molecular graphs where vertices represent the set of atoms and edges represent chemical bonds. A topological index is designed by transforming a molecular graph into a number, and basically it expresses in numerical form the two-dimensional topology of the chemical species it presents. The construction and investigation of topological indices that could be used to describe molecular structures is one of the important directions of mathematical chemistry.[1−11] A number of such indices have been suggested in more than the last 50 years.[12−24] They usually correlate more or less satisfactorily with the molecular properties but could not discriminate well between structural isomers often providing one and the same value of a topological index for the different isomers. Bonchev and Trinajstić[25] applied information theory to the problem of characterizing molecular structures and molecular topology[26−29] by means of information indices[4] that are just the quantitative measure of a given topological structure. The advantage of such a kind of indices is in that they normally have greater discriminating power for isomers than the respective topological indices. The reasons for this are that information indices are not restricted to integral values as topological indices frequently are, and information indices are formed from a summation of different magnitudes which is usually greater in number than that for the corresponding topological indices.

The discrimination power[30,31] is one of the basic characteristics of a topological index I and corresponds to a measure of its ability to distinguish among the nonisomorphic graphs (the structural isomers) by distinct numerical values of index I. The theoretical evaluation of index sensitivity S on a fixed set M of nonisomorphic graphs can be achieved by the formula

$$S = \frac{N - N_I}{N} \qquad (1)$$

where $N = |M|$ is the number of graphs in a set M and N_I is the number of degeneracies of the index I within a set M. According to the definition, $S = 1$ means that among the elements of the set considered no two nonisomorphic graphs have the same value of the index I.

Bonchev and Trinajstić[25] investigated the discrimination power of information and topological indices between 45 alkane trees up to 8 vertices. Basak et al.[32] have continued these investigations on the set of 45 alkane trees as well as on the set of 19 monocyclic graphs. Razinger, Chretien, and Dubois[31] explicitly pointed out the fact that the discrimination power of the Wiener index[12] is very low in alkane series. Randić has shown[20] that the smallest degenerate alkane trees with the same molecular identification (ID) numbers occurred for $n = 15$ vertices. Müller et al.[33] have investigated the Balaban index (BID),[34] and it was shown that BID numbers are unique for alkane trees with up to 18 vertices.

At first discriminating tests among the polycyclic graphs were done by Konstantinova and Paleev[35] on the set of 1020 subgraphs of the regular square lattice.[36] Later Konstantinova[37] has examined information and topological indices for 2562 subgraphs of the regular hexagonal lattice. Graphs of this class represent the molecular structures of unbranched cata-condensed benzenoid hydrocarbons. The discrimination powers of topological and information indices as well as the Wiener polynomial derivatives were studied by Konstantinova and Diudea[38] for 3006 subgraphs of the regular hexagonal lattice and for the set of 347 cycle-containing graphs with 10 vertices and three- to eight-membered cycles.

* Corresponding author phone: 7 + 3832 + 331238; e-mail: e_konsta@math.nsc.ru.

10.1021/ci025659y CCC: $25.00
Published on Web 09/12/2003

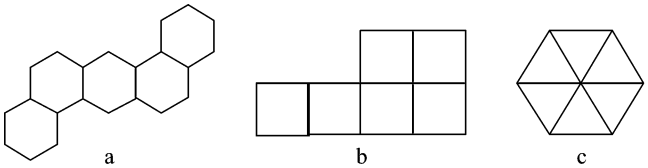


Figure 1. Simply connected hexagonal (a), square (b), and triangular (c) animals.

An exhaustive analysis of the weighting topological index based on the extended adjacency matrix was performed by Chang-Yu Hu and Lu Xu[39] for 3 807 434 alkane trees, 202 558 complex cyclic or polycyclic graphs, and 430 472 structures containing heteroatoms.

The present paper continues these investigations and gives the results of discriminating tests of 13 information and topological indices on the sets of trees and animals. We examine the indices on the set of 1 528 775 alkane trees as well as on the set of ordinary 3 490 528 trees up to 21 vertices. Alkane trees have a maximum vertex degree of at the most 4. There are no restrictions on the vertex degree for ordinary trees. We also study these indices on the set of 849 285 hexagonal, 298 382 square, and 295 365 triangular simply connected animals. The paper is organized in the following way. First of all, we define the classes of animals considered and give the formulas for the information and topological indices investigated. Then the numerical results of discriminating tests are presented.

2. CLASSES OF ANIMALS

The graph-theoretical problem known as the constructive enumeration of polycyclic graphs corresponds to the combinatorial problem known as the cell-growth problem that was formulated as the following one.[40−43] The name stems from an analogy with an animal which, starting from a single cell of some specified basic polygonal shape, grows step by step in the plane by adding at each step a cell of the same shape to its periphery. Thus if the basic shape is a regular hexagon, the animals are the hexagonal ones (Figure 1a). If the basic shape is a square or an equilateral triangle, we obtain square and triangular animals looking like those in Figure 1b,c. Animals are defined as simply connected ones if they have no holes and as multiply connected ones otherwise. All animals presented in Figure 1 are simply connected ones. Free animals are considered distinct if they have different shapes. Their orientation and location in the plane is no importance. Two free animals are isomorphic if one animal can be removed to another one under rotation and reflection conditions on the plane. From the graph-theoretical point of view simply connected animals embedded to the regular lattices are the subgraphs of infinite planar lattices.

In this paper we investigate free simply connected hexagonal, square, and triangular animals without and with internal vertices up to 12, 13, and 16 cells correspondingly.[44] The hexagonal animals that are also called polyhexes correspond to the structural formulas of planar polycyclic aromatic hydrocarbons.[45−48] That is the reason polyhexes have found a considerable interest in chemical enumerations.[49−65]

3. TOPOLOGICAL AND INFORMATION INDICES

We study 13 topological and information indices based on the distances within a graph. Let G be a finite connected graph without loops and multiple edges.[66] $V(G)$ is the set of vertices of a graph G with a cardinality $p = |V(G)|$, and $E(G)$ is the set of edges of a graph G with a cardinality $q = |E(G)|$. The distance $d(u,v)$ between vertices u and v in a graph G is the length of the shortest path that connects vertices u and v in a graph G.

Harry Wiener[12] proposed one of the first topological indices reflecting the topological structure of a molecular graph. The Wiener number W of a graph G is denoted by

$$W(G) = \frac{1}{2} \sum_{u,v \in V(G)} d(u,v) \tag{2}$$

The discriminating power of the Wiener number was investigated in many papers,[20,25,31,32,35,37,38,67] and there are the numbers of reviews[22,68−71] concerning its investigations and its chemical applications.

In 1994 Gutman[72] proposed to extend the formula that is satisfied for the Wiener index of trees to arbitrary graphs. The idea is the following one. Let us define $n_1(e)$ and $n_2(e)$ for an edge $e = (u,v)$ as $n_1(e) = |\{w|w \in V(G), d(w,u) < d(w,v)\}|$ and $n_2(e) = |\{w|w \in V(G), d(w,u) > d(w,v)\}|$. Then the Szeged index $Sz(G)$ of a graph G is equal to

$$Sz(G) = \frac{1}{2} \sum_{e \in E(G)} n_1(e) \cdot n_2(e) \tag{3}$$

where the summation embraces all edges of G. If G is a tree, then $Sz(G) = W(G)$. The basic results achieved in the theory and applications of the Szeged index were presented by Gutman and Dobrynin.[73] In particular, as it was mentioned in this review, the Sz and W values of benzenoid molecules were found to be well correlated, yet for very large benzenoids this correlation is curvilinear. This means that, for this class of polycyclic molecules, the structural information contained in Sz is quite similar to that contained in W. It will be confirmed also for animals by the present investigations.

One more topological index related to the Wiener number is called Schultz molecular topological index[74−76] and defined by

$$MTI(G) = \sum_{v \in V(G)} \deg(v) \cdot d(v) + \sum_{v \in V(G)} \deg(v)^2 \tag{4}$$

where $d(v)$ is the distance of a vertex v calculated by $d(v) = \sum_{u \in V(G)} d(v,u)$ and $\deg(v)$ is the vertex degree. This index has found interesting applications in chemistry.[22] Its discrimination power was considered by Dobrynin[77] for cata-condensed benzenoid graphs.

The average distance sum connectivity was introduced by Balaban[19] and defined by

$$J(G) = \frac{q}{q - p + 2} \sum_{u,v \in V(G)} (d(u) \cdot d(v))^{-1/2} \tag{5}$$

It was shown by Konstantinova[37] that the degeneracy of this index is very low for graphs of unbranched hexagonal systems. We will see the same results on the sets of animals considered.

The Randić index $\chi(G)$[15,20] is based on the molecular connectivity[6,7] and achieved by the formula

$$\chi(G) = \sum_{u,v \in V(G)} (\deg(u) \cdot \deg(v))^{-1/2} \quad (6)$$

It correlates well with biochemical properties[6] and has a good discrimination power for 45 alkane trees.[25]

For reducing degeneracies, Bonchev and Trinajstić[25] devised information theoretical indices. Let X be a set having n elements and let us assume that the elements are divided into N equivalence classes so that $n = \sum_{i=1}^{N} n_i$, where n_i is the number of elements in a subset X_i. Then the Shannon relation[78] is defined by

$$I = -\sum_{i=1}^{N} p_i \log_2 p_i \quad (7)$$

where $p_i = n_i/n$ is the probability for an element to belong to the ith subset.

The distance matrix and the Wiener number were considered to introduce the new indices applying this formula. Since the Wiener number is also given by $W = \sum_{i=1}^{d(G)} g_i \cdot i$, where g_i is the number of vertex pairs being at a distance i from each other and $d(G) = \max_{u,v \in V(G)} d(u,v)$ is the diameter of a graph G. Then following eq 7, one can obtain so-called[25] information on the realized distances in a given graph

$$\bar{I}_D^W = -\sum_{i=1}^{d(G)} g_i \cdot \frac{i}{W} \cdot \log_2 \frac{i}{W} \quad (8)$$

It was shown[25,37] that $\bar{I}_D^W$ has a great discrimination power among 45 alkane trees up to 8 vertices and 2562 unbranched hexagonal systems.

Actually the Shannon relation is called as the entropy in information theory.[78−80] D'yachkov and Konstantinova[81] define the entropy H_D, the marginal entropy H_D^i, and the information I_D based on the distance matrix $D = \|d_{ij}\|$, $i, j = 1, ..., p$, as follows:

$$H_D \equiv -\sum_{i=1}^{p} \sum_{j=1}^{p} \frac{d_{ij}}{2W} \cdot \log_2 \frac{d_{ij}}{2W} = 1 + \log_2 W - \frac{1}{W} \cdot \sum_{i=1}^{d(G)} g_i \cdot i \cdot \log_2 i \quad (9)$$

$$H_D^i \equiv -\sum_{i=1}^{p} \frac{d(i)}{2W} \cdot \log_2 \frac{d(i)}{2W} = 1 + \log_2 W - \frac{1}{2W} \cdot \sum_{i=1}^{p} d(i) \cdot \log_2 d(i) \quad (10)$$

$$I_D \equiv 2 \cdot H_D^i - H_D = 1 + \log_2 W - \frac{1}{W}\left(\sum_{i=1}^{p} d(i) \cdot \log_2 d(i) - \sum_{i=1}^{d(G)} g_i \cdot i \cdot \log_2 i\right) \quad (11)$$

In this paper we study the sensitivity of the marginal entropy H_D^i (this index was also presented in another form by Bonchev[4] as the information index on the distance degrees of the graph) and the information I_D.

The same approach is applied to the layer matrix $\lambda = \|\lambda_{ij}\|$, $i = 1, ..., p$, $j = 1, ..., d(G)$, where λ_{ij} is equal to the number of vertices located at distance j from vertex i. Let k_i be the number of matrix elements equal to i. The entropy H_λ, the marginal entropies H_λ^i and H_λ^j, and the information I_λ are defined by

$$H_\lambda \equiv -\sum_{i=1}^{p} \sum_{j=1}^{d(G)} \frac{\lambda_{ij}}{p(p-1)} \cdot \log_2 \frac{\lambda_{ij}}{p(p-1)} = \log_2 p(p-1) - \frac{1}{p(p-1)} \cdot \sum_{i=1}^{\max} k_i \cdot i \cdot \log_2 i \quad (12)$$

$$H_\lambda^i \equiv -p \cdot \frac{p-1}{p(p-1)} \cdot \log_2 \frac{p-1}{p(p-1)} = \log_2 p \quad (13)$$

$$H_\lambda^j \equiv -\sum_{j=1}^{d(G)} \frac{2g_j}{p(p-1)} \cdot \log_2 \frac{2g_j}{p(p-1)} = \log_2 p(p-1) - \frac{1}{p(p-1)} \cdot \sum_{j=1}^{d(G)} 2g_j \cdot \log_2 2g_j \quad (14)$$

$$I_\lambda \equiv H_\lambda^i + H_\lambda^j - H_\lambda = \log_2 p - \frac{1}{p(p-1)} \cdot \left(\sum_{j=1}^{d(G)} 2g_j \cdot \log_2 2g_j - \sum_{i=1}^{\max} k_i \cdot i \cdot \log_2 i\right) \quad (15)$$

The information indices $\bar{I}_D^W$, H_D, and H_λ^j are based on the vector g_i and on the constants p and W that lead to their correlations. In particular, since $\bar{I}_D^W = \log_2 W - (1/W) \cdot \sum_{i=1}^{d(G)} g_i \cdot i \cdot \log_2 i$, then following eq 9 we obtain that $H_D = \bar{I}_D^W + 1$. So it is enough to study the only index among them. We take $\bar{I}_D^W$ as the well-known information index.

One more information index based on the distance matrix was considered by Skorobogatov et al.[82] in the structure−activity correlations. The information index H_2 is defined by the formula

$$H_2 = -\sum_{i=1}^{n} \frac{d(i) \cdot n_i}{2W} \cdot \log_2 \frac{d(i) \cdot n_i}{2W} \quad (16)$$

where n_i, $i = 1, ..., n$, is the number of vertices having the distance $d(i)$. This index gives the linear correlations with information mass-spectrum indices on the several classes of organic and organometallic compounds.[83−86]

Two information indices having high discrimination powers were introduced by Konstantinova and Paleev[35] on the basis of the distance and layer matrices. Let us define the information distance index of a vertex i as follows

$$H_D(i) = -\sum_{j=1}^{p} \frac{d_{ij}}{d(i)} \cdot \log_2 \frac{d_{ij}}{d(i)} \quad (17)$$

Then the information distance index of graph vertices takes the form

$$H_D^p = \sum_{i=1}^{p} H_D(i) \quad (18)$$

Table 1. Number of Degeneracies on the Sets of N Hexagonal Animals with n Cells and i Internal Vertices[a]

$n-i-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
4−0−5										2			
5−0−12								2		8	2	4	2
5−1−6										5			
5−2−3										2			
6−0−36								10		34	21	18	21
6−1−24								4		21	4	6	11
6−2−14								4		11	2	2	4
6−3−4										4			
6−4−3										2			
7−1−106								20		104	37	58	81
7−2−68								12		66	14	20	31
7−3−25								2		22	2	4	4
7−4−10										10			2
7−5−3										2			
8−0−411			2	4			2	193		407	398	370	398
8−1−453								102		451	254	377	434
8−2−329				2			2	71		326	143	176	245
8−3−144								31		142	42	63	92
8−4−67								2		67	12	7	28
8−5−21								3		21	3	4	5
8−6−9										8		2	
9−0−1489			8	12			8	832		1486	1475	1457	1475
9−1−1966			2	6			4	542		1965	1463	1898	1944
9−2−1601		4		2		2	2	343	2	1598	1137	1259	1455
9−3−825							2	196		823	442	638	769
9−4−396	2	2	2	2		2	2	39	2	395	145	178	286
9−5−154								27		153	45	51	109
9−6−55										55	4	4	20
9−7−15										14		5	
9−8−4										4			
10−0−5572	18	18	31	47	2	20	26	3337	26	5569	5556	5534	5556
10−1−8395			20	40			4	2756		8391	7755	8336	8381
10−2−7652	2	6	6	16		12	12	1956	16	7650	7153	7209	7517
10−3−4518			4	8	2	2	6	1128	4	4516	3938	4378	4467
10−4−2340			6	6	4	6	6	370	6	2339	1812	1892	2186
10−5−1018								167		1018	584	824	968
10−6−416								32		414	160	173	319
10−7−123								17		123	24	58	69
10−8−42								4		41	4	4	11
10−9−9										8			
11−0−21115	20	24	175	303	6	26	108	13543	60	21110	21101	21085	21101
11−1−35885			98	188			66	13208		35883	35419	35806	35865
11−2−36109	12	63	46	127		69	44	10123	101	36105	35749	35740	35993
11−3−24020	24	24	40	62	4	12	34	6240	24	24017	23656	23903	23979
11−4−13415	36	44	38	40	16	32	26	2601	42	13413	13004	13091	13290
11−5−6405	8	8	10	10	6	4	6	1107	10	6404	5907	6265	6353
11−6−2811			2	2			2	352		2809	2257	2362	2663
11−7−1008			2	2				131	2	1007	538	796	957
11−8−333								27		332	107	145	223
11−9−100								8		99	16	37	54
11−10−26										25		2	2
12−0−81121	308	330	936	1587	86	418	555	53785	652	81118	81111	81087	81111
12−1−152688	2	2	577	1205	2	14	431	62447	26	152687	152323	152609	152664
12−2−168318	52	222	274	850	4	214	302	52027	389	168313	168004	167933	168188
12−3−124119	32	40	251	539	34	64	240	33653	122	124116	123827	123976	124082
12−4−74985	18	50	218	322	76	131	168	16921	229	74983	74644	74686	74860
12−5−38727	30	30	88	100	38	52	64	7329	82	38725	38372	38624	38683
12−6−18306	10	12	24	32	2	12	12	2722	18	18303	17914	17987	18195
12−7−7375	14	14	20	20	4	8	14	1001	14	7373	6855	7224	7326
12−8−2713			2	2			2	275		2709	2105	2196	2558
12−9−878								82		877	456	670	825
12−10−279								21		277	84	101	178
12−11−61								2		59	6	18	24
12−12−13								2		13	2	2	2

[a] No data means no degeneracies in Tables 1−3.

In the same way the information layer index of graph vertices is defined by the formula

$$H_\lambda^p = \sum_{i=1}^{p} H_\lambda(i) = -\sum_{i=1}^{p}\sum_{j=0}^{e(i)} \frac{\lambda_{ij}}{p}\cdot\log_2\frac{\lambda_{ij}}{p} \qquad (19)$$

where $e(i) = \max_{v\in V(G)} d(i,v)$ is the vertex eccentricity. The discrimination powers of H_D^p and H_λ^p were investigated by Konstantinova[37] and Diudea.[38] The information index H_D^p showed greater discrimination power regarding the structural isomers than the known topological indices.

Table 2. Number of Degeneracies on the Sets of *N* Square Animals with *n* Cells and *i* Internal Vertices

$n-i-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
5−0−11								4			4	4	4
6−0−27			2	4			4	17		4	17	13	23
6−1−7								3		2	2	2	3
7−0−82			4	6		2	4	54		45	60	67	76
7−1−21			2	2			2	14		11	10	16	15
8−0−250			4	18			12	187		190	211	220	242
8−1−90		2	2	4		2	2	69	2	67	57	74	85
8−2−21			2	2				10		9	7	6	13
9−0−815	2	4	44	105	4	4	38	639	2	726	755	782	810
9−1−334			14	26			14	267		289	270	321	330
9−2−89	2	2	10	19	2	2	10	54	2	66	52	59	76
9−3−9								5		3	2	2	5
10−0−2685	6	14	139	341	4	18	233	2185	16	2573	2609	2656	2676
10−1−1311	2	4	68	128	2	4	54	1102	4	1233	1246	1296	1306
10−2−391	2	6	29	50	4	10	21	277	10	337	308	344	379
10−3−67			9	17			14	48	2	50	33	46	61
10−4−6										2			
11−0−9072	18	30	464	1409	6	53	790	7647	57	8942	8990	9035	9065
11−1−4978	10	12	224	506	6	12	307	4286	14	4877	4936	4963	4972
11−2−1674	22	24	107	199	16	22	105	1292	22	1579	1603	1622	1658
11−3−324	4	4	26	40	4	4	22	243	4	278	244	300	317
11−4−45								25		34	13	25	29
12−0−30889	94	195	1394	5569	28	225	2588	26590	252	30732	30808	30854	30879
12−1−19030	14	34	800	2550	20	54	1446	16736	68	18893	18992	19012	19025
12−2−7089	40	50	375	795	32	85	476	5786	94	6950	7036	7048	7074
12−3−1630	22	28	119	195	6	24	127	1330	26	1548	1548	1610	1627
12−4−275		2	10	28		2	8	176	2	232	194	209	254
12−5−23								9		9	8	4	12
13−0−106290	334	679	5096	23710	110	682	11243	93199	807	106079	106209	106253	106276
13−1−72082	164	184	2985	12198	106	182	6172	64126	232	71909	72045	72045	72075
13−2−29433	254	278	1409	3813	192	365	2004	24790	375	29238	29381	29380	29412
13−3−7629	28	34	400	798	46	86	469	6443	98	7513	7570	7606	7620
13−4−1498	4	14	76	137	8	20	71	1089	22	1422	1417	1433	1479
13−5−174	2	2	4	8	2	2	4	136	2	150	116	151	169
13−6−11				2						4			2

Table 3. Number of Degeneracies on the Sets of *N* Triangular Animals with *n* Cells and *i* Internal Vertices

$n-i-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
4−0−3								2			2		2
6−0−11								8			4	2	8
7−0−23	2	2	4	4			4	17			9	12	19
8−0−62	6	6	8	12	4	4	12	54	4	6	31	46	59
8−1−4								2				2	2
9−0−148			14	30	2	2	20	136	2	12	98	133	146
9−1−11								7			2	2	8
10−0−405	12	12	55	92	5	8	39	372	9	54	351	387	402
10−1−38			4	6				31		6	20	21	38
11−0−1041	13	17	132	272	14	18	193	993	14	217	992	1030	1039
11−1−118	4	4	17	28	4	4	15	104		10	75	99	115
12−0−2825	42	58	413	832	26	53	559	2718	53	833	2773	2812	2821
12−1−386	14	15	64	101	8	13	60	362	6	60	337	378	383
12−2−15								9		4	7	2	9
13−0−7541	144	159	1106	2485	46	122	1899	7308	102	3101	7507	7531	7539
13−1−1189	32	44	187	311	18	40	192	1130	14	248	1149	1177	1186
13−2−54			2	3			2	41		8	24	38	52
14−0−20525	295	329	3008	8064	202	355	6016	20020	283	11083	20472	20510	20520
14−1−3691	106	129	672	1081	44	94	795	3559	46	1304	3649	3681	3685
14−2−233	4	8	39	51	2	6	35	202		59	181	205	230
15−0−55633	684	739	8352	24946	416	936	19556	54516	782	34376	55593	55622	55632
15−1−11223	254	285	1932	3828	144	321	2768	10912	166	5293	11168	11206	11221
15−2−841	18	26	146	201	6	16	119	781	6	276	786	821	840
15−3−19				2				12		2	2	7	13
16−0−152181	2404	2531	23417	84366	1114	2552	65020	149571	2146	96618	152137	152162	152177
16−1−33966	632	722	5790	13513	355	850	9968	33325	525	18929	33919	33959	33960
16−2−3059	67	81	510	810	53	92	517	2902	33	1407	2998	3045	3054
16−3−102			9	15			4	75		37	53	69	100

4. NUMERICAL RESULTS AND DISCUSSION

4.1. Animals. The values of 13 topological and information indices, introduced in the previous paragraph, were calculated for 849 285 hexagonal, 298 382 square, and 295 365 triangular simply connected animals. The obtained data are kept in the Data Base MySQL. The calculation accuracy for all indices is 10^{-13}.

At first the number of degeneracies for all indices were calculated on the sets of animals with *n* cells and *i* internal vertices. The data are given in Tables 1−3 for hexagonal, square, and triangular animals correspondingly. In all tables *N* is the number of animals in the set considered. There are no data for sets not having degeneracies. As one can see from the tables, the information indices H_D^p, I_D, H_D^i, and H_2

Table 4. Discrimination Power of Indices on the Sets of N Hexagonal Animals with n Cells

$n-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
1−1	1	1	1	1	1	1	1	1	1	1	1	1	1
2−1	1	1	1	1	1	1	1	1	1	1	1	1	1
3−3	1	1	1	1	1	1	1	1	1	1	1	1	1
4−7	1	1	1	1	1	1	1	1	1	0.714	1	1	1
5−22	1	1	1	1	1	1	1	0.909	1	0.318	0.909	0.818	0.909
6−81	1	1	1	1	1	1	1	0.778	1	0.111	0.667	0.679	0.556
7−331	1	1	1	1	1	1	1	0.755	1	0.033	0.526	0.438	0.329
8−1435	1	1	0.999	0.996	1	1	0.997	0.720	1	0.009	0.366	0.268	0.136
9−6505	0.999	0.999	0.998	0.997	1	0.999	0.997	0.696	0.999	0.002	0.271	0.101	0.049
10−30086	0.999	0.999	0.998	0.996	0.999	0.999	0.998	0.675	0.998	0.0006	0.039	0.026	0.010
11−141229	0.999	0.999	0.997	0.995	0.999	0.999	0.998	0.665	0.998	0.0002	0.014	0.005	0.002
12−669584	0.999	0.999	0.996	0.993	0.999	0.999	0.997	0.656	0.998	0.00005	0.001	0.001	0.0004

Table 5. Discrimination Power of Indices on the Sets of N Square Animals with n Cells

$n-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
1−1	1	1	1	1	1	1	1	1	1	1	1	1	1
2−1	1	1	1	1	1	1	1	1	1	1	1	1	1
3−2	1	1	1	1	1	1	1	1	1	1	1	1	1
4−5	1	1	1	1	1	1	1	1	1	1	1	1	1
5−12	1	1	1	1	1	1	1	0.667	1	1	0.667	0.667	0.667
6−35	1	1	0.943	0.886	1	1	0.886	0.429	1	0.829	0.457	0.571	0.257
7−107	1	1	0.944	0.925	1	0.981	0.944	0.364	1	0.477	0.346	0.224	0.150
8−363	1	0.994	0.978	0.934	1	0.994	0.961	0.267	0.994	0.267	0.242	0.174	0.063
9−1248	0.997	0.995	0.946	0.880	0.995	0.995	0.950	0.227	0.997	0.131	0.135	0.063	0.019
10−4460	0.998	0.995	0.945	0.880	0.998	0.993	0.928	0.190	0.993	0.059	0.047	0.025	0.008
11−16094	0.997	0.996	0.949	0.866	0.998	0.994	0.924	0.162	0.994	0.024	0.019	0.006	0.002
12−58937	0.997	0.995	0.954	0.845	0.999	0.993	0.921	0.141	0.993	0.010	0.003	0.003	0.0009
13−217117	0.996	0.995	0.954	0.813	0.998	0.994	0.908	0.126	0.993	0.004	0.001	0.0006	0.0002

Table 6. Discrimination Power of Indices on the Sets of N Triangular Animals with n Cells

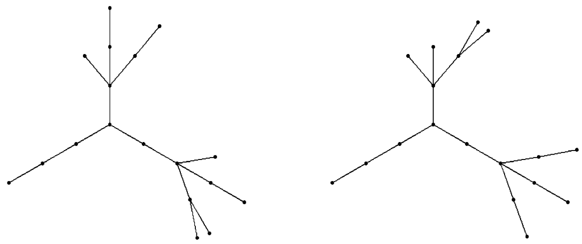
$n-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
1−1	1	1	1	1	1	1	1	1	1	1	1	1	1
2−1	1	1	1	1	1	1	1	1	1	1	1	1	1
3−1	1	1	1	1	1	1	1	1	1	1	1	1	1
4−3	1	1	1	1	1	1	1	0.333	1	1	0.333	1	0.333
5−4	1	1	1	1	1	1	1	1	1	1	1	1	1
6−12	1	1	1	1	1	1	1	0.333	1	1	0.667	0.833	0.333
7−24	0.917	0.917	0.833	0.833	1	1	0.833	0.292	1	1	0.625	0.500	0.208
8−66	0.909	0.909	0.879	0.818	0.939	0.939	0.818	0.152	0.939	0.909	0.530	0.273	0.076
9−159	1	1	0.912	0.811	0.987	0.987	0.874	0.101	0.987	0.925	0.371	0.151	0.031
10−444	0.973	0.973	0.867	0.779	0.989	0.982	0.912	0.092	0.980	0.865	0.164	0.059	0.009
11−1161	0.985	0.982	0.872	0.742	0.984	0.981	0.821	0.055	0.988	0.804	0.081	0.017	0.006
12−3226	0.983	0.977	0.852	0.711	0.989	0.980	0.808	0.042	0.982	0.722	0.033	0.007	0.004
13−8785	0.980	0.977	0.853	0.681	0.993	0.982	0.762	0.035	0.987	0.618	0.010	0.002	0.0008
14−24453	0.983	0.981	0.848	0.624	0.990	0.981	0.720	0.027	0.987	0.491	0.004	0.0009	0.0005
15−67716	0.986	0.984	0.846	0.572	0.992	0.981	0.669	0.022	0.986	0.410	0.001	0.0003	0.0001
16−189309	0.984	0.982	0.843	0.479	0.992	0.982	0.601	0.018	0.986	0.382	0.0005	0.0001	0.00004

Table 7. Discrimination Power of Indices on the Sets of N Hexagonal Animals with i Internal Vertices

$i-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
0−109883	0.997	0.997	0.990	0.982	0.999	0.996	0.994	0.347	0.993	0.0003	0.001	0.002	0.001
1−199525	0.999	0.999	0.997	0.993	0.999	0.999	0.997	0.604	0.999	0.00009	0.011	0.002	0.0007
2−214095	0.999	0.999	0.998	0.995	0.999	0.999	0.998	0.699	0.998	0.0001	0.007	0.008	0.002
3−153656	0.999	0.999	0.998	0.996	0.999	0.999	0.998	0.732	0.999	0.0001	0.011	0.004	0.001
4−91216	0.999	0.999	0.997	0.996	0.999	0.998	0.998	0.781	0.997	0.00008	0.016	0.015	0.006
5−46328	0.999	0.999	0.998	0.998	0.999	0.999	0.998	0.814	0.998	0.0001	0.031	0.012	0.004
6−21598	0.999	0.999	0.999	0.998	0.999	0.999	0.999	0.856	0.999	0.0004	0.058	0.050	0.019
7−8522	0.998	0.998	0.997	0.997	0.999	0.999	0.998	0.865	0.998	0.0006	0.130	0.052	0.020
8−3092	1	1	0.999	0.999	1	1	0.999	0.901	1	0.002	0.283	0.242	0.097
9−987	1	1	1	1	1	1	1	0.909	1	0.003	0.522	0.284	0.109
10−306	1	1	1	1	1	1	1	0.931	1	0.013	0.725	0.663	0.412
11−63	1	1	1	1	1	1	1	0.968	1	0.063	0.905	0.714	0.619
12−13	1	1	1	1	1	1	1	0.846	1		0.846	0.846	0.846
13−1	1	1	1	1	1	1	1	1	1	1	1	1	1

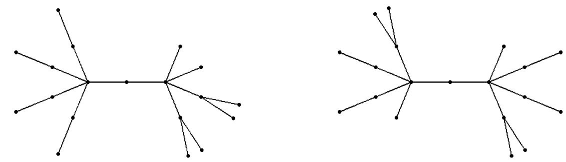
and the topological index J have the lowest degeneracies. The indices H_D^p and I_D are the most sensitive ones on the sets of hexagonal and square animals. The indices H_D^p and J show the minimal numbers of degeneracies for almost all sets of triangular animals.

Then the animals were united in the sets with the same number of cells and the same number of internal vertices for all considered classes of animals. The discrimination powers of indices were calculated in accordance with the formula (1), and the values were rounded to three digits after p = 18



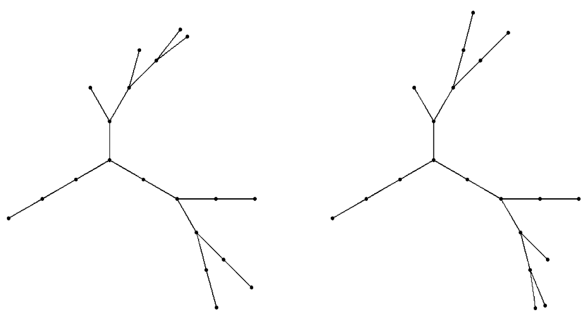
$H_D^p = 71.06719582653618$

p = 19

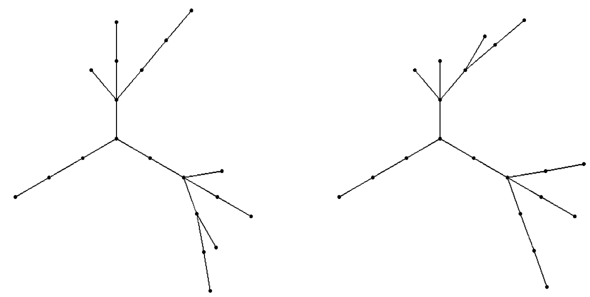


$H_D^p = 76.97820768153596$

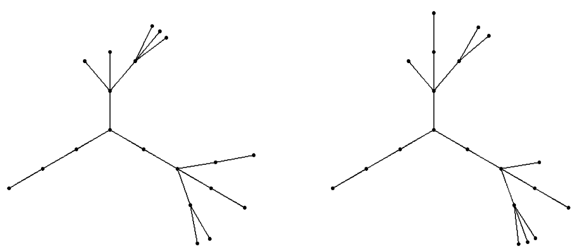
p = 20



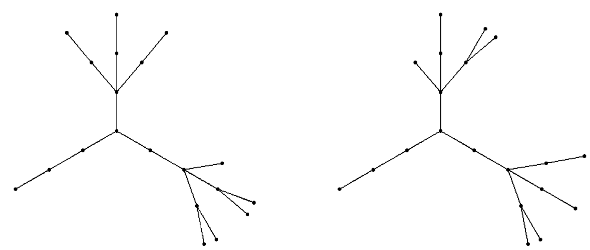
$H_D^p = 81.80354946997848$



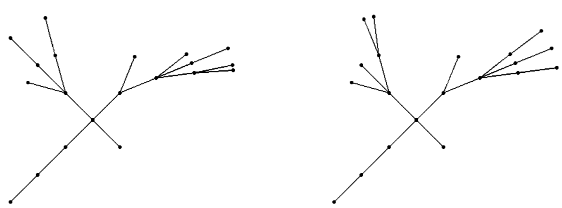
$H_D^p = 82.07740940300192$



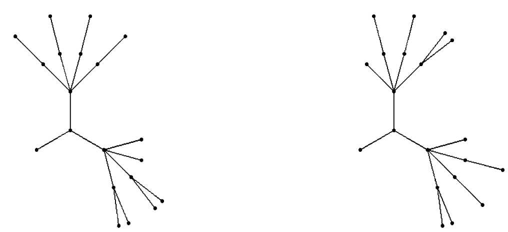
$H_D^p = 82.16707259857088$



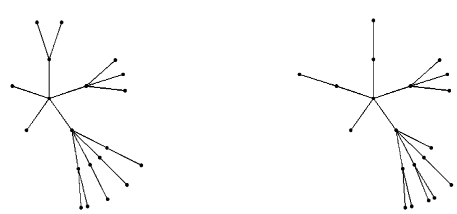
$H_D^p = 82.23982070813867$



$H_D^p = 82.46161593939024$

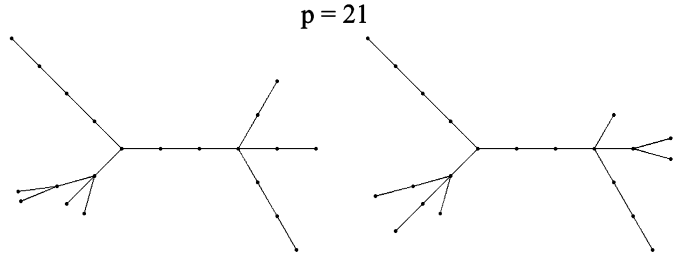


$H_D^p = 82.74262835086996$



$H_D^p = 83.25510903225370$

p = 21



$H_D^p = 87.30751562364020$

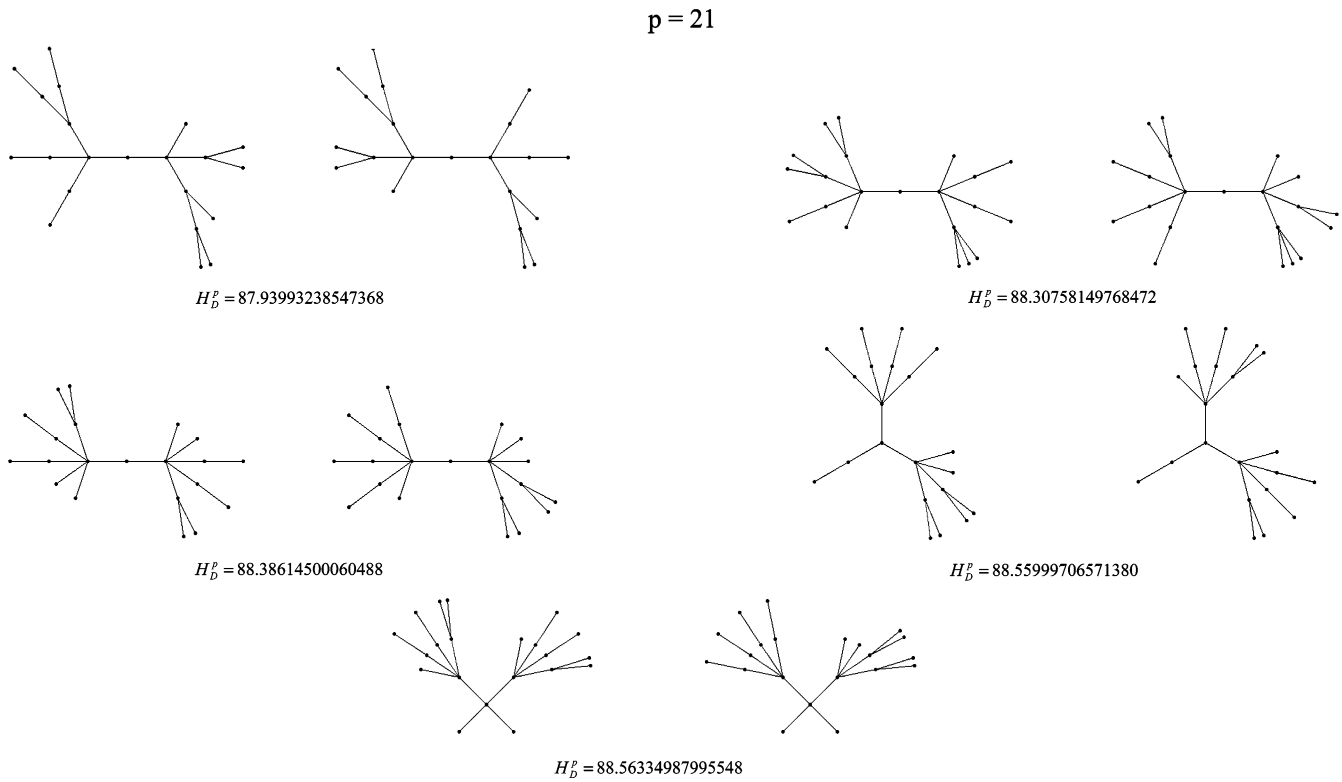


Figure 2. The pairs of trees having the same values of the information index H_D^p.

Table 8. Discrimination Power of Indices on the Sets of *N* Square Animals with *i* Internal Vertices

$i-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	*MTI*	*Sz*	*W*
0−150129	0.997	0.994	0.952	0.792	0.999	0.993	0.901	0.131	0.992	0.006	0.002	0.002	0.0005
1−97855	0.998	0.998	0.958	0.842	0.999	0.997	0.918	0.115	0.997	0.006	0.003	0.001	0.0003
2−38702	0.992	0.991	0.950	0.874	0.994	0.987	0.932	0.168	0.987	0.014	0.007	0.006	0.002
3−9661	0.994	0.993	0.943	0.891	0.994	0.988	0.935	0.165	0.987	0.028	0.027	0.010	0.003
4−1825	0.998	0.991	0.953	0.910	0.996	0.988	0.957	0.293	0.987	0.074	0.110	0.087	0.035
5−198	0.990	0.990	0.980	0.960	0.990	0.990	0.980	0.268	0.990	0.197	0.374	0.217	0.086
6−12	1	1	1	0.833	1	1	1	1	1	0.667	1	1	0.833

Table 9. Discrimination Power of Indices on the Sets of *N* Triangular Animals with *i* Internal Vertices

$i-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	*MTI*	*Sz*	*W*
0−240405	0.985	0.984	0.848	0.496	0.992	0.983	0.612	0.020	0.986	0.391	0.001	0.0003	0.0001
1−50628	0.979	0.976	0.829	0.627	0.989	0.974	0.727	0.024	0.985	0.489	0.004	0.001	0.0004
2−4205	0.979	0.973	0.834	0.747	0.985	0.973	0.840	0.064	0.991	0.583	0.050	0.014	0.004
3−126	1	1	0.929	0.865	1	1	0.968	0.310	1	0.690	0.563	0.397	0.103
4−1	1	1	1	1	1	1	1	1	1	1	1	1	1

the decimal point. The data are presented in Tables 4−6 (for the sets with *n* cells) and in Tables 7−9 (for the sets with *i* internal vertices). The real numbers of degeneracies could be easily obtained by the summation of corresponding data from Tables 1−3 for all indices except *MTI*, *Sz*, and *W*. There are degeneracies between the sets within Tables 1−3 for these indices. All another indices do not have degeneracies between the sets. As one can see from Table 4, the highest discrimination power, $S = 0.999$, corresponds to the information indices H_D^p, I_D, H_D^i, and H_2 for all sets of hexagonal animals. The square animals are discriminated well by these indices as well as the topological index *J*. Their discrimination powers range from 0.981 to 0.999 (see Table 5). It follows from Table 6 that indices H_D^p and I_D have the highest discrimination powers on the sets of triangular animals. Their discrimination powers range from 0.909 to 0.993, and they are lower than what we have on the sets of hexagonal and square animals. The Randić index χ has the lowest discrimination powers on the sets of hexagonal animals, and the Wiener number *W* gives the lowest discrimination powers on the sets of square and triangular animals.

The highest sensitivity, $S = 0.999$, is showed by index H_D^p for all sets of hexagonal animals with *i* internal vertices (see Table 7). This index is the best one for square animals as well as for triangular animals. The worst sensitivities are shown by the Randić index χ for hexagonal animals and by the Wiener number *W* for square and triangular animals.

The discrimination powers of indices on the sets of all considered hexagonal, square, and triangular animals are given in Tables 10−12. The data in the tables show that the information indices give a much more discrimination power. The indices H_D^p, I_D, H_D^i, and H_2 have the best results ($S = 0.999$) for 849.285 hexagonal animals. All topological indices, except *J*, could not discriminate well between these animals. The same situation is observed for square and

Table 10. Discrimination Power of Indices on the Set of N Hexagonal Animals

N	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
849285	0.999	0.999	0.997	0.993	0.999	0.999	0.997	0.659	0.998	0.0001	0.004	0.002	0.0006

Table 11. Discrimination Power of Indices on the Set of N Square Animals

N	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
298382	0.997	0.995	0.954	0.811	0.998	0.994	0.906	0.133	0.993	0.005	0.002	0.001	0.0003

Table 12. Discrimination Power of Indices on the Set of N Triangular Animals

N	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
295365	0.984	0.982	0.844	0.466	0.992	0.981	0.585	0.021	0.986	0.407	0.0008	0.0002	0.0001

Table 13. Discrimination Power of Indices on the Set of N Hexagonal, Square, and Triangular Animals

N	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	Sz	W
1443032	0.996	0.995	0.957	0.848	0.998	0.994	0.894	0.419	0.994	0.084	0.002	0.001	0.0003

Table 14. Discrimination Power of Indices on the Sets of N Ordinary Trees with p Vertices

$p-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	W
4−2	1	1	1	1	1	1	1	1	1	1	1	1
5−3	1	1	1	1	1	1	1	1	1	1	1	1
6−6	1	1	0.667	1	1	1	1	1	1	1	1	1
7−11	1	1	0.818	0.818	1	1	0.818	1	1	1	1	0.636
8−23	1	1	1	0.913	1	1	0.913	1	1	0.826	0.739	0.739
9−47	1	1	0.957	0.915	1	1	0.957	0.915	1	0.702	0.596	0.170
10−106	1	0.981	1	0.858	1	0.981	0.896	0.962	0.981	0.575	0.443	0.217
11−235	1	0.991	0.991	0.851	1	0.991	0.894	0.834	0.991	0.455	0.217	0.060
12−551	1	0.953	0.993	0.751	1	0.949	0.822	0.895	0.949	0.332	0.102	0.042
13−1301	1	0.983	0.989	0.765	1	0.983	0.836	0.836	0.983	0.248	0.032	0.012
14−3159	0.999	0.921	0.984	0.666	1	0.917	0.764	0.841	0.916	0.177	0.017	0.009
15−7741	0.999	0.968	0.988	0.676	1	0.967	0.748	0.791	0.967	0.125	0.007	0.002
16−19320	0.998	0.891	0.987	0.581	1	0.886	0.669	0.795	0.885	0.087	0.004	0.002
17−48629	0.998	0.957	0.987	0.556	1	0.956	0.654	0.747	0.955	0.060	0.001	0.0004
18−123867	0.997	0.869	0.986	0.465	0.999	0.861	0.555	0.750	0.859	0.040	0.0005	0.0003
19−317955	0.999	0.946	0.986	0.423	0.999	0.943	0.526	0.711	0.942	0.027	0.0002	0.00007
20−823065	0.997	0.847	0.985	0.337	0.999	0.838	0.438	0.702	0.837	0.018	0.0001	0.00004
21−2144505	0.999	0.934	0.984	0.282	0.999	0.931	0.393	0.664	0.930	0.012	0.00004	0.000009

Table 15. Number of Degeneracies on the Sets of N Ordinary Trees with p Vertices[a]

$p-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	W
4−2												
5−3												
6−6			2									
7−11			2	2			2					4
8−23				2			2			4	6	6
9−47			2	4			2	4		14	19	39
10−106		2		15		2	11	4	2	45	59	83
11−235		2	2	35		2	25	39	2	128	184	221
12−551		26	4	137		28	98	58	28	368	495	528
13−1301		22	14	306		22	214	214	22	979	1259	1286
14−3159	4	249	52	1056		261	747	502	265	2601	3105	3131
15−7741	2	250	92	2510		258	1947	1621	258	6776	7684	7724
16−19320	43	2115	250	8094		2196	6399	3958	2217	17641	19251	19289
17−48629	43	2080	631	21568		2164	16846	12295	2192	45690	48561	48610
18−123867	332	16211	1699	66310	2	17255	55131	30962	17453	118956	123800	123832
19−317955	284	17269	4505	183338	2	18175	150720	91967	18378	309482	317883	317932
20−823065	2766	126026	12461	545993	14	133071	462160	245143	134508	808323	822980	823029
21−2144505	2494	141282	34048	1538831	12	147986	1302117	720699	149854	2119187	2144429	2144486

[a] No data means no degeneracies in Tables 15 and 17.

triangular animals. The degeneracy is high for W, MTI, and Sz and very low for H_D^p, I_D, and J. The well-known information index $\bar{I}_D^W$ discriminates well among hexagonal animals, but it does not discriminate among square and triangular animals. The opposite situation is observed for the Randić index χ. Its discrimination power is the lowest one on hexagonal animals and the highest one on triangular animals (see Tables 10−12).

At last the resulting data on the united set of all 1 443 032 animals are presented in Table 13. It is easy to see that the discrimination power is very low for W, MTI, Sz, and χ and high for H_D^p, I_D, H_D^i, J, and H_2. Moreover there are no

Table 16. Discrimination Power of Indices on the Sets of N Alkane Trees with p Vertices

$p-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	W
4−2	1	1	1	1	1	1	1	1	1	1	1	1
5−3	1	1	1	1	1	1	1	1	1	1	1	1
6−5	1	1	0.600	1	1	1	1	1	1	1	1	1
7−9	1	1	1	0.778	1	1	0.778	1	1	1	1	0.556
8−18	1	1	1	0.889	1	1	0.889	1	1	0.778	0.778	0.778
9−35	1	1	1	0.886	1	1	0.943	0.943	1	0.600	0.657	0.171
10−75	1	1	1	0.800	1	1	0.853	0.973	1	0.427	0.560	0.253
11−159	1	1	1	0.805	1	1	0.868	0.830	1	0.283	0.296	0.057
12−355	1	0.966	0.989	0.648	1	0.966	0.758	0.904	0.966	0.158	0.144	0.054
13−802	1	0.993	0.985	0.661	1	0.993	0.771	0.848	0.993	0.086	0.056	0.015
14−1858	0.999	0.949	0.984	0.546	1	0.947	0.681	0.850	0.946	0.039	0.023	0.011
15−4347	1	0.982	0.986	0.536	1	0.981	0.628	0.802	0.981	0.019	0.011	0.003
16−10359	0.998	0.922	0.988	0.435	1	0.918	0.544	0.810	0.918	0.009	0.005	0.003
17−24894	0.999	0.972	0.986	0.401	1	0.971	0.505	0.764	0.970	0.005	0.002	0.0006
18−60523	0.998	0.906	0.986	0.326	0.999	0.899	0.413	0.773	0.897	0.002	0.0008	0.0005
19−148284	0.999	0.962	0.985	0.292	1	0.961	0.389	0.733	0.960	0.001	0.0004	0.0001
20−366319	0.997	0.886	0.984	0.237	0.999	0.880	0.327	0.734	0.878	0.0004	0.0002	0.00008
21−910726	0.999	0.954	0.982	0.200	0.999	0.951	0.292	0.697	0.950	0.0002	0.00007	0.00001

Table 17. Number of Degeneracies on the Sets of N Alkane Trees with p Vertices

$p-N$	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	W
4−2												
5−3												
6−5			2									
7−9				2			2					4
8−18				2			2			4	4	4
9−35				4			2	2		14	12	29
10−75				15			11	2		43	33	56
11−159				31			21	27		114	112	150
12−355		12	4	125		12	86	34	12	299	304	336
13−802		6	12	272		6	184	122	6	733	757	790
14−1858	2	94	30	844		98	593	278	100	1785	1816	1837
15−4347		78	60	2019		82	1619	862	82	4263	4300	4333
16−10359	19	812	128	5856		847	4720	1966	853	10264	10304	10333
17−24894	10	691	341	14916		727	12331	5887	737	24781	24836	24880
18−60523	139	5688	847	40770	2	6115	35551	13768	6205	60418	60472	60492
19−148284	106	5603	2282	104914		5845	90539	39659	5939	148142	148224	148267
20−366319	1183	41746	6021	279459	10	44000	246510	97579	44606	366190	366245	366289
21−910726	894	42338	16083	728741	4	44798	645178	276262	45552	910538	910662	910713

Table 18. Discrimination Power of Indices on the Set of Ordinary Trees up to 21 Vertices

N	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	W
3 490 528	0.998	0.912	0.985	0.321	0.999	0.907	0.428	0.683	0.907	0.017	0.00004	0.00002

Table 19. Discrimination Power of Indices on the Set of Alkane Trees up to 21 Vertices

N	I_D	H_D^i	I_λ	H_λ	H_D^p	H_2	H_λ^p	$\bar{I}_D^W$	J	χ	MTI	W
1 528 775	0.998	0.937	0.983	0.229	0.999	0.933	0.321	0.715	0.932	0.0006	0.0001	0.00004

degeneracies between classes of animals for indices H_D^p, I_D, H_D^i, and J.

4.2. Trees. The values of 12 topological and information indices (the Wiener index and the Szeged index give one and the same values for trees) presented in the previous paragraph were calculated for all 3 490 528 ordinary trees and all 1 528 775 alkane trees up to 21 vertices. The data obtained are kept in the Data Base MySQL. The calculation accuracy for all indices is 10^{-13}.

At first the discrimination powers of indices were calculated on the sets of ordinary trees having p vertices. The data are given in Table 14. The corresponding numbers of degeneracies for indices on these sets are presented in Table 15. In all tables N is the number of trees in the set considered. In Table 15 there are no data for sets not having degeneracies. As one can see from the tables, the information index H_D^p has the lowest degeneracy. There is no degeneracy of this index, i.e., $S = 1$, on the set of trees up to 17 vertices (the number of trees is $N = 81.134$). Two trees with 18 vertices and two trees having 19 vertices give the same values of H_D^p. There are only 14 degeneracies on the set of 20 vertices trees, $N = 823\,065$, and only 12 degeneracies on the set of 21 vertices trees, $N = 2\,144\,505$. All pairs of trees having the same values of this index are presented in Figure 2.

The same discriminating tests were done on the sets of alkane trees with p vertices. The obtained data on discrimination powers and numbers of degeneracies are given in Tables 16 and 17 correspondingly. The information index H_D^p shows the best sensitivity on alkane trees among the indices investigated. The best topological index with respect to the discrimination ability is the index J.

The discrimination powers of indices on the set of 3 490 528 ordinary trees and on the set of 1 528 775 alkane trees up to 21 vertices are given in Tables 18 and 19. The data in the tables show that the discrimination power is very low for W, MTI, and χ and high for H_D^p and I_D. All topological indices, except J, could not discriminate well between the trees.

Thus the presented results demonstrated that the information index H_D^p has a high discriminating ability regarding the structural alkane isomers. This index could be used for characterizing molecular structures.

ACKNOWLEDGMENT

This work was supported by the Russian Fund of Fundamental Investigations through Grant No. 02-01-01118 and partially was supported by Com^{2}MaC-KOSEF, The Republic of Korea. The authors are grateful to the referees for their constructive remarks and comments.

REFERENCES AND NOTES

(1) *Chemical Applications of Graph Theory*; Balaban, A. T., Ed.; Academic Press: London, 1976.
(2) Balaban, A. T. Applications of graph theory in chemistry. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 334−343.
(3) Basak, C.; Niemi, G. J.; Veith, G. D. Predicting properties of molecules using graph invariants. *J. Math. Chem.* **1991**, *7*, 243−272.
(4) Bonchev, D. *Information Theoretic Indices for Characterization of Chemical Structures*; Research Studies Press: Chichester, 1983; p 249.
(5) Gutman, I.; Polansky, O. E. *Mathematical Concepts in Organic Chemistry*; Springer: Berlin, 1986.
(6) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: London, 1976.
(7) Kier, L. B.; Hall L. H. *Molecular Connectivity in Structure--Activity Analysis*; Research Studies Press: Letchworth, 1986.
(8) *Graph Theory and Topology in Chemistry*; King, R. B., Rouvray, D. H., Eds.; Elsevier: Amsterdam, 1987.
(9) Rouvray, D. H. In *Chemical Applications of Topology and Graph Theory. Studies in Physical and Theoretical Chemistry*; King, R. B., Ed.; Elsevier: Amsterdam, 1983; Vol. 28, pp 159−177.
(10) Trinajstić, N., Ed. *Mathematics and Computational Concepts in Chemistry*; Ellis Horwood: Chichester, 1986.
(11) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, FL, 1992.
(12) Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **1947**, *69*, 17−20.
(13) Wiener, H. Vapor pressure−temperature relationships among the branched paraffin hydrocarbons. *J. Chem. Phys.* **1948**, *52*, 425−430.
(14) Hosoya, H. Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332−2339.
(15) Randić, M. On characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *69*, 6609−6615.
(16) Entringer, R. C.; Jackson, D. E.; Snyder, D. A. Distance in graphs. *Czechoslovak Math. J.* **1976**, *2*, 283−297.
(17) Doyle, J. K.; Graver, J. E. Mean distance in graph. *Discrete Math.* **1977**, *17*, 147−154.
(18) Kier, L. B.; Hall, L. H. Derivation and significance of valence molecular connectivity. *J. Charm. Sci.* **1981**, *70*, 583−589.
(19) Balaban, A. T. Topological indices based on topological distances in molecular graphs. *Pure Appl. Chem.* **1983**, *55*, 199−206.
(20) Randić, M. On molecular identification numbers. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 164−175.
(21) Randić, M. Generalized molecular descriptors. *J. Math. Chem.* **1991**, *7*, 155−168.
(22) Nikolić, S.; Trinajstić, N.; Mihalić, Z. The Wiener index: developments and applications. *Croat. Chem. Acta* **1995**, *68*, 105−129.
(23) Entringer, R. C. Distance in graphs: trees. *J. Combin. Math. Combin. Comput.* **1997**, *24*, 65−84.
(24) Diudea, M.; Gutman, I. Wiener-type topological indices. *Croat. Chem. Acta* **1998**, *71*, 21−51.
(25) Bonchev, D.; Trinajstić, N. Information theory, distance matrix, and molecular branching. *J. Chem. Phys.* **1977**, *38*, 4517−4533.
(26) Bonchev, D.; Trinajstić, N. On topological characterization of molecular branching. *Int. J. Quantum Chem.* **1978**, *S12*, 293−303.
(27) Bonchev, D.; Knop, J. V.; Trinajstić, N. Mathematical models of branching. *MATCH* **1979**, *6*, 21−47.
(28) Bonchev, D.; Mekenyan, O.; Trinajstić, N. Topological characterization of cyclic structure. *Int. J. Quantum Chem.* **1980**, *17*, 845−893.
(29) Bonchev, D.; Trinajstić, N. Chemical information theory. Structural aspects, *Intern. J. Quantum Chem. Symp.* **1982**, *16*, 463−480.
(30) Bonchev, D.; Mekenyan, O.; Trinajstić, N. Isomer discrimination by topological information approach. *J. Comput. Chem.* **1981**, *2*, 127−148.
(31) Razinger, M.; Chretien, J. R.; Dubois, J. K. Structural selectivity of topological indices in alkane series. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 23−27.
(32) Raychaudhary, C.; Ray, S. K.; Ghosh, J. J.; Roy, A. B.; Basak, S. C. Discrimination of isomeric structures using information theoretic topological indices. *J. Comput. Chem.* **1984**, *5*, 581−588.
(33) Müller, W. R.; Szymanski, K.; Knop, J. V.; Miholić, Z.; Trinajstić, N. The walk ID numbers revisited. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 231−233.
(34) Balaban, A. T. Numerical modelling of chemical structures: local graph invariants and topological indices. In *Graph Theory and Topology*; King, R. B., Rouvray, D. H., Eds.; Elsevier: Amsterdam, 1987; pp 159−176.
(35) Konstantinova, E. V.; Paleev, A. A. Sensitivity of topological indices of polycyclic graphs. *Vychisl. Sistemy* **1990**, *136*, 38−48 (Russian).
(36) Müller, W. R.; Szymanski, K.; Knop, J. V.; Trinajstić, N. On the number of square-cell configurations. *Theor. Chim. Acta* **1993**, *86*, 269−278.
(37) Konstantinova, E. V. The discrimination ability of some topological and information distance indices for graphs of unbranched hexagonal systems. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 54−57.
(38) Konstantinova, E. V.; Diudea, M. V. The Wiener polynomial derivatives and other topological indices in chemical research. *Croat. Chem. Acta* **2000**, *73*, 383−403.
(39) Chang-Yu, H.; Lu, X. On highly discriminating molecular topological index. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 82−90.
(40) Harary, F. Graphical enumeration problems. In *Graph Theory and Theoretical Physics*; Academic Press: London, 1967; Chapter 1, pp 1−41.
(41) Harary, F.; Palmer, E. M. *Graphical Enumeration*; Academic Press: New York, 1973.
(42) Pólya, G.; Read, R. C. *Combinatorial Enumeration of Groups, Graphs and Chemical Compounds*; Springer-Verlag: New York, 1987.
(43) Klarner, D. A. Polyominoes. In *Handbook of Discrete and Computational Geometry*; Goodman, J. E., et al., Eds.; CRC Press Series on Discrete Mathematics and its Applications: Boca Raton, FL, 1997; pp 225−240.
(44) Konstantinova, E. V.; Vidyuk, M. V. Fast generation of animals. Manuscript in preparation.
(45) Clar, E. *The Aromatic Sextet*; Wiley: London, 1972.
(46) Dias, J. R. *Handbook of Polycyclic Hydrocarbons. Part A. Benzenoid Hydrocarbons*; Elsevier: Amsterdam, 1987.
(47) Gutman, I.; Cyvin, S. J. *Introduction to the Theory of Benzenoid Hydrocarbons*; Springer-Verlag: Berlin, 1989.
(48) *Advances in the Theory of Benzenoid Hydrocarbons*; Gutman, I., Cyvin, S. J., Eds.; Springer-Verlag: Berlin, 1990.
(49) Balaban, A. N.; Harary, F. Enumeration and proposed nomenclature of benzenoid cata-condensed polycyclic aromatic hydrocarbons. *Tetrahedron* **1968**, *24*, 2505−2516.
(50) Trinajstić, N.; Jericević, Z.; Knop, J. V.; Müller, W. R.; Szymanski, K. Computer generation of isomeric structures. *Pure Appl. Chem.* **1983**, *55*, 379−390.
(51) Knop, J. V.; Szymanski, K.; Jericević Z.; Trinajstić, N. On the total number of polyhexes. *MATCH* **1984**, *16*, 119−134.
(52) Stojmenović, I.; Tošić R.; Doroslavčki, R. An algorithm for generating and counting hexagonal systems. In *Proc. 6 Yugoslav Seminar on Graph Theory*; Dubrovnik, 1985; pp 189−198.
(53) Knop, J. V.; Müller, W. R.; Szymanski, K.; Trinajstić, N. *Computer Generation of Certain Classes of Molecules*; SKTN: Zagreb, 1985.
(54) Wenchen, He; Wenjie, He. Generation and enumeration of planar polycyclic aromatic hydrocarbons. *Tetrahedron* **1986**, *42*, 5291−5299.
(55) Cyvin, B. N.; Brunvoll, J.; Cyvin, S. J.; Gutman, I. All-benzenoid systems: enumeration and classification of benzenoid hydrocarbons. VI. *MATCH* **1988**, *23*, 163−173.
(56) Trinajstić, N.; Nikolić, S.; Knop, J. V.; Müller, W. R.; Szymanski, K. *Computational chemical graph theory: characterization, enumeration and generation of chemical structures by computer methods*; Simon & Schuster: New York, 1991.
(57) Trinajstić, N. On the classification of polyhexes. *J. Math. Chem.* **1992**, *9*, 373−380.
(58) Cyvin, S. J. Generalization of extremal hexagonal animals (polyhexes). *J. Math. Chem.* **1992**, *9*, 389−390.
(59) Cyvin, B. N.; Brunvoll, J.; Rongsi Chen; Cyvin, S. J. Coronenic coronoids: a course in chemical enumeration. *MATCH* **1993**, *29*, 131−142.
(60) Cyvin, B. N.; Zhang, F.; Guo, X.; Brunvoll, J.; Cyvin, S. J. On the total number of polyhexes with 10 hexagons. *MATCH* **1993**, *29*, 143−163.
(61) Cyvin, S. J.; Cyvin, B. N.; Brunvoll, J. Polycyclic conjugated hydrocarbons with arbitrary ring sizes. *J. Mol. Struct.* **1993**, *300*, 9−22.

(62) Cyvin, B. N.; Brunvoll, J.; Cyvin, S. J. Isomer enumeration of unbranched cata-condensed polygonal systems with pentagons and heptagons. *MATCH* **1996**, *34*, 109−121.
(63) Konstantinova, E. V. The constructive enumeration of square animals, preprint at Com[2]MaC, POSTECH, 2000; Vol. 10, 55pp. (http://com2mac.postech.ac.kr/resorce/pre00_text.htm).
(64) Konstantinova, E. V. The constructive enumeration of triangular animals, preprint at Com[2]MaC, POSTECH, 2000; Vol. 21, 33 pp. (http://com2mac.postech.ac.kr/resorce/pre00_text.htm).
(65) Konstantinova, E. V. Enumeration and generation of animals. In *Proc. 2000 Year-End Symposium on Combinatorial and Computational Mathematics*; Pohang, 2001; pp 53−67.
(66) Harary, F. *Graph Theory*; Addison-Wesley Publishing Company: London, 1971.
(67) Lepović, M.; Gutman, I. A collective property of trees and chemical trees. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 626−629.
(68) Gutman, I.; Yeh, Y. N.; Lee, S. L.; Luo, Y. L. Some recent results in the theory of the Wiener number. *Indian J. Chem.* **1993**, *32A*, 651−661.
(69) Gutman, I.; Potgiete, J. H. Wiener number and intermolecular forces. *J. Serb. Chem. Soc.* **1997**, *62*, 185−192.
(70) Dobrynin, A. A.; Entringer, R.; Gutman I. Wiener index of trees: theory and applications. *Acta Appl. Math.* **2001**, *66*, 211−249.
(71) *Topology in Chemistry: Discrete Mathematics of Molecules*; Rouvray, D. H., King, R. B., Eds.; Proceedings of the International Wiener Conference: Horwood, Chichester, U.K., 2002.
(72) Gutman, I. A formula for the Wiener number of trees and its extension to graph containing cycles. *Graph Theory Notes of New York* **1994**, *27:2*, 9−15.
(73) Gutman, I.; Dobrynin, A. A. The Szeged index − a success story. *Graph Theory Notes New York* **1998**, *34*, 37−44.
(74) Schultz, H. P. Topological organic chemistry. 1. Graph theory and topological indices of alkanes. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 227−228.
(75) Müller, W. R.; Szymanski, K.; Knop, J. V.; Trinajstić, N. Molecular topological index. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 160−163.
(76) Gutman, I. Selected properties of the Schultz molecular index. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1087−1089.
(77) Dobrynin, A. A. Discriminating power of the Schultz index for cata-condensed benzenoid graphs. *MATCH* **1998**, *38*, 19−32.
(78) Shannon, C.; Weaver, W. *Mathematical Theory of Communications*; University of Illinois: Urbana, 1949.
(79) Ashby, W. R. *An Introduction to Cybernetics*; Wiley: New York, 1956.
(80) Brillouin, L. *Science and Information Theory*; Academic: New York, 1956.
(81) D'yachkov, A. G.; Konstantinova, E. V. On entropy and information of trees. Manuscript in preparation.
(82) Skorobogatov, V. A.; Konstantinova, E. V.; Nekrasov, Yu. S.; Sukharev, Yu. N.; Tepfer, E. E. On the correlation between the molecular information topological and mass-spectra indices of organometallic compounds. *MATCH* **1991**, *26*, 215−228.
(83) Sukharev, Yu. N.; Nekrasov, Yu. S.; Molgacheva, N. S.; Tepfer, E. E. Computer processing and interpretation of mass-spectral information. Part IX − Generalized characteristics of mass-spectra. *Org. Mass Spectrom.* **1993**, *28*, 1555−1561.
(84) Nekrasov, Yu. S.; Tepfer, E. E.; Sukharev, Yu. N. On the relationship between the mass-spectral and structural indices of arylsilanes. *Russ. Chem. Bull.* **1993**, *42*, 343−346.
(85) Nekrasov, Yu. S.; Sukharev, Yu. N.; Molgacheva, N. S.; Tepfer, E. E. Generalized characteristics of mass-spectra of aromatic compounds and their correlation with the constants of substituents. *Russ. Chem. Bull.* **1993**, *42*, 1986−1990.
(86) Nekrasov, Yu. S.; Sukharev, Yu. N.; Tepfer, E. E.; Molgacheva, N. S. Establishment of correlations between the structure and reactivity of molecules in the gas-phase based on information theory. *Russ. Chem. Bull.* **1996**, *45*, 2542−2546.

CI025659Y