CROATICA CHEMICA ACTA CCACAA **71** (3) 473–488 (1998)

ISSN-0011-1643
CCA-2506

Original Scientific Paper

Szeged Fragmental Indices

Ovidiu M. Minailiuc,[a] *Gabriel Katona,*[a] *Mircea V. Diudea*[a,*]
Mate Strunje,[b] *Ante Graovac,*[c] *and Ivan Gutman*[d]

[a]*Department of Chemistry, The Babes-Bolyai University, R-3400 Cluj, Romania*

[b]*Faculty of Chemical Engineering and Technology, Marulićev trg 19, HR-10000 Zagreb, Croatia*

[c]*Rugjer Bošković Institute, P.O.Box 1016, HR-10001 Zagreb, Croatia*

[d]*Institute of Physical Chemistry, The Attila Jozsef University, H-6701 Szeged, P.O.Box 105, Hungary*

Received February 23, 1998; revised August 8, 1998; accepted August 8, 1998

Novel Szeged indices, defined on unsymmetric matrices, which collect various fragmental topological properties, are proposed. They are illustrated on selected graphs possessing heteroatoms, multiple bonds and stereoisomers. Correlations on organic structures with herbicidal activity and explosive properties support the usefulness of these newly proposed indices.

INTRODUCTION

Wiener index,[1] W, one of the most studied topological indices (see some of the reviews)[2–4] is connected with the problem of distances in graph. In acyclic structures, the Wiener index can be defined by

$$\boldsymbol{W} = \boldsymbol{W}(\Gamma) = \sum_{\text{all}(i,j)\in \mathrm{E}(\Gamma)} N_{i(i,j)}\, N_{j(i,j)} \qquad (1)$$

where $N_{i,(i,j)}$ and $N_{j,(i,j)}$ denote the number of vertices lying on the two sides of the edge $(i,j)\in \mathrm{E}(\Gamma)$, $\mathrm{E}(\Gamma)$, being the set of edges in a connected graph, Γ. The summation runs over all edges in Γ. The product $N_{i,(i,j)}\, N_{j,(i,j)}$ represents

* Author to whom correspondence should be addressed.

the contribution of edge *(i,j)* to the global index, *W*. It is just the *(i,j)*-entry in the edge-defined Wiener matrix[5,6]

$$[\boldsymbol{W}_e]_{ij} = N_{i,(i,j)}\, N_{j,(i,j)} \; ; \quad (i,j) \in E(\Gamma) \tag{2}$$

from which *W* can be calculated as a half sum of its entries

$$W = (1/2)\, \Sigma_i \Sigma_j\, [\boldsymbol{W}_e]_{ij} \; . \tag{3}$$

Note that for non-adjacent vertices *i* and *j*, $(i,j) \notin E(\Gamma)$, the non-diagonal entries in $\boldsymbol{W}_e$ are zero. When *(i,j)* represents a path, $(i,j) \in P(\Gamma)$, with $P(\Gamma)$ being the set of paths in graph, then a relation similar to Eq. (1) will define the *hyper-Wiener* index,[7] *WW*

$$WW = WW(\Gamma) = \sum_{\text{all}(i,j) \in P(\Gamma)} N_{i(i,j)}\, N_{j(i,j)} \; . \tag{4}$$

The summation runs over all paths in Γ. $N_{i,(i,j)}$ and $N_{j,(i,j)}$ refer now to the number of vertices lying on the two sides of the path $(i,j) \in P(\Gamma)$ and the product $N_{i,(i,j)} N_{j,(i,j)}$ is the contribution of the path *(i,j)* to the global index, *WW*. It is the *(i,j)*-entry in the path-defined Wiener matrix[5,6]

$$[\boldsymbol{W}_p]_{ij} = N_{i,(i,j)}\, N_{j,(i,j)} \; ; \quad (i,j) \in P(\Gamma) \; . \tag{5}$$

From W_p the *WW* index is calculated as a half sum of its entries

$$WW = (1/2)\, \Sigma_i \Sigma_j\, [\boldsymbol{W}_p]_{ij} \; . \tag{6}$$

In both $\boldsymbol{W}_e$ and $\boldsymbol{W}_p$ matrices, the diagonal entries are zero.

In cycle-containing graphs, Wiener matrices are not defined. Wiener indices are here calculated by means of the distance-type matrices.[8,9]

The idea of extending the validity of relations (1) to (6) in cycle-containing graphs supplied some Wiener analogue indices,[10–13] one of them coming from Gutman.[14–20]

Szeged index, *SZ*, is defined[14] according to Eq. (1) but the quantities $N_{i,(i,j)}$ and $N_{j,(i,j)}$ are defined so that definition in Eq. (1) remains valid both in acyclic and cycle-containing graphs

$$N_{i,(i,j)} = |\{v | v \in V(\Gamma);\ (i,j) \in E(\Gamma);\ D_{iv} < D_{jv}\}| \tag{7}$$

$$N_{j,(i,j)} = |\{v | v \in V(\Gamma);\ (i,j) \in E(\Gamma);\ D_{jv} < D_{iv}\}| \tag{8}$$

where V(Γ) denotes the set of vertices in graph and D_{iv}, D_{jv} are the topological distances (*i.e.*, the number of edges joining vertices i and j on the shortest path (i,j)).

Thus, $N_{i,(i,j)}$ and $N_{j,(i,j)}$ represent the cardinality of the sets of vertices closer to i and to j, respectively; vertices equidistant to i and j are not counted. Edge-defined Szeged index, SZ_e is calculated by summing up all the edge contributions in graph

$$SZ_e = SZ_e(\Gamma) = \sum_{\text{all}(i,j)\in E(\Gamma)} N_{i(i,j)} N_{j(i,j)} \quad (9)$$

By analogy to the Wiener matrix, an edge-defined Szeged matrix,[21] $\boldsymbol{SZ}_e$, can be defined with the aid of edge contributions

$$[\boldsymbol{SZ}_e]_{i,j} = N_{i,(i,j)} N_{j,(i,j)} \; ; \quad (i,j)\in E(\Gamma) \quad (10)$$

As in the case of $\boldsymbol{W}_e$, for non-adjacent vertices i and j, $(i,j)\notin E(\Gamma)$, the non-diagonal entries in $\boldsymbol{SZ}_e$ are zero. The global index SZ_e can be calculated by

$$SZ_e = (1/2) \sum_i \sum_j [\boldsymbol{SZ}_e]_{i,j} \quad (11)$$

Similarly, a path-defined Szeged matrix is conceivable

$$[\boldsymbol{SZ}_p]_{i,j} = N_{i,(i,j)} N_{j,(i,j)} \; ; \quad (i,j)\in P(\Gamma) \quad (12)$$

and the index calculated on it is the *hyper-Szeged index*,[21] SZ_p

$$SZ_p = SZ_p(\Gamma) = (1/2) \sum_i \sum_j [\boldsymbol{SZ}_p]_{i,j} \quad (13)$$

$\boldsymbol{SZ}_p$ can be obtained by the Hadamard product[22] (*i.e.*, $[\boldsymbol{M}_a \bullet \boldsymbol{M}_b]_{ij} = [\boldsymbol{M}_a]_{ij} [\boldsymbol{M}_b]_{ij}$, where $\boldsymbol{M}$ is a square matrix) between $\boldsymbol{SZ}_p$ and $\boldsymbol{A}$ (the adjacency matrix, whose entries are 1 if two vertices are adjacent and zero otherwise)

$$\boldsymbol{SZ}_e = \boldsymbol{SZ}_p \bullet \boldsymbol{A} \quad (14)$$

A similar relation is valid for $\boldsymbol{W}_e$ and $\boldsymbol{W}_p$. Both Szeged matrices, $\boldsymbol{SZ}_e$ and $\boldsymbol{SZ}_p$ have the diagonal entries zero.

A Szeged unsymmetric matrix, $\boldsymbol{SZ}_u$, was also defined[21]

$$[\boldsymbol{SZ}_u]_{ij} = N_{i,(i,j)} \; ; \quad (i,j)\in P(\Gamma) \quad (15)$$

SZ_u is a square array of dimension $N \times N$, general by unsymmetric (N being the number of vertices in graph). The diagonal entries are zero. It allows the construction of symmetric matrices $\boldsymbol{SZ}_e$ and $\boldsymbol{SZ}_p$ by relation

$$[\boldsymbol{SZ}_{e/p}]_{ij} = [\boldsymbol{SZ}_u]_{ij} [\boldsymbol{SZ}_u]_{ji} \tag{16}$$

and the derivation of two Szeged indices, as

$$SZ_e = \sum_{\text{all}(i,j)\in E(G)} [\boldsymbol{SZ}_u]_{ij} [\boldsymbol{SZ}_u]_{ji} \tag{17}$$

$$SZ_p = \sum_{\text{all}(i,j)\in P(G)} [\boldsymbol{SZ}_u]_{ij} [\boldsymbol{SZ}_u]_{ji} \ . \tag{18}$$

In tree graphs, the index defined on edges is identical in both Wiener and Szeged matrices.Analytical relations for calculating the Szeged indices in paths, P_N, and simple cycles, C_N, are derived[23]

$$SZ_e (P_N) = N(N^2-1) / 6 \tag{19}$$

$$SZ_p (P_N) = (5N^4 - 10N^3 + 16N^2 - 8N - 6zN + 3z) / 48 \tag{20}$$

$$SZ_e (C_N) = N(N-z)^2 / 4 \tag{21}$$

$$SZ_p (C_N) = N(N-1)^{(2z+1)} (N^2-2N+4)^{(1-z)} / 8 \ ; \quad z = N \bmod 2 \ . \tag{22}$$

The present paper presents some extensions of the Szeged index, which account for fragments and their chemical nature as well as for their 3D-geometry.

BASIC SZEGED PROPERTY MATRICES

By analogy to $\boldsymbol{SZ}_u$ (see Eqs. (7), (8), (10), (12), (14–16)), Szeged property matrices are defined[24]

$$[\boldsymbol{SZ}_u\boldsymbol{P}]_{ij} = P_{i,(i,j)} \tag{23}$$

$$P_{i,(i,j)} = f(P_v)|v\in V(\Gamma) \ ; \quad D_{iv} < D_{jv}| \tag{24}$$

$$f(P_v) = m \sum_v P_v \tag{25}$$

$$f(P_v) = (\Pi_v P_v)^{1/N} \ . \tag{26}$$

The summation and product in Eqs. (25) and (26) run over all vertices in graph.

Entries in a Szeged property matrix (see Eq. (23)), in fact properties $P_{i,(i,j)}$ of vertex i (with respect to the edge/path *(i,j)*), are defined by a function $f(P_v)$, evaluated on vertices v, which obey the Szeged index condition (see Eq. 7). In other words, the set of such vertices can be viewed as a fragment (*i.e.*, a subgraph) since a molecular graph is always a connected one. Consequently, $P_{i,(i,j)}$ can be viewed as a fragmental property. $P_{i,(i,j)}$ is mainly a topological (local) property (*e.g.*, a topological index) but other physicochemical properties are also considered (e.g., atomic mass or group electronegativities – see below). Two types of $f(P_v)$ are proposed here: an additive and a multiplicative one.

Several cases of additive function (Eq. (25)) are considered:

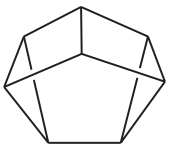
(a) $P_v = 1$ (*i.e.*, the cardinality) and the weighting factor $m = 1$ (classical $\boldsymbol{SZ}_u$ matrix).

(b) P_v = some vertex property; $m = 1$; (property matrix, $\boldsymbol{SZ}_u\boldsymbol{P}$).

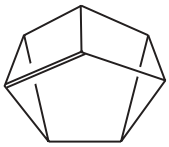
(c) P_v = some vertex property; $m = 1/P(\Gamma)$; $P(\Gamma)$ = a pertinent global property of Γ;
(*normalized* property matrix, $\boldsymbol{SZ}_u\boldsymbol{NP}$). Indices calculated on matrix $\boldsymbol{SZ}_u\boldsymbol{NP}$ (*cf.* Eqs. (17), (18) or (24)) are analogues of the *X'/X* indices of Randić[25] or *SP* indices of Diudea *et al.*[26]

(d) $P_v = \sum_v A_v$; $m = 1/12$; A_v is the atomic mass and the matrix is $\boldsymbol{SZ}_u\boldsymbol{A}$. Factor m indicates that $f(P_v)$ is a fragmental mass, relative to the carbon atomic mass.

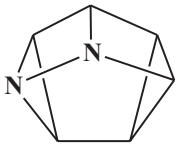
Matrices $\boldsymbol{SZ}_u\boldsymbol{P}$, with $P = {}^kW$; k=1–3 (*i.e.*, walk numbers[27] of elongation (1)–(3)), for the cuneane graphs Γ_1 and Γ_2, are illustrated in Chart 1. Matrices $\boldsymbol{SZ}_u\boldsymbol{NP}$, with $P = {}^3W$, are illustrated in Chart 2.



Γ_1



Γ_2



Γ_3

The multiplicative function (Eq. (26)) was used for group electronegativities: $P_v = X_v$; ($\boldsymbol{SZ}_u\boldsymbol{X}$ matrix).

X_v local electronegativity considered is the *EC* group electronegativity[28] (a Sanderson-type group electronegativity) for heteroatoms and fragments. It is defined by

$$EC_v = ES_v \,/\, (mlc_v EC_C) \tag{27}$$

$\boldsymbol{SZ}_u{}^1\boldsymbol{W}(\Gamma_1)$

0	9	9	12	12	6	9	9
12	0	12	9	9	12	6	12
9	9	0	15	12	9	12	6
9	3	9	0	6	9	6	9
9	6	9	9	0	9	6	9
6	9	9	12	12	0	9	9
9	6	12	9	9	9	0	12
6	9	6	15	12	9	12	0

SZ_pP: 2448
SZ_eP: 1152

$\boldsymbol{SZ}_u{}^1\boldsymbol{W}(\Gamma_2)$

0	10	10	13	14	6	9	9
13	0	14	9	10	13	6	12
9	9	0	15	13	9	12	6
10	3	11	0	7	10	6	9
9	6	10	9	0	9	6	9
6	10	10	13	14	0	9	9
11	7	14	10	11	11	0	13
7	11	8	17	14	10	13	0

SZ_pP: 2874
SZ_eP: 1365

$\boldsymbol{SZ}_u{}^2\boldsymbol{W}(\Gamma_1)$

0	27	27	36	36	18	27	27
36	0	36	27	27	36	18	36
27	27	0	45	36	27	36	18
27	9	27	0	18	27	18	27
27	18	27	27	0	27	18	27
18	27	27	36	36	0	27	27
27	18	36	27	27	27	0	36
18	27	18	45	36	27	36	0

SZ_pP: 22032
SZ_eP: 10368

$\boldsymbol{SZ}_u{}^2\boldsymbol{W}(\Gamma_2)$

0	34	34	44	48	20	29	29
43	0	48	30	34	43	19	38
28	28	0	48	43	29	38	18
33	10	38	0	24	33	19	28
28	19	33	29	0	28	18	28
19	33	34	43	48	0	28	29
38	24	48	34	38	38	0	43
24	38	28	58	48	34	43	0

SZ_pP: 31210
SZ_eP: 14856

$\boldsymbol{SZ}_u{}^3\boldsymbol{W}(\Gamma_1)$

0	81	81	108	108	54	81	81
108	0	108	81	81	108	54	108
81	81	0	135	108	81	108	54
81	27	81	0	54	81	54	81
81	54	81	81	0	81	54	81
54	81	81	108	108	0	8	81
81	54	108	81	81	81	0	108
54	81	54	135	108	81	108	0

SZ_pP: 198288
SZ_eP: 93312

$\boldsymbol{SZ}_u{}^3\boldsymbol{W}(\Gamma_2)$

0	115	115	148	163	67	96	96
142	0	163	100	115	142	62	125
90	90	0	158	142	94	123	58
109	32	128	0	80	109	61	90
90	62	110	96	0	90	58	91
62	110	115	144	163	0	91	96
128	81	163	115	130	128	0	144
80	128	96	196	161	113	142	0

SZ_pP: 343353
SZ_eP: 163692

Chart 1. Matrices $\boldsymbol{SZ_uP}$; $\boldsymbol{P} = {}^k\boldsymbol{W}$; k = 1–3, for Graphs Γ_1 and Γ_2.

$\boldsymbol{SZ_uNP}$ (Γ_1); $\boldsymbol{P} = {}^3\boldsymbol{W}$

0.0000	0.3750	0.3750	0.5000	0.5000	0.2500	0.3750	0.3750
0.5000	0.0000	0.5000	0.3750	0.3750	0.5000	0.2500	0.5000
0.3750	0.3750	0.0000	0.6250	0.5000	0.3750	0.5000	0.2500
0.3750	0.1250	0.3750	0.0000	0.2500	0.3750	0.2500	0.3750
0.3750	0.2500	0.3750	0.3750	0.0000	0.3750	0.2500	0.3750
0.2500	0.3750	0.3750	0.5000	0.5000	0.0000	0.3750	0.3750
0.3750	0.2500	0.5000	0.3750	0.3750	0.3750	0.0000	0.5000
0.2500	0.3750	0.2500	0.6250	0.5000	0.3750	0.5000	0.0000

SZ_pNP: 4.25000

SZ_eNP: 2.00000

$\boldsymbol{SZ_uNP}$ (Γ_2); $\boldsymbol{P} = {}^3\boldsymbol{W}$

0.0000	0.4021	0.4021	0.5175	0.5699	0.2343	0.3357	0.3357
0.4965	0.0000	0.5699	0.3497	0.4021	0.4965	0.2168	0.4371
0.3147	0.3147	0.0000	0.5524	0.4965	0.3287	0.4301	0.2028
0.3811	0.1119	0.4476	0.0000	0.2797	0.3811	0.2133	0.3147
0.3147	0.2168	0.3846	0.3357	0.0000	0.3147	0.2028	0.3182
0.2168	0.3846	0.4021	0.5035	0.5699	0.0000	0.3182	0.3357
0.4476	0.2832	0.5699	0.4021	0.4545	0.4476	0.0000	0.5035
0.2797	0.4476	0.3357	0.6853	0.5629	0.3951	0.4965	0.0000

SZ_pNP: 4.19767

SZ_eNP: 2.00122

Chart 2. Matrices $\boldsymbol{SZ_uNP}$; $\boldsymbol{P} = {}^3\boldsymbol{W}$, for Graphs Γ_1 and Γ_2.

where ES_v is the Sanderson group electronegativity[29] of vertex v, mlc_v is the mean length of covalent bond (relative to the tetraconnected carbon atom) around the vertex v, and $EC_C = 2.746/1.4996$. For the tetraconnected carbon atom, $EC_v = 1$. Appendix contains EC_v values for some frequently used groups.

Matrices $\boldsymbol{SZ_uX}$ for graphs Γ_1–Γ_3 , are illustrated in Chart 3.

Topological indices on these matrices are calculated by the general relation

$$TI_{e/p} = \sum_{e/p} [\boldsymbol{SZ_uP}]_{ij} [\boldsymbol{SZ_uP}]_{ji};$$
$$TI = SZ; SZP; SZNP; SZA; SZX . \quad (28)$$

When summation goes over all edges in Γ, e = all $(i,j) \in E(\Gamma)$, and TI_e is an edge-defined index (*e.g.*, SZ_e index). When it goes over all paths in Γ, p =

$\boldsymbol{SZ_uX}(\Gamma_1)$							
0.000	0.972	0.972	0.972	0.972	0.972	0.972	0.968
0.969	0.000	0.972	0.972	0.972	0.969	0.972	0.969
0.965	0.965	0.000	0.970	0.969	0.968	0.967	0.967
0.968	0.962	0.968	0.000	0.967	0.968	0.967	0.965
0.965	0.967	0.968	0.968	0.000	0.965	0.967	0.968
0.967	0.968	0.972	0.969	0.972	0.000	0.968	0.968
0.968	0.972	0.972	0.972	0.972	0.968	0.000	0.969
0.967	0.968	0.972	0.972	0.969	0.968	0.969	0.000
SZ_pX: 26.290							
SZ_eX: 10.330							
$\boldsymbol{SZ_uX}(\Gamma_2)$							
0.000	1.005	1.005	0.996	1.022	0.972	0.972	0.968
0.994	0.000	1.022	0.972	1.005	0.994	0.972	0.969
0.965	0.965	0.000	0.970	0.994	0.968	0.967	0.967
1.002	0.962	1.036	0.000	1.017	1.002	0.967	0.965
0.965	0.967	1.002	0.968	0.000	0.965	0.967	0.968
0.967	1.002	1.005	0.994	1.022	0.000	0.968	0.968
1.036	1.022	1.022	1.005	1.039	1.036	0.000	0.994
1.017	1.036	1.075	1.012	1.019	1.002	0.994	0.000
$SZpX$: 27.637							
SZ_eX: 10.889							
$\boldsymbol{SZ_uX}(\Gamma_3)$							
0.000	1.049	1.049	1.029	1.091	0.972	0.972	0.968
1.027	0.000	1.091	0.972	1.049	1.027	0.972	0.969
0.965	0.965	0.000	0.970	1.027	0.968	0.967	0.967
1.046	0.962	1.130	0.000	1.085	1.046	0.967	0.965
0.965	0.967	1.046	0.968	0.000	0.965	0.967	0.968
0.967	1.046	1.049	1.027	1.091	0.000	0.968	0.968
1.130	1.091	1.091	1.049	1.133	1.130	0.000	1.027
1.085	1.130	1.224	1.066	1.088	1.046	1.027	0.000
SZ_pX: 29.497							
SZ_eX: 11.663							

Chart 3. Matrices $\boldsymbol{SZ_uX}$ for Graphs Γ_1–Γ_3.

all $(i,j)\in P(\Gamma)$, and TI_p is a path-defined index (or a hyper-index, *e.g.*, SZ_pA index). Symbol P is taken *ad-hoc*. Values of indices *SZP, SZNP* and *SZX for cuneanes* Γ_1–Γ_3 are included in Charts 1–3.

DISTANCE EXTENDED SZEGED PROPERTY MATRICES

Tratch *et al.*[30] proposed an *extended distance matrix*, $\boldsymbol{E}$, whose entries are the product of entries in the distance matrix $\boldsymbol{D}$ (*i.e.*, the matrix with the non-diagonal entries equal to the topological distances between vertices i and j) and a multiplyer, m_{ij}, which is the number of paths in graph having the path (i,j) as a subpath. In acyclic structures, m_{ij} equals the entries in the $\boldsymbol{W}_\mathrm{p}$ matrix, so that $\boldsymbol{E}$ is further referred to as $\boldsymbol{D\_W}_\mathrm{p}$ matrix

$$[\boldsymbol{D\_W}_\mathrm{p}]_{ij} = [\boldsymbol{D}]_{ij}\, m_{ij} = [\boldsymbol{D}]_{ij}\, [\boldsymbol{W}_\mathrm{p}]_{ij} = \boldsymbol{D}\bullet\boldsymbol{W}_\mathrm{p} \ . \tag{29}$$

Thus, $\boldsymbol{D\_W}_\mathrm{p}$ matrix results as the Hadamard product, $\boldsymbol{D}\bullet\boldsymbol{W}_\mathrm{p}$. It is a square symmetric matrix of dimensions $N \times N$, having the diagonal entries zero. The half sum of its entries gives an expanded Wiener number.

In full analogy, Diudea *et al.*[31] proposed a distance-extended Szeged unsymmetric matrix

$$\boldsymbol{D\_SZ}_\mathrm{u} = \boldsymbol{D}\bullet\boldsymbol{SZ}_\mathrm{u} \tag{30}$$

which offered a new distance-extended index

$$D^2\_SZ_\mathrm{p} = \sum_\mathrm{p}[\boldsymbol{D\_SZ}_\mathrm{u}]_{ij}\, [\boldsymbol{D\_SZ}_\mathrm{u}]_{ji} = (1/2)\, \sum_i \sum_j [(\boldsymbol{D\_SZ}_\mathrm{u})(\boldsymbol{D\_SZ}_\mathrm{u})^T)]_{ij} \ . \tag{31}$$

where summation goes over all paths in Γ: p = all $(i,j) \in \mathrm{P}(\Gamma)$. Note that index $D^2\_SZ_\mathrm{p}$ involves squared distances (see the superscript number), which are used for calculating the moment of inertia of molecules.[32]

When the Hadamard multiplication is performed using the reciprocal distance matrix,[33,34] $\boldsymbol{RD}$, (*i.e.*, the matrix having entries $1/[\boldsymbol{D}]_{ij}$): $\boldsymbol{RD}\bullet\boldsymbol{M} = \boldsymbol{RD\_M}$, it results in new (reciprocal) distance-extended indices, or Harary-type indices[23]

$$H^2\_SZ_\mathrm{p} = \sum_\mathrm{p}[\boldsymbol{RD\_SZ}_\mathrm{p}]_{ij}\, [\boldsymbol{RD\_SZ}_\mathrm{p}]_{ji} \ . \tag{32}$$

Values $H^2\_SZ_\mathrm{p}$ for octanes and the correlating ability of indices $D^2\_SZ_\mathrm{p}$ and $H^2\_SZ_\mathrm{p}$ have been presented elswere.[31]

In the present paper, matrix $\boldsymbol{SZ}_\mathrm{u}$ is changed by $\boldsymbol{SZ}_\mathrm{u}\boldsymbol{P}$ in deriving the corresponding matrices and indices

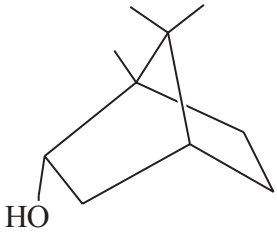
$$\boldsymbol{M\_SZ}_\mathrm{u}\boldsymbol{P} = \boldsymbol{M}\bullet\boldsymbol{SZ}_\mathrm{u}\boldsymbol{P} \ ; \quad \boldsymbol{M} = \boldsymbol{D}; \boldsymbol{RD} \tag{33}$$

$$M^2\_SZ_\mathrm{p}P = \sum_\mathrm{p}[\boldsymbol{M\_SZ}_\mathrm{u}\boldsymbol{P}]_{ij}\, [\boldsymbol{M\_SZ}_\mathrm{u}\boldsymbol{P}]_{ji} \ ; \quad M = D \ ; \quad H \ . \tag{34}$$

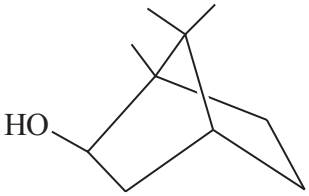
Edge defined indices are similarly calculated. Values of the above indices are listed in Ref. 31.

An interesting index is $H^2\_SZ_pA$, which is a Gravitational index analogue.[35]

When matrix ***D*** is changed by a 3D-***D*** matrix (*i.e.*, a matrix whose entries are genuine distances), which is called the geometric matrix,[36] ***G***, the index $G^2\_SZ_pP$ can distinguish among stereoisomers The corresponding Harary index is denoted by $RG^2\_SZ_pP$. Thus, the two isomers of borneol, Γ_4 and Γ_5, are easily discriminated (see Table I). The basic Szeged indices for these graphs are given in the footnote to Table I.

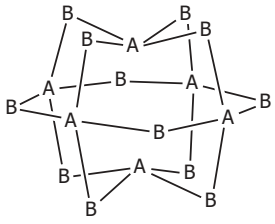


Γ_4

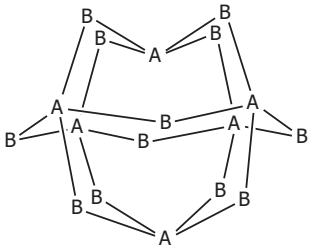


Γ_5

Conformational isomers may also be separated. A homeomorphic transform of the square dipyramid[37] shows two main conformers, $\Gamma_{6,a}$ and $\Gamma_{6,b}$. Indices are calculated for A = C; Si; Ge; Sn and B = CH2, NH, O, S; values are listed in Table II. The basic Szeged indices for this graph are given in the footnote to Table II.



$\Gamma_{6,a}$



$\Gamma_{6,b}$

The geometric matrix, ***G***, was calculated using the SPARTAN PC program and Szeged matrices and indices by our program, SZEGED 2.0, written in DELPHY.

CORRELATING TEST

In a QSAR study,[38] a set of 18 aromatic structures,[39] containing nitrogen both in a heterocycle and in a side chain (Table III) and having inhibition

TABLE I

Szeged – Type Indices for Stereoisomeric Graphs Γ_4 and Γ_5.

Topological Index	Borneol Γ_4	Isoborneol Γ_5
$D^2\_SZ_p$	4477.00	4477.00
$G^2\_SZ_p$	5963.73	5888.15
$H^2\_SZ_p$	261.37	261.37
$RG^2\_SZ_p$	139.13	139.18
$D^2\_SZ_e$	174.00	174.00
$G^2\_SZ_e$	411.02	411.17
$H^2\_SZ_e$	174.00	174.00
$RG^2\_SZ_e$	73.81	73.78
$D^2\_SZ_pA$	6205.90	6205.90
$G^2\_SZ_pA$	8266.95	8160.54
$H^2\_SZ_pA$	364.59	364.59
$RG^2\_SZ_pA$	194.02	194.10
$D^2\_SZ_eA$	243.03	243.03
$G^2\_SZ_eA$	573.12	573.38
$H^2\_SZ_eA$	243.03	243.03
$RG^2\_SZ_eA$	103.30	103.26
$D^2\_SZ_pX$	313.99	313.99
$G^2\_SZ_pX$	428.78	422.03
$H^2\_SZ_pX$	19.40	19.40
$RG^2\_SZ_pX$	10.43	10.45
$D^2\_SZ_eX$	11.92	11.92
$G^2\_SZ_eX$	27.90	27.91
$H^2\_SZ_eX$	11.92	11.92
$RG^2\_SZ_eX$	5.11	5.10

SZ_p : 730; SZ_e : 174; SZ_pX : 54.625; SZ_eX : 11.916; SZ_pA : 1014.35; SZ_eA : 243.03

activity on the growth of tetrahymena, was modeled by the Szeged fragmental indices for some nitrogen containing compounds. The statistics of the regression equations are given in Table IV. For comparison, correlations given by the *EATI* index[39] are included. The biological activity is taken as log(IGC50). IGC50 is the concentration – in mmol/L – which inhibits 50% of the growth of tetrahymena cultures.

TABLE II

Szeged – Type Indices for the Conformers of Graph Γ_6.

A; B (G)	$RG^2\_SZ_p$	$RG^2\_SZ_e$	SZ_pA	SZ_eA	$H^2\_SZ_pA$	$RG^2\_SZ_pA$	$RG^2\_SZ_eA$
C; CH_2 (6a)	1292.5	656.7	9642.2	2133.3	3064.7	1599.3	809.5
C; NH (6a)	1545.0	774.7	10645.1	2352.0	3381.5	2110.0	1054.5
C; O (6b)	1685.5	852.6	11697.7	2581.3	3713.9	2529.3	1273.6
C; S (6b)	1694.3	857.3	35281.7	7701.3	11153.7	7651.6	3820.9
Si; NH (6b)	990.9	509.7	19911.8	4485.3	6378.8	2545.0	1323.0
Si; O (6b)	1161.0	602.4	21343.0	4800.0	6832.3	3193.0	1673.3
Si; S (6b)	746.7	376.5	50985.7	11285.4	16208.0	4886.3	2458.7
Ge; O (6b)	945.3	486.0	63256.9	14560.3	20478.7	7767.3	4095.1
Sn; O (6b)	784.1	402.4	130768.6	30401.2	42555.5	13365.8	7080.1

SZ_p : 7785.0; SZ_e : 1728.0; $H^2\_SZ_p$: 2477.7; $H^2\_SZ_e$: 1728.0.

TABLE III

Szeged Fragmental Indices and log(IGC50) (Inhibition of the Growth of Tetrahymena Cultures) of Some Nitrogen Containing Compounds

No.	Graph	SZ_p	SZ_e	SZ_pA	SZ_eA	SZ_pX	SZ_eX	log(IGC50)
1	pyridine	105	54	126.389	65.000	16.859	6.740	1.1853
2	3-picoline	182	78	224.701	96.264	23.120	7.702	1.0175
3	4-picoline	182	78	225.292	96.458	23.261	7.745	0.8921
4	3,4-lutidine	296	106	371.479	133.361	30.553	8.696	0.5051
5	quinoline	783	243	910.993	282.361	49.680	12.165	–0.0132
6	4-phenylpyridine	1684	360	1970.458	421.875	72.835	14.396	–0.6576
7	acridine	3149	656	3586.694	748.333	99.422	17.522	–1.3979
8	aniline	182	72	225.042	96.306	22.466	7.486	0.2201
9	3-toluidine	296	108	372.757	136.028	29.614	8.454	0.4133
10	4-toluidine	310	110	391.653	138.944	29.641	8.463	0.1271
11	3,4-xilidine	474	144	604.354	184.306	37.824	9.429	0.2878
12	1-naphthylamine	1107	300	1311.201	355.194	59.020	12.896	–0.2218
13	4-aminobiphenyl	2350	450	2791.792	356.167	83.451	14.994	–0.8239
14	nitrobenzene	182	78	381.917	161.097	35.286	11.573	0.0645
15	3-nitrotoluene	296	108	604.242	218.528	43.954	12.548	–0.3098
16	4-nitrotoluene	310	110	646.653	227.278	44.368	12.704	–0.2366
17	4-nitro-*o*-xilene	474	144	956.854	290.347	54.123	13.634	–0.6383
18	4-nitrobyphenyl	2350	450	3988.944	776.611	107.022	20.024	–1.0000

TABLE IV

Statistics of Regression Equations: $Y = a + \sum_i b_i X_i$, for the Set of Table III.

No.	TI	*b*	*a*	*r*	*s*	*F*
1	*EATI*	–0.084	1.823	0.9313	0.265	104.54
2	ln *EATI*	–1.911	5.762	0.9419	0.244	125.79
3	SZ_pX	–0.024	1.110	0.9087	0.304	75.83
4	ln SZ_pX	–1.236	4.580	0.9424	0.243	126.94
5	SZ_pA	–0.0005	0.508	0.8365	0.399	37.29
6	ln SZ_pA	–0.646	4.167	0.9300	0.267	102.46
7	$1/SZ_pX$	22.695	0.486	0.9440	0.248	61.44
	SZ_eX	–0.099				
8	ln SZ_pX	–0.775	4.776	0.9459	0.244	63.75
	ln SZ_eX	–0.801				
9	$1/SZ_eX$	17.223	–1.501	0.9475	0.240	65.88
	SZ_p	–0.0002				
10	ln SZ_pX	–1.644	4.725	0.9480	0.239	66.53
	ln SZ_p	0.223				
11	ln SZ_pA	–1.103	3.877	0.9463	0.243	64.31
	ln SZ_e	0.648				
12	ln SZ_pA	–1.289	4.345	0.9510	0.232	70.98
	ln SZ_p	0.606				

In single variable regression, (Table IV, entries 1–6), the best correlation was found for SZ_pX (as a natural logarithm – entry 4):

$$\log(\text{IGC50}) = 4.58 - 1.236 \ln SZ_pX$$

$N = 18$; $r = 0.9424$; $s = 0.243$; $F = 126.94$

This result surpasses the best correlation reported for the *EATI* index (entry 2). Note the beneficial action of logarithmation of values of the Szeged-type indices (compare entries 3 and 4, and also 5 and 6).

In the two variable regression, the standard error was slightly reduced (entries 9–12) in comparison with the best single variable regression. The best equation was

$$\log(\text{IGC50}) = 4.345 - 1.289 \ln \sum Z_\pi A + 0.606 \ln SZ_p$$

$N = 18$; $r = 0.9510$; $s = 0.232$; $F = 70.98$

The modeling is not satisfactory and not suitable for prediction. However, it demonstrates the dependency of the biological response on the molecular structure.

In a QSPR test,[40] two physicochemical properties: diffusion coefficient in water, *DC* water, and in air, *DC* air, of a set of 15 explosives[41] (Table V) have been modeled. Statistics of the single variable regression: $Y = a + bX$ for the compounds of Table V and the cross validation test (»Leave one out«), *cv*, are included in Table VI.

Table VI indicates a good estimative and predictive ability of regression equations. It appears that fragmental descriptors are suitable for correlating studies.

TABLE V

Szeged Fragmental Indices and Diffusion Coefficients, *DC*, for Some Explosives.

Graph*	SZ_e	SZ_p	SZ_eA	SZ_pA	SZ_eX	SZ_pX	DC water $\times 10^6$ (cm^2/s)	*DC* air (cm^2/s)
1	594	4348	601.049	4169.958	74.949	510.786	6.71	0.0639
2	360	2050	378.451	2049.049	57.524	333.915	7.31	0.0670
3	348	1993	366.007	1978.160	49.430	290.101	7.31	0.0670
4	296	1542	303.292	1500.896	49.076	268.231	7.94	0.0729
5	516	3450	509.917	3239.271	66.837	431.457	7.20	0.0679
6	516	3450	555.083	3528.729	68.038	439.675	7.15	0.0739
7	1156	11794	1228.833	12083.000	100.785	824.478	6.02	0.0629
8	1014	10342	1022.236	9969.632	99.907	820.577	5.99	0.0590
9	594	4348	613.896	4248.014	78.289	528.620	7.03	0.0660
10	968	11514	1083.000	12479.986	94.923	832.585	5.61	0.0570
11	424	3677	475.729	3961.931	68.805	455.598	6.95	0.0700
12	48	159	61.389	194.514	36.330	133.953	10.40	0.1019
13	151	827	172.771	904.382	43.192	206.946	8.72	0.0839
14	344	2518	408.569	2913.472	45.670	270.017	7.05	0.0689
15	184	1153	214.278	1274.937	50.820	262.697	7.93	0.0769

* 1. Trinitrotoluene; 2. 2,4-Dinitrotoluene; 3. 2,6-Dinitrotoluene; 4. 1,3-Dinitrobenzene; 5. 1,3,5-Trinitrobenzene; 6. Hexahydro-1,3,5-trinitro-1,3,5-triazine; 7. Octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine; 8. *N*-2,4,6-Tetranitro-*N*-methylaniline; 9. Picric acid; 10. Pentaerythrytol tetranitrate; 11. Nitroglycerin; 12. Nitroguanidine; 13. Ethylene glycol dinitrate; 14. Diethylene glycol dinitrate; 15. Propylene glycol dinitrate

TABLE VI

Statistics of the Single Variable Regression: $Y = a + bX$ for the Set of Table V and the Cross Validation Test (Leave One Out)

No	Y	X	a	b	r	s	F	r_{cv}	s_{cv}
1	CD_{air}	ln SZ_e			0.943	0.0037	105.09		
2		1/ln SZ_eA	0.00073	0.4157	0.960	0.0031	155.21	0.951	0.0035
3		1/ln SZ_pA			0.954	0.0034	133.61		
4	CD_{water}	ln SZ_e	15.60561	–1.3934	0.973	0.2801	231.72	0.960	0.338
5		ln SZ_eA	16.31334	–1.4915	0.979	0.2443	310.57	0.968	0.302
6		ln SZ_pA	15.82862	–1.0766	0.984	0.2123	413.59	0.975	0.266
7		$1/SZ_pX$			0.961	0.3321	160.95		

CONCLUSION

Extension of the basic Szeged indices to indices which take into account fragments of a graph, with their chemical and geometrical characteristics, makes it possible to discriminate between various classes of isomeric chemical structures. Correlations of some of the newly proposed indices with molecular properties have also been found.

Acknowledgements . – This work was supported in part by the GRANT of the National Education Ministry, No. 74/339/1997 and in part by the GRANT of the Romanian Academy of Sciences, No. 2627/1997.

REFERENCES

1. H. Wiener, *J. Am. Chem. Soc*. **69** (1947) 17–20.
2. I. Gutman, Y. N. Yeh, S. L. Lee, and Y. L. Luo, *Indian J. Chem*. **32A** (1993) 651–661.
3. S. Nikolić, N. Trinajstić, and Z. Mihalić, *Croat. Chem. Acta.* **68** (1995) 105–129.
4. M. V. Diudea and I. Gutman, *Croat. Chem. Acta*, (in press).
5. M. Randić, X. Guo, T. Oxley, and H. Krishnapriyan, *J. Chem. Inf. Comput. Sci.* **33** (1993) 700–716.
6. M. Randić, X. Guo, T. Oxley, H. Krishnapriyan, and L. Naylor, *J. Chem. Inf. Comput. Sci.* **34** (1994) 361–367.
7. M. Randić, *Chem. Phys. Lett*. **211** (1993) 478–483.
8. H. Hosoya, *Bull. Chem. Soc. Jpn.* **44** (1971) 2332–2339.
9. M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **36** (1996) 535–540.
10. M. V. Diudea, *Commun. Math. Comput. Chem*. (*MATCH*) **35** (1997) 169–183.
11. M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **37** (1997) 300–305.
12. I. Gutman and M. V. Diudea, *J. Serb. Chem. Soc*. (submitted).
13. M. V. Diudea, B. Parv, and Gutman, I. *J. Chem. Inf. Comput. Sci.* **37** (1997) 1101–1108.
14. I. Gutman, *Graph Theory Notes New York* **27** (1994) 9–15.
15. A. A. Dobrynin and I. Gutman, *Publ. Inst. Math.* (Beograd) **56** (1994) 18–22.

16. A. A. Dobrynin, I. Gutman, and G. Dömötör, *Appl. Math. Lett*. **8** (1995) 57–62.
17. A. A. Dobrynin and I. Gutman, *Graph Theory Notes New York*, **28** (1995) 21–23.
18. P. V. Khadikar, N. V. Deshpande, P. P. Kale, A. A. Dobrynin, I. Gutman, and G. Dömötör, *J. Chem. Inf. Comput. Sci.* **35** (1995) 547–550.
19. I. Gutman and S. Klavžar, *J. Chem. Inf. Comput. Sci.* **35** (1995) 1011–1014.
20. A. A. Dobrynin and I. Gutman, *Croat. Chem. Acta* **69** (1996) 845–856.
21. M. V. Diudea, O. M. Minailiuc, G. Katona, and I. Gutman, *Commun. Math. Comput. Chem. (MATCH*) **35** (1997) 129–143.
22. R. A. Horn, C. R. Johnson, *Matrix Analysis,* Cambridge Univ. Press, Cambridge, 1985.
23. M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **37** (1997) 292–299.
24. Szeged fragmental indices were first put forward in a lecture given by M. V. D. at the Rugjer Bošković Institute, Zagreb, Croatia, on July 7, 1996.
25. M. Randić, *J. Math. Chem.* **7** (1991) 155–168.
26. M. V. Diudea, O. M. Minailiuc, and G. Katona, *Croat. Chem. Acta* **69** (1996) 857–871.
27. M. V. Diudea, M. I. Topan, and A. Graovac, *J. Chem. Inf. Comput. Sci.* **34** (1994) 1072–1078.
28. M. V. Diudea, I. E. Kacso, and M. I. Topan, *Rev. Roum. Chim.* **41** (1996) 141–157.
29. R. T. Sanderson, *Polar Covalence*, Acad. Press, New York, 1983.
30. S. S. Tratch, M. I. Stankevitch, and N. S. Zefirov, *J. Comput. Chem*. **11** (1990) 899–908.
31. M. V. Diudea, B. Parv, and M. I. Topan, *J. Serb. Chem. Soc.* **62** (1997) 267–276.
32. M. Randić and M. Razinger, *J. Chem. Inf. Comput. Sci.* **35** (1995) 140–147.
33. O. Ivanciuc, T. S. Balaban, and A. T. Balaban, *J. Math. Chem.* **12** (1993) 309–318.
34. Z. Mihalić, S. Nikolić, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **32** (1992) 28–37.
35. A. R. Katritzky, L. Mu, and V. S. Lobanov, *J. Phys. Chem*. (in press).
36. G. M. Crippen, *J. Comput. Phys*. 24 (1977) 96–107.
37. The structure was presented in a lecture given by M. V. D. at the Faculty of Science, University of Kragujevac, Yugoslavia, Feb. 5, 1997.
38. I. Schirger and M. V. Diudea, *Studia Univ. »Babes-Bolyai«* **42** (1997) 000–000.
39. M. Guo, L. Xu, C. Y. Hu, and S. M. Yu, *Commun. Math. Comput. Chem. (MATCH)* **35** (1997) 185–197.
40. A. A. Kiss, I. E. Kacso, O. M. Minailiuc, M. V. Diudea, S. Nikolić, and I. Gutman, *Studia Univ. »Babes-Bolyai«* **42** (1997) 135–142.
41. S. Nikolić, M. Medić-Šarić, S. Rendić, and N. Trinajstić, *Drug Metabolism Reviews* **26** (1994) 717.

SAŽETAK

Szegedski fragmentalni indeksi

Ovidiu M. Minailiuc, Gabriel Katona, Mircea V. Diudea
Mate Strunje, Ante Graovac i Ivan Gutman

Predloženi su novi szegedski indeksi, definirani na nesimetričnim matricama, koji obuhvaćaju razna fragmentalna topološka svojstva molekula. Indeksi su ispitani na raznim klasama grafova što opisuju molekule s heteroatomima i višestrukim vezama te stereoizomere. Korisnost novouvedenih indeksa potvrđena je njihovom korelacijom s herbicidalnom učinkovitošću i eksplozivnim svojstvima niza organskih molekula.