CROATICA CHEMICA ACTA CCACAA **71** (3) 459–471 (1998)

ISSN-0011-1643
CCA-2505 *Original Scientific Paper*

Detour and Cluj-Detour Indices*

Mircea V. Diudea,[a,**] *Gabriel Katona,*[a] *István Lukovits,*[b] *and Nenad Trinajstić*[c]

[a]*Department of Chemistry, Babes-Bolyai University, 11 Arany Janos Str., R-3400 Cluj, Romania*
[b]*Central Research Institute for Chemistry, Hungarian Academy of Sciences, H-1525 Budapest, P. O. Box 17, Hungary*
[c]*Rugjer Bošković Institute, P.O.B. 1016, HR-10001 Zagreb, Croatia*

Received January 28, 1998; revised April 6, 1998; accepted May 22, 1998

Recently proposed Cluj matrices: the Cluj-distance matrix and the Cluj-detour matrix are reviewed. New Harary-type indices on the detour and Cluj-detour matrices are defined. Additionally, the formulae for calculating these indices of cycles are derived. Modeling of boiling points for a set of 32 acyclic and cyclic octanes using Cluj indices and their Harary counterparts is presented. The best structure-boiling point relationships are obtained by means of the multiple linear regression using either combinations of two reciprocal paths numbers ($1/p_2$, $1/p_3$) and detour and hyper-detour indices or a combination of the same two reciprocal path numbers and Harary indices derived from the edge-defined and path-defined Cluj matrices.

INTRODUCTION

In an undirected connected acyclic graph, a given pair of vertices (i,j) is joined by a unique path $p(i,j)$, that is, a continuous sequence of edges, with the property that all are distinct and any two subsequent edges are adja-

* Reported in part at MATH/CHEM/COMP 1997, an International Course and Conference on the Interface between Mathematics, Chemistry and Computer Science, Dubrovnik, Croatia, June 23–28, 1997.

**Author to whom correspondence should be addressed.
(e-mail address: diudea@chem.ubbcluj.ro)

cent.[1,2] The length of the path p(i,j) is equal to the number of edges in the path between vertices i and j.

In an undirected connected cycle-containing graph between any two vertices, there is at least one path connecting them. If more than one path connects a given pair of vertices (i,j), we denote the k-th path by the symbol $p_k(i,j)$. The shortest path joining vertices i and j is called *geodesic* and its length is the topological distance, $(D)_{i,j}$. The longest path is the *elongation* and its length is equal to the detour distance, $(\boldsymbol{\Delta})_{ij}$. The square arrays which collect the lengths of the two path types are called the *distance matrix,*[1–3] denoted as $\boldsymbol{D}$, and the *detour matrix,*[3–6] denoted as $\boldsymbol{\Delta}$, respectively:

$$(\boldsymbol{D}_{\mathrm{e}})_{ij} = \begin{cases} N_{\mathrm{e,p}(i,j)}\text{: p}(i,j) \text{ is a geodesic if } i \neq j \\ 0 \quad \text{if } i = j \end{cases} \tag{1}$$

$$(\Delta_{\mathrm{e}})_{ij} = \begin{cases} N_{\mathrm{e,p}(i,j)}\text{: p}(i,j) \text{ is an elongation if } i \neq j \\ 0 \quad \text{if } i = j \end{cases} \tag{2}$$

where $N_{\mathrm{e,p}(i,j)}$ is the number of edges on the shortest/longest path p(i,j). The subscript e in the symbols of the above matrices means that they are edge-defined, that is, their entries count edges on the path p(i,j). Matrices $\boldsymbol{D}_{\mathrm{e}}$ and Δ_{e} for graph G_1, corresponding for example to pinane (see Figure 1), are given in Table I.

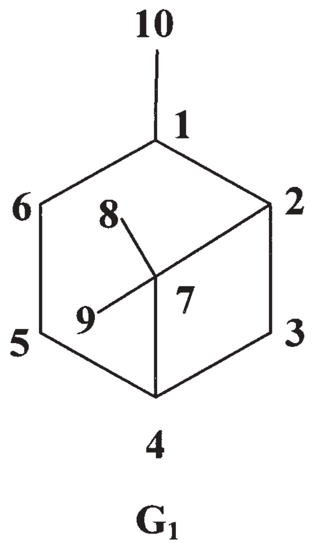

Figure 1. A graph G_1 corresponding to pinane

TABLE I

Distance, detour and Cluj matrices for graph G_1 and the related indices

$\boldsymbol{D}_e$

	1	2	3	4	5	6	7	8	9	10
1	0	1	2	3	2	1	2	3	3	1
2	1	0	1	2	3	2	1	2	2	2
3	2	1	0	1	2	3	2	3	3	3
4	3	2	1	0	1	2	1	2	2	4
5	2	3	2	1	0	1	2	3	3	3
6	1	2	3	2	1	0	3	4	4	2
7	2	1	2	1	2	3	0	1	1	3
8	3	2	3	2	3	4	1	0	2	4
9	3	2	3	2	3	4	1	2	0	4
10	1	2	3	4	3	2	3	4	4	0
	18	16	20	18	20	22	16	24	24	26

$W = 102$

$\boldsymbol{\Delta}_e$

	1	2	3	4	5	6	7	8	9	10
1	0	5	6	3	4	5	6	7	7	1
2	5	0	5	4	3	4	5	6	6	6
3	6	5	0	5	6	5	6	7	7	7
4	3	4	5	0	5	4	5	6	6	4
5	4	3	6	5	0	5	6	7	7	5
6	5	4	5	4	5	0	5	6	6	6
7	6	5	6	5	6	5	0	1	1	7
8	7	6	7	6	7	6	1	0	2	8
9	7	6	7	6	7	6	1	2	0	8
10	1	6	7	4	5	6	7	8	8	0
	44	44	54	42	48	46	42	50	50	52

$\omega = 236$

$\boldsymbol{D}_p$

	1	2	3	4	5	6	7	8	9	10
1	0	1	3	6	3	1	3	6	6	1
2	1	0	1	3	6	3	1	3	3	3
3	3	1	0	1	3	6	3	6	6	6
4	6	3	1	0	1	3	1	3	3	10
5	3	6	3	1	0	1	3	6	6	6
6	1	3	6	3	1	0	6	10	10	3
7	3	1	3	1	3	6	0	1	1	6
8	6	3	6	3	6	10	1	0	3	10
9	6	3	6	3	6	10	1	3	0	10
10	1	3	6	10	6	3	6	10	10	0
	30	24	35	31	35	43	25	48	48	55

$WW = 187$

$\boldsymbol{\Delta}_p$

	1	2	3	4	5	6	7	8	9	10
1	0	15	21	6	10	15	21	28	28	1
2	15	0	15	10	6	10	15	21	21	21
3	21	15	0	15	21	15	21	28	28	28
4	6	10	15	0	15	10	15	21	21	10
5	10	6	21	15	0	15	21	28	28	15
6	15	10	15	10	15	0	15	21	21	21
7	21	15	21	15	21	15	0	1	1	28
8	28	21	28	21	28	21	1	0	3	36
9	28	21	28	21	28	21	1	3	0	36
10	1	21	28	10	15	21	28	36	36	0
	145	134	192	123	159	143	138	187	187	196

$\omega\omega = 802$

$\boldsymbol{CJD}_u$

	1	2	3	4	5	6	7	8	9	10
1	0	4	3	3	3	7	3	5	5	9
2	6	0	7	3	6	5	5	5	5	6
3	2	3	0	4	2	5	1	6	6	6
4	5	2	6	0	6	5	4	4	4	5
5	2	2	2	4	0	6	2	5	5	6
6	3	2	3	3	4	0	3	3	3	3
7	4	5	3	6	4	5	0	9	9	6
8	1	1	1	1	1	1	1	0	1	1
9	1	1	1	1	1	1	1	1	0	1
10	1	1	1	1	1	1	1	1	1	0
	25	21	27	26	28	36	21	39	39	43

$CD_p = 432$

$CD_e = 214$

$\boldsymbol{CJ\Delta}_u$

	1	2	3	4	5	6	7	8	9	10
1	0	2	2	3	2	2	2	2	2	9
2	4	0	4	1	6	4	2	2	2	4
3	1	1	0	1	1	1	1	1	1	1
4	5	1	4	0	4	4	2	2	2	5
5	1	2	1	1	0	1	1	1	1	2
6	1	1	3	1	1	0	3	3	3	1
7	3	3	3	3	3	3	0	9	9	3
8	1	1	1	1	1	1	1	0	1	1
9	1	1	1	1	1	1	1	1	0	1
10	1	1	1	1	1	1	1	1	1	0
	18	13	20	13	20	18	14	22	22	27

$C\Delta_p = 168$

$C\Delta_e = 62$

When paths of length $1 \le |\mathrm{p}| \le |\mathrm{p}(i,j)|$ are counted on path $\mathrm{p}(i,j)$, another pair of matrices can be constructed

$$(\boldsymbol{D}_{\mathrm{p}})_{ij} = \begin{cases} N_{\mathrm{p},\mathrm{p}(i,j)}: \mathrm{p}(i,j) \text{ is a geodesic if } i \neq j \\ 0 \quad \text{if } i = j \end{cases} \tag{3}$$

$$(\Delta_{\mathrm{p}})_{ij} = \begin{cases} N_{\mathrm{p},\mathrm{p}(i,j)}: \mathrm{p}(i,j) \text{ is an elongation if } i \neq j \\ 0 \quad \text{if } i = j \; . \end{cases} \tag{4}$$

They are path-defined matrices and the number of paths $N_{\mathrm{p},\mathrm{p}(i,j)}$ is obtained from entries $(\boldsymbol{M}_{\mathrm{e}})_{ij}$, $\boldsymbol{M}_{\mathrm{e}} = \boldsymbol{D}_{\mathrm{e}}$ or Δ_{e}, by:[7,8]

$$N_{\mathrm{p},\mathrm{p}(i,j)} = \{[(\boldsymbol{M}_{\mathrm{e}})_{ij}]^2 + (\boldsymbol{M}_{\mathrm{e}})_{ij}\}/2 \; . \tag{5}$$

Matrices $\boldsymbol{D}_{\mathrm{p}}$ and Δ_{p} for the pinane graph G_1 are also given in Table I.

Several graph descriptors (topological indices) *TI* can be calculated as the half-sum of entries in the above matrices:

$$TI_{\mathrm{e/p}} = (1/2) \sum_i \sum_j (\boldsymbol{M}_{\mathrm{e/p}})_{ij} \tag{6}$$

where (the edge-defined index) TI_{e} represents the Wiener index[9] W and the detour index[4–6,10–13] ω, while (the path-defined index) TI_{p} is the hyper-Wiener index[14–16] WW and the hyper-detour index[8,10] $\omega\omega$, respectively. Values of these indices for the pinane graph G_1 are given in Table I.

The detour and hyper-detour indices have been recently introduced and tested in structure-property modeling.[6,10,11] The obtained results encouraged us to continue the investigation along the same line using the recently proposed Cluj matrices.[17,18] In the present paper, Harary-type indices[19–22] will be derived from the detour and Cluj-detour matrices. Modeling of the boiling points for a set of 32 acyclic and cyclic octanes using these indices will be reported.

DEFINITION OF CLUJ MATRICES

Cluj matrices $\boldsymbol{CJD}_{\mathrm{u}}$ and $\boldsymbol{CJ\Delta}_{\mathrm{u}}$ have been recently proposed by Diudea.[17,18,23–25] These matrices are $n \times n$ square matrices, which are unsymmetrical. Note that subscript u denotes the unsymmetricity of matrices. The non-diagonal entries, $(\boldsymbol{M}_{\mathrm{u}})_{ij}$, $\boldsymbol{M}_{\mathrm{u}} = \boldsymbol{CJD}_{\mathrm{u}}$ or $\boldsymbol{CJ\Delta}_{\mathrm{u}}$, in the two Cluj matrices are defined as:

$$(\boldsymbol{M}_{\mathrm{u}})_{ij} = N_{i,\,\mathrm{p}_k(i,j)} = \max|V_{i,\,\mathrm{p}_k(i,j)}| \tag{7}$$

$$V_{i,p_k(i,j)} = \{v|\ v \in V(\mathrm{G});\ \boldsymbol{D}_{iv} < \boldsymbol{D}_{jv};\ \mathrm{p}_h(i,v) \cap \mathrm{p}_k(i,j) = \{i\} : \mathrm{p}_k(i,j) \text{ is a geodesic}\};\ k = 1,2, \ldots;\ h = 1,2\ldots \tag{8}$$

or

$$V_{i,p_k(i,j)} = \{v|\ v \in V(\mathrm{G});\ \varDelta_{iv} < \varDelta_{jv};\ \mathrm{p}_h(i,v) \cap \mathrm{p}_k(i,j) = \{i\} : \mathrm{p}_k(i,j) \text{ is an elongation}\};\ k = 1,2, \ldots;\ h = 1,2\ldots \tag{9}$$

Quantity $V_{i,p_k(i,j)}$ denotes the set of vertices lying *closer* to vertex i than to vertex j, (condition $D_{iv} < D_{jv}$ – previously proposed by Gutman[26] in defining the Szeged index) and are *external* with respect to path $\mathrm{p}_k(i,j)$ (condition $\mathrm{p}_h(i,v) \cap \mathrm{p}_k(i,j) = \{i\}$). Since in cycle-containing structures, various shortest paths $\mathrm{p}_k(i,j)$, in general, lead to various sets $V_{i,p_k(i,j)}$, by definition, the (ij)-entries in the Cluj matrices are taken as max $|V_{i,p_k(i,j)}|$. The diagonal entries are zero. For paths $\mathrm{p}_h(i,v)$, no restrictions related to their length are imposed. The above definitions (Eqs. (7)–(9)) are valid in any connected graph.

Cluj matrices are also given for the pinane graph in Table I. One can see that all entries in $\boldsymbol{CJ\varDelta}_{\mathrm{u}}$ related to vertex 3 in G_1 are equal to 1, as are those related to the external vertices 8, 9 and 10 (which are at the same time endpoints of the paths that contain them). This property has been called[25] *the internal ending of all longest paths* joining vertex i and the remaining vertices in G_1 and vertex i, like 3, an *internal endpoint*. More details about the Cluj matrices can be found elsewhere.[17,18,23–25]

The two Cluj matrices $\boldsymbol{M}_{\mathrm{u}}$ allow the construction of the corresponding symmetric matrices $\boldsymbol{M}_{\mathrm{p}}$ (defined on paths) and $\boldsymbol{M}_{\mathrm{e}}$ (defined on edges) by:

$$\boldsymbol{M}_{\mathrm{p}} = \boldsymbol{M}_{\mathrm{u}} \bullet (\boldsymbol{M}_{\mathrm{u}})^{\boldsymbol{T}} \tag{10}$$

$$\boldsymbol{M}_{\mathrm{e}} = \boldsymbol{M}_{\mathrm{p}} \bullet \boldsymbol{A} \tag{11}$$

where $\boldsymbol{A}$ is the adjacency matrix (having the non-diagonal entries equal to 1 if vertices i and j are adjacent and zero otherwise).[27] Symbol • means the Hadamard matrix product,[28] *i.e.*, $(\boldsymbol{M}_{\mathrm{a}} \bullet \boldsymbol{M}_{\mathrm{b}})_{ij} = (\boldsymbol{M}_{\mathrm{a}})_{ij}\ (\boldsymbol{M}_{\mathrm{b}})_{ij}$.

In the case of acyclic structures, the two variants of Cluj matrices coincide, as a consequence of the uniqueness of the paths. The symmetric matrices, edge-defined and path-defined ones, in both variants, are identical to the Wiener matrices[29,30] $\boldsymbol{W}_{\mathrm{e}}$ (edge-defined) and $\boldsymbol{W}_{\mathrm{p}}$ (path-defined), respectively.

Recall that for trees, the Wiener index can be calculated[8] by:

$$W = \sum_{ij} N_{i,\mathrm{p}(i,j)} N_{j,\mathrm{p}(i,j)} \tag{12}$$

where $N_{i,\mathrm{p}(i,j)}$ and $N_{j,\mathrm{p}(i,j)}$ have the same meaning in trees as the quantity $N_{i,p_k(i,j)}$ in Eq. (7). The summation runs over all edges. Product $N_{i,\mathrm{p}(i,j)} N_{j,\mathrm{p}(i,j)}$ is the (i,j)-entry in the Wiener matrix $\boldsymbol{W}_\mathrm{e}$ from which W can be calculated as the half-sum of its entries:

$$W = (1/2) \sum_{ij} (\boldsymbol{W}_\mathrm{e})_{ij} \ . \tag{13}$$

One can consider $\boldsymbol{W}_\mathrm{e}$ as the weighted adjacency matrix since $(\boldsymbol{W}_\mathrm{e})_{ij} \neq 0$ if and only if vertices i and j are adjacent. When $\mathrm{p}(i,j)$ represents a path, then a relation similar to (12) defines the *hyper-Wiener* index *WW*:[14–16]

$$WW = \sum_{ij} N_{i,\mathrm{p}(i,j)} N_{j,\mathrm{p}(i,j)} \tag{14}$$

where the product $N_{i,\mathrm{p}(i,j)} N_{j,\mathrm{p}(i,j)}$ is the (i,j)-entry in the Wiener matrix $\boldsymbol{W}_\mathrm{p}$, from which *WW* can be calculated as the half-sum of its entries. The summation in Eq. (14) is over all paths.

In cycle-containing graphs, the Wiener matrices are not defined. Wiener indices are calculated by means of the distance-type matrices as shown above. In such graphs, the two versions of Cluj matrices are different.

Several indices can be derived from the Cluj matrices,[18] either as the half-sum of entries in the corresponding symmetric matrices or directly from the unsymmetric matrices:

$$TI_{\mathrm{e/p}} = \sum_{\mathrm{e/p}} (\boldsymbol{M}_\mathrm{u})_{ij} (\boldsymbol{M}_\mathrm{u})_{ji} \ . \tag{15}$$

When defined on edges, TI_e is a Cluj index: denoted by CD_e or $C\Delta_\mathrm{e}$, depending on whether it is derived from the Cluj-distance or Cluj-detour matrix. Similarly, when defined on paths, TI_p is a hyper-Cluj index denoted by CD_p or $C\Delta_\mathrm{p}$. Values of these indices for a set of 32 acyclic and cyclic octanes are given in Table II.

For cycles, the Cluj-detour indices can be calculated by the formulae:[25]

$$C\Delta_\mathrm{e} = N \tag{16}$$

$$C\Delta_\mathrm{p} = (k + 1)\, N\, (4k^2 + 3yk + 2k + 3y) / 6 \tag{17}$$
$$k = [(N-1)/4];\ y = (N-1) \bmod 4 \ .$$

The edge-defined Cluj-detour index $C\Delta_\mathrm{e}$ is equal to N, the number of vertices, or to the number of edges in a cycle. The path-defined Cluj-detour

TABLE II

Cluj indices of acyclic and cyclic octanes

No	Oct ane	CD_e	CD_p	H_{CDe}	H_{CDp}	$C\Delta_e$	$C\Delta_p$	$H_{C\Delta e}$	$H_{C\Delta p}$
1	N8	84	210	0.648	5.86	84	210	0.64821	5.8593
2	2MN7	79	185	0.708	7.89	79	185	0.70774	7.8939
3	3MN7	76	170	0.724	8.58	76	170	0.72440	8.5244
4	4MN7	75	165	0.729	8.69	75	165	0.72857	8.6897
5	3EN6	72	150	0.745	9.30	72	150	0.74524	9.2952
6	25MN 6	74	161	0.767	10.178	74	161	0.76726	10.1783
7	24MN6	71	147	0.784	10.892	71	147	0.78393	10.8922
8	23MN6	70	143	0.788	11.099	70	143	0.78810	11.0992
9	34MN6	68	134	0.801	11.634	68	134	0.80060	11.6339
10	3E2MN5	67	129	0.805	11.788	67	129	0.80476	11.7881
11	22MN6	71	149	0.784	10.959	71	149	0.78393	10.9589
12	33MN6	67	131	0.805	11.855	67	131	0.80476	11.8547
13	234MN5	65	122	0.848	13.759	65	122	0.84762	13.7587
14	3E3MN5	64	118	0.821	12.571	64	118	0.82143	12.5714
15	224MN5	66	127	0.843	13.577	66	127	0.84345	13.5767
16	223MN5	63	115	0.860	14.402	63	115	0.86012	14.4017
17	233MN5	62	111	0.864	14.598	62	111	0.86429	14.5976
18	2233MN4	58	97	0.919	17.420	58	97	0.91964	17.4196
19	112MC5	67	150	1.004	9.28	34	74	3.42857	15.8452
20	113MC5	71	170	0.970	7.75	32	70	3.09524	17.5119
21	IPC5	73	186	1.002	5.94	40	92	3.85238	15.7523
22	PC5	78	215	0.943	4.37	45	113	3.79286	13.2595
23	11MC6	104	197	0.686	6.73	24	75	4.95238	17.9940
24	12MC6	106	202	0.677	6.46	25	81	4.53571	15.5634
25	13MC6	108	211	0.669	5.78	24	80	4.28571	16.2579
26	14MC6	110	220	0.661	5.61	24	80	4.28571	16.2857
27	EC6	109	226	0.626	4.49	29	94	4.89286	15.1845
28	C8	128	288	0.500	3.28	8	64	8.00000	19.0000
29	123MC5	70	164	0.956	8.22	34	77	2.92857	14.3452
30	1M2EC5	72	178	0.944	6.49	39	93	3.36905	13.2857
31	1M3EC5	76	199	0.911	5.50	37	88	3.03571	14.7023
32	MC7	88	225	0.754	4.62	16	71	6.14286	17.5595

Note that the symbols have the following meaning: N = chain length; M = Methyl; E = Ethyl; P = Propyl; IP = Isopropyl; CN = N–membered cycle. Hence, for example, the sets of symbols 2233MN4 and 1M2EC5 should be read as 2,2,3,3-dimethylbutane and 1-methyl-2-ethyl-cyclopentane, respectively.

index $C\Delta_p$ depends on *mod 4* in a manner similar to that found for the path-defined Cluj-distance index CD_p.[18]

RECIPROCAL DETOUR AND CLUJ-DETOUR INDICES

Harary indices H are constructed on reciprocal matrices $\boldsymbol{M}^r$, *i.e.* matrices having $(\boldsymbol{M}^r)_{ij} = 1/(\boldsymbol{M})_{ij}$:[20–22,31]

$$H = (1/2) \sum_i \sum_j 1/(\boldsymbol{M})_{ij} \tag{18}$$

the symbol $\boldsymbol{M}$ stands for detour and Cluj-detour matrices Δ_e, Δ_p, $\boldsymbol{CJ\Delta}_e$ and $\boldsymbol{CJ\Delta}_p$.

In the case of simple cycles C_N, the Harary-type indices defined on detour and Cluj-detour matrices can be expressed in closed form:

$$H_{\Delta_e} = z\,[N \sum_{i=1}^{(N-1)/2} (N-i)^{-1}] + (1-z)\,[N \sum_{i=1}^{(N-2)/2} (N-i)^{-1} + 1] \tag{19}$$

$$H_{\Delta_p} = z\,\mathrm{N} \sum_{i=1}^{(N-1)/2} [(N-\mathrm{i}+1)\,(N-\mathrm{i})/2]^{-1} + (1-z)\,\{N \sum_{i=1}^{(N-2)/2} [(N-i+1)\,(N-i)/2]^{-1} +$$

$$+ (N/2)\,[(N/2)+1\,(N/2)/2]^{-1}\} \tag{20}$$

$$H_{\boldsymbol{CJ\Delta}_p} = 2\,N \sum_{i=1}^{k} i^{-2} + N\,y\,(k+1)^{-2}/2 \tag{21}$$

where

$$z = N \bmod 2 \tag{22}$$

$$k = [(N-1)/4] \tag{23}$$

$$y = (N-1) \bmod 4\ . \tag{24}$$

Expansion of sums in the above equations leads to:

$$H_{\Delta_e} = -z\,N\,\psi\,[(1-N)/2] - N\,\psi\,(-N/2) + N\,\psi\,(1-N) + 1 + z\,N\psi\,(-N/2) - z \tag{25}$$

$$H_{\Delta_p} = 2\,(N^2 + N - 2\,z)\,/\,(N+1)\,(N+2) \tag{26}$$

$$H_{\boldsymbol{CJ\Delta}_p} = 2N\,[-\psi\,(1,\,k+1) + \pi^2/6] + Ny/2\,(k+1)^2 \tag{27}$$

where $\psi(x) = \Delta[\ln(\gamma(x), x]$, $\psi(N,x) = \Delta(\psi(x), x \in N)$, $\psi(0,x) = \psi(x)$ and $\gamma(x) =$ interpol(exp($-t$) $*t^{\wedge(x-1)}$); $t = 0,\ldots, \infty$. The reader should note that Δ is a mathematical function, which should not be confused with the symbol Δ which stands for the detour matrix.

For cycles, $C\Delta_e = H_{CJ\Delta_e} = N$. Values of reciprocal detour and Cluj-detour indices for the octanes and cyclooctanes are also given in Table II.

MODELING THE BOILING POINTS OF ACYCLIC AND CYCLIC OCTANES

Lukovits[10] used detour-type indices in explaining the variation of boiling points BP, of alkanes. He considered 77 alkanes and cycloalkanes up to $N = 10$ (all acyclic alkanes from methane up to octanes and some cycloalkanes). In that study, the number of carbon atoms N and their square roots $N^{1/2}$ were used as the simplest descriptors and the correlation coefficients obtained ($r = 0.977$, 0.986) were fair. Fractional exponents were used for the Wiener and the detour indices, in single variable regression (equations of the type $BP = a + bI^{1/m}$); none of these indices surpassed the correlation coefficients obtained with $N^{1/2}$. Composite indices, for example, of the type $(W\omega)^{1/m}$ or two variable regressions combining the Wiener and the detour indices yielded much higher correlation coefficients, *e.g.* $(W\omega)^{1/8}$; $r = 0.994$; but the standard deviation s was still high; $s = 6.4$.

When the correlation analysis was repeated on the subset of isomeric (acyclic and cyclic) octanes ($n = 29$), parameter N obviously could not be used any more. Wiener and hyper-Wiener indices did not correlate with boiling points. However, detour and hyper-detour indices showed significantly higher r values, although these were also far from acceptable values (ω: $r = 0.747$; $\omega\omega$: $r = 0.759$). Fractional exponents of these indices or composite indices slowly increased the correlation coefficient, which, however, did not surpass 0.800.

A part of Lukovits' analysis[10] was repeated by Trinajstić and coworkers[6] in their work on the uses of detour matrix in chemistry. They considered 76 lowest alkanes and cycloalkanes. The best structure-boiling point correlation was obtained by $BP = a + b(W\omega)^c$; r being 0.995 and $s = 6.2$.

In the present work, a set of 32 octanes, also investigated by Lukovits,[10] have been considered. Cluj indices and their Harary counterparts have been calculated and listed in Tables II and III. The number of paths of length 2 and 3 (p_2 and p_3) are also given in Table III.

As in the previous study,[18] none of the indices reported in Tables II and III produced an acceptable structure-boiling point correlation. Only the hyper-detour index $\omega\omega$, distance-Cluj index CD_p and the related Harary index H_{CD_p}, surpassed the 0.800 limit ($r = 0.816$, 0.844 and 0.808, respectively).

TABLE III

Boiling points[32] and Wiener, detour and path indices of acyclic and cyclic octanes.

No.	Octane	*BP*	*W*	*WW*	ω	$\omega\omega$	$H_{\boldsymbol{D}\mathrm{e}}$	$H_{\boldsymbol{D}\mathrm{p}}$	$H_{\Delta\mathrm{e}}$	$H_{\Delta\mathrm{p}}$	p_2	p_3
1	N8	125.8	84	210	84	210	13.74	10.56	13.74	10.56	6	5
2	2MN7	117.6	79	185	79	185	14.10	10.86	14.10	10.86	7	5
3	3MN7	118.8	76	170	76	170	14.26	10.98	14.26	10.98	7	6
4	4MN7	117.7	75	165	75	165	14.31	11.01	14.31	11.01	7	6
5	3EN6	118.9	72	150	72	150	14.48	11.13	14.48	11.13	7	7
6	25MN6	108.4	74	161	74	161	14.46	11.16	14.46	11.16	8	5
7	24MN6	109.4	71	147	71	147	14.65	11.30	14.65	11.30	8	6
8	23MN6	115.3	70	143	70	143	14.73	11.36	14.73	11.36	8	7
9	34MN6	118.7	68	134	68	134	14.86	11.46	14.86	11.46	8	8
10	3E2MN5	115.6	67	129	67	129	14.91	11.50	14.91	11.50	8	8
11	22MN6	107.0	71	149	71	149	14.76	11.43	14.76	11.43	9	5
12	33MN6	112.0	67	131	67	131	15.03	11.63	15.03	11.63	9	7
13	234MN5	113.4	65	122	65	122	15.16	11.73	15.16	11.73	9	8
14	3E3MN5	118.2	64	118	64	118	15.25	11.80	15.25	11.80	9	9
15	224MN5	99.3	66	127	66	127	15.16	11.76	15.16	11.76	10	5
16	223MN5	110.5	63	115	63	115	15.41	11.96	15.41	11.96	10	8
17	233MN5	114.6	62	111	62	111	15.50	12.03	15.50	12.03	10	9
18	2233MN4	106.0	58	97	58	97	16.00	12.50	16.00	12.50	12	9
19	112MC5	113.5	56	92	106	278	16.66	13.33	9.45	5.76	12	13
20	113MC5	115.5	58	100	104	266	16.50	13.20	9.52	5.80	12	11
21	IPC5	126.4	62	114	106	286	16.00	12.73	9.78	6.12	10	11
22	PC5	131.0	67	135	111	315	15.56	12.36	9.50	5.90	9	10
23	11MC6	119.5	59	103	119	337	16.33	13.03	8.17	4.49	11	10
24	12MC6	126.6	60	106	124	362	16.16	12.86	7.81	4.19	10	11
25	13MC6	122.3	61	110	123	355	16.08	12.80	7.83	4.21	10	10
26	14MC6	121.8	62	115	122	349	16.03	12.76	7.87	4.22	10	10
27	EC6	131.8	64	122	124	368	15.78	12.53	8.00	4.40	9	10
28	C8	146.0	64	120	160	552	15.66	12.40	5.08	11.60	8	8
29	123MC5	115.0	58	99	109	290	16.41	13.10	9.15	5.50	11	13
30	1M2EC5	124.0	61	110	111	307	16.08	12.80	9.26	5.65	10	12
31	1M3EC5	121.0	63	119	109	294	15.95	12.70	9.37	5.68	10	11
32	MC7	134.0	61	109	142	451	16.00	12.70	6.34	2.80	9	9

Note that the symbols have the following meaning: N = chain length; M = Methyl; E = Ethyl; P = Propyl; IP = Isopropyl; CN = N-membered cycle. Hence, for example, the sets of symbols 2233MN4 and 1M2EC5 should be read as 2,2,3,3-dimethylbutane and 1-methyl-2-ethyl-cyclopentane, respectively.

Multiple linear regressions, with two and three variables, were not successful, either. However, when $1/p_2$ and $1/p_3$ were associated with the detour indices (see Table IV; entries 1 and 2), a correlation coefficient higher than 0.9 was obtained. Combination of four variables, the first two being p_2 and p_3 or their reciprocals, and the last two detour indices, yielded a higher correlation coefficient ($r > 0.950$; entries 4 and 5); on the contrary, Wiener indices did not surpass $r = 0.94$ limit.

Similar results have been obtained by using the Cluj indices and their inverses (see Table IV; entries 6 to 9). Predicting ability of the best regression equations was tested by a cross validation procedure (leave 1/3 out – see below).

TABLE IV

Statistics of multivariable regression.

No.	Variables	r	s	F
1	$1/p_2$; $1/p_3$; ω	0.939	3.326	69.892
2	$1/p_2$; $1/p_3$; $\omega\omega$	0.952	2.980	89.387
3	p_2; p_3; W; WW	0.932	3.588	44.303
4	p_2; p_3; ω; $\omega\omega$	0.966	2.563	93.291
5	$1/p_2$; $1/p_3$; ω; $\omega\omega$	0.956	2.892	71.842
6	p_2; p_3; $H_{\boldsymbol{CD}\mathrm{e}}$, $H_{\boldsymbol{C}\Delta\mathrm{e}}$	0.958	2.826	75.538
7	p_2; p_3; CD_{p}; $H_{\boldsymbol{C}\Delta\mathrm{p}}$	0.961	2.714	82.522
8	p_2; p_3; $H_{\boldsymbol{C}\Delta\mathrm{e};}$ $H_{\boldsymbol{C}\Delta\mathrm{p}}$	0.959	2.802	76.965
9	$1/p_2$; $1/p_3$; $H_{\boldsymbol{C}\Delta\mathrm{e};}$ $H_{\boldsymbol{C}\Delta\mathrm{p}}$	0.966	2.542	94.995

The symbols have the following meaning: r = correlation coefficient; s = standard error of estimate; F = Fischer test.

Note that the best structure-boiling point correlation ($r = 0.966$, $s = 2.5$, $F = 95$) was achieved using the following set of indices: $1/p_2$, $1/p_3$, $H_{\boldsymbol{CD}_e}$ and $H_{\boldsymbol{CD}_p}$ obtained from the Cluj matrices $\boldsymbol{CJ}\Delta_{\mathrm{e}}$ and $\boldsymbol{CJ}\Delta_{\mathrm{p}}$. A cross-validation procedure (leave 1/3 out) led to $r(cv) = 0.934$ and $s(cv) = 3.3$

CONCLUSIONS

Two variants of Cluj matrices have been illustrated. New Harary-type indices, based on the detour and Cluj-detour matrices, have been introduced. The formulae for calculating these indices for simple cycles have been derived. Modeling the boiling points of a set of 32 acyclic and cycle-containing octanes illustrated the superior ability of the Cluj indices and their Harary derivatives in comparison to that of the Wiener indices, to account for the variation in boiling points. This study indicates the possible usefulness of both detour and distance-Cluj indices in modeling physicochemical properties of chemical structures.

Acknowledgements. – This work was supported in part by the Ministry of Science and Technology of the Republic of Croatia through Grant No. 1-07-159. We thank Dr Dejan Plavšić (Zagreb) for his extensive comments on the early version of the manuscript. We also thank the referees for their constructive comments.

REFERENCES

1. F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1969.
2. N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, 1983.
3. F. Buckley and F. Harary, *Distance in Graphs*, Addison-Wesley, Reading, MA, 1990.
4. O. Ivanciuc and A. T. Balaban, *Commun. Math. Comput. Chem.* **30** (1994) 141–152.
5. D. Amić and N. Trinajstić, *Croat. Chem. Acta* **68** (1995) 53–62; see also S. Nikolić, N. Trinajstić, A. Jurić, and Z. Mihalić, *Croat. Chem. Acta* **69** (1996) 1577–1591.
6. N. Trinajstić, S. Nikolić, B. Lučić, D. Amić, and Z. Mihalić, *J. Chem. Inf. Comput. Sci*. **37** (1997) 631–638.
7. M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **36** (1996) 535–540.
8. D. J. Klein, I. Lukovits, and I. Gutman, *J. Chem. Inf. Comput. Sci.* **35** (1995) 50–52.
9. H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17–20; se also S. Nikolić, N.Trinajstić, and Z. Mihalić, *Croat. Chem. Acta* **68** (1995) 105–129.
10. I. Lukovits, *Croat. Chem. Acta* **69** (1996) 873–882.
11. I. Lukovits and M. Razinger, *J. Chem.I nf. Comput. Sci*. **37** (1997) 283–286: see also G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci.*, in press.
12. N. Trinajstić, S. Nikolić, and Z. Mihalić, *Int. J. Quantum Chem.: Quantum Chem. Symp.* **65** (1997) 415–419.
13. M. Randić, L. M. DeAlba, and F. E. Harris, *Croat. Chem. Acta* **71** (1998) 53–68.
14. M. Randić, *Chem. Phys. Lett*. **211** (1993) 478–483.
15. I. Lukovits, *Comput. Chem*. **19** (1995) 27–31.
16. I. Gutman, W. Linert, I. Lukovits, and A. A. Dobrynin, *J. Chem. Inf. Comput. Sci.* **37** (1997) 349–354.
17. M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **36** (1996) 833–836.
18. M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **37** (1997) 300–305.
19. T.-S. Balaban, P. A. Filip, and O. Ivanciuc, *J. Math. Chem.* **11** (1992) 79–105.
20. D. Plavšić, S. Nikolić, N. Trinajstić, and Z. Mihalić, *J. Math. Chem*. **12** (1993) 235–250.
21. O. Ivanciuc, T.-S. Balaban, and A.T. Balaban, *J. Math. Chem.* **12** (1993) 309–318.
22. M. V. Diudea, *J. Chem. Inf. Comput. Sci*. **37** (1997) 292–299.
23. M. V. Diudea, *Commun. Math. Comput. Chem*. **35** (1997) 169–183.
24. I. Gutman and M. V. Diudea, *Commun. Math. Comput. Chem.*, in press.
25. M.V. Diudea, B. Parv, and I. Gutman, *J. Chem. Inf. Comput. Sci*. **37** (1997) 1101–1108.
26. I. Gutman, *Graph Theory Notes New York* **27** (1994) 9–15.
27. *e.g.*, N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, FL, 1992.
28. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1985.

29. M. Randić, X. Guo, T. Oxley, and H. Krishnapriyan, *J. Chem. Inf. Comput. Sci.* **33** (1993) 700–716.
30. M. Randić, X. Guo, T. Oxley H. Krishnapriyan, and L. Naylor, *J. Chem. Inf. Comput. Sci.* **34** (1994) 361–367.
31. O. Ivanciuc, T. Ivanciuc, and M. V. Diudea, *SAR & QSAR Environ. Res.* **7** (1997) 63–87; see also M. V. Diudea and I. Gutman, *Croat. Chem. Acta* **71** (1998) 21–51.
32. Beilsteins Handbuch der Organischen Chemie, Band 1 und 5, Springer-Verlag, Berlin, 1925.

SAŽETAK

Indeks zaobilaženja i Cluj-ski indeks zaobilaženja

Mircea V. Diudea, Gabriel Katona, István Lukovits i Nenad Trinajstić

Prikazane su nedavno uvedene Cluj-ske matrice: Cluj-ska matrica udaljenosti i Cluj-ska matrica zaobilaženja. Definirani su novi indeksi Hararyjeve vrste na matrici zaobilaženja i Cluj-skoj matrici zaobilaženja. Također su izvedene formule za računanje tih indeksa za prstenove. S pomoću Cluj-skih indeksa i njihovih Hararyjevih inačica predviđena su vrelišta za 32 aciklička i ciklička oktana. Najbolje je predviđanje postignuto s pomoću višestruke linearne regresije, uporabom kombinacije dvaju recipročnih brojeva staza ($1/p_2$, $1/p_3$) s indeksima zaobilaženja i hiperzaobilaženja ili kombiniranjem istih recipročnih brojeva staza s Hararyjevim indeksima izvedenim iz dviju Cluj-skih matrica, jedne definirane s pomoću bridova i druge, s pomoću staza.