CROATICA CHEMICA ACTA CCACAA 69 (1) 63–74 (1996)

ISSN 0011-1643
CCA-2288

Original Scientific Paper

Characterization of Chemical Structures by the Atomic Counts of Self-Returning Walks: On the Construction of Isocodal Graphs

Ovidiu Ivanciuc and Alexandru T. Balaban

Department of Organic Chemistry, Faculty of Chemical Technology, University »Politehnica« of Bucharest, Splaiul Independentei 313, 77206 Bucharest, Roumania

Received October 20, 1994; revised January 19, 1995; accepted March 29, 1995

A characterization of chemical structures based on counting self-returning walks in a molecular graph is found to be degenerate for certain pairs of isospectral graphs. On the basis of endospectral graphs, we present (without proof) two theorems for constructing pairs of nonisomorphic graphs with identical atomic counts of self-returning walks.

INTRODUCTION

Chemical graph theory has attracted an increasing research interest in recent years.[1–9] Among the large variety of topics treated, the graph isomorphism problem has received considerable attention. The identification and recognition of identical chemical structures (graph isomorphism problem) remains one of the central problems in many chemical studies involving the chemical species generation and enumeration, computer storage and retrieval of chemical compounds, computer-assisted organic synthesis, chemical data-bases, chemical similarity and structure-property relationships. Out of the large class of graph invariants, we mention here the graph theoretic polynomials and spectra, spectral moments, topological indices, distances, walks and paths in graphs.

By removing all hydrogen atoms from the chemical formula of a chemical compound containing covalent bonds, one obtains the hydrogen-depleted graph (or molecular graph) of that compound, whose vertices correspond to non-hydrogen atoms. In the particular case of hydrocarbons, the vertices of the molecular graph denote carbon atoms.

A number of useful graph definitions will be introduced. Let G = (V,E) be a graph G with N vertices, without loops and multiple edges. The adjacency matrix of graph G, $\boldsymbol{A} = A(\mathrm{G})$, is the square $N \times N$ symmetric matrix which contains information about the connnectivity of the vertices in G. Its entries are defined as:

$$(A)_{ij} = \begin{cases} 1, \text{ for vertices } i, j \text{ adjacent} \\ 0, \text{ otherwise} \end{cases}$$

A walk in a graph is a sequence of edges which can be continuously traversed, starting from any vertex and ending on any vertex. Repeated use of the same edge or edges is allowed. A self-returning walk is a walk starting and finishing at the same vertex. The length of a walk is the total number of edges that are traversed.

Self-returning walks of length k may be computed by considering the diagonal elements of the first k powers of the adjacency matrix $\boldsymbol{A}$, due to the fact that each diagonal element $(A^k)_{ii}$ of matrix $\boldsymbol{A}^k$ can be interpreted as the sum of all self-returning walks of lengths k from/to vertex i.[10,11] The sequence of integers $\{(A^1)_{ii}, (A^2)_{ii}, \ldots, (A^N)_{ii}\}$ defines the self-returning walk atomic code (SRWAC) of atom i in a molecule.[12] The SRWAC characterizes the environment of a given atom in a molecule.

Randić[12] conjectured that the atomic codes defined on the basis of self-returning walks are a complete set of graph invariants, *i.e.* there is no pair of nonisomorphic graphs with the same collection of atomic codes.

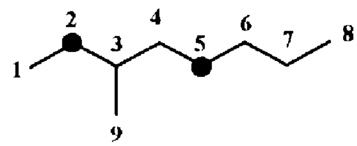
The characteristic or spectral polynomial Ch(h,x) of the molecular graph G is the characteristic polynomial of its adjacency matrix:[13]

$$\mathrm{Ch}(\mathrm{G}, x) = \det(x\boldsymbol{I} - \boldsymbol{A}) \tag{1}$$

where $\boldsymbol{I}$ is the $N \times N$ unit matrix. Although it was initially conjectured that the characteristic polynomial might be used as a unique descriptor of graphs, nonisomorphic graphs with the same characteristic polynomial were found,[14–20] and called isospectral or cospectral graphs.

An important connection between the characteristic polynomial of a molecular graph and the count of walks in the graph is stated by the Cayley-Hamilton theorem.[21] According to this theorem, if Ch(x) is the characteristic polynomial of matrix $\boldsymbol{A}$, then:

$$\mathrm{Ch}(\boldsymbol{A}) = 0\ . \tag{2}$$



$\mathbf{T_1}$

Scheme I

For example, the characteristic polynomial of tree T_1 (Scheme I) is

$$\mathrm{Ch}(T_1) = x^9 - 8x^7 + 20x^5 - 17x^3 + 4x \ . \tag{3}$$

From the Calyley-Hamilton theorem, the following equation is satisfied for the corresponding elements $(A^k)_{ij}$ of powers of the adjacency matrix of tree T_1:

$$(A^9)_{ij} - 8\,(A^7)_{ij} + 20\,(A^5)_{ij} - 17\,(A^3)_{ij} + 4\,(A)_{ij} = 0 \ . \tag{4}$$

The structural code of vertex i (SC_i) was defined as:[22]

$$SC_i = \sum_{k=1}^{N} (A^k)_{ii} \ . \tag{5}$$

Based on the SC, Barysz and Trinajstić[22] defined the ordered structural code (OSC) as the ascending ordered sequence of SCs in a molecule.

As an example, the SRWACs and SCs of vertices in tree T_1 are given in Table I. Only even-length walks are given, because in trees there are no odd-length self-returning walks.

For tree T_1 the OSC sequence is given below:

$$\mathrm{OSC}(T_1) = \{22,\ 30,\ 57,\ 63,\ 90,\ 107,\ 107,\ 143,\ 219\} \ .$$

On the basis of the OSC, Barisz and Trinajstić proposed the following conjecture: Two trees are isomorphic if and only if they have identical ordered structural codes.

In certain cases nonequivalent vertices in a molecular graph have identical SRWACs.[23,24] Such vertices are termed endospectral vertices, and the corresponding graph is termed an endospectral graph.

The concept of endospectral graphs appeared in connection with the problem of the graph isomorphism problem[25,26] and isospectral graphs (distinct graphs with identical spectrum, spectral moments and characteristic

TABLE I

Self-returning walk atomic codes and structural counts of tree T_1

Vertex	Walk length 2	4	6	8	SC
1	1	2	6	21	30
2 5	2	6	21	78	107
3	3	11	42	163	219
4	2	7	27	107	143
6	2	6	19	63	90
7	2	5	14	42	63
8	1	2	5	14	22
9	1	3	11	42	57

polynomial).[27–30] For example, tree T_1, studied by Schwenk,[29] has two endospectral vertices, namely 2 and 5; any subgraph attached to either vertex 2 or vertex 5 produces a pair of isospectral graphs. The endospectral vertices are depicted as distinct circles.

As a consequence of the Cayley-Hamilton theorem, if the SRWACs of two nonequivalent vertices in a graph are identical up to the Nth power of the adjacency matrix, they will present identical values also for higher powers of the adjacency matrix.

Recently, the collection of irreducible endospectral trees up to 16 vertices was reported.[31] Endospectral graphs are responsible for the occurence of a great number of isospectral trees, leading to, when one considers trees of increasing size, the situation that led Schwenk[29] to give the proposition:

Proposition. – If P_n denotes the probability that a random tree on n vertices has another tree cospectral with it, then P_n tends to one as n tends to infinity.

The simplest way of producing a pair of isospectral graphs from tree T_1 is to connect a vertex by a single edge to either vertex 2 or vertex 5. Trees T_2 and T_3, obtained by the above procedure from tree T_1 (Scheme II), exhibit the same characteristic polynomial:

T_2 T_3

Scheme II

$$\mathrm{Ch}(\mathrm{T}_2) = \mathrm{Ch}(\mathrm{T}_3) = x^{10} - 9x^8 + 26x^6 - 27x^4 + 8x^2 \tag{6}$$

Isocodal vertices can also occur in different graphs, as illustrated in Figure 1 for trees $T_4 - T_{11}$, where isocodal vertices are represented as black enlarged circles. The SRWACs of the isocodal vertices of the graphs in Figure 1, corresponding to even-length walks up to the 20th power of the adjacency matrix, are presented below:

SRWAC(T_4,v) = SRWAC(T_5,v) = {2 6 20 68 232 792 2704 9232 31520 107616}

SRWAC(T_6,v) = SRWAC(T_7,v) = {3 11 43 171 683 2731 10923 43691 174763 699051}

SRWAC(T_8,v) = SRWAC(T_9,v) = {2 7 29 124 533 2293 9866 42451 182657 785932}

SRWAC(T_{10},v) = SRWAC(T_{11},v) = {2 6 22 86 342 1366 5462 21846 87382 349526}

If one connects with an edge two isocodal vertices in two different graphs, an endospectral graph is obtained. This procedure, if applied to the four pairs of trees in Figure 1, gives the four irreducible endospectral trees

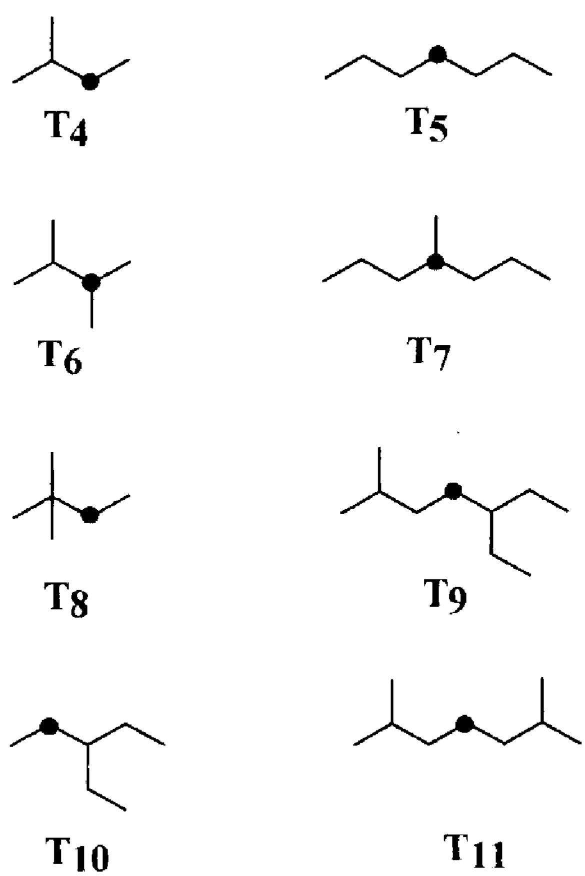


Figure 1. Pairs of trees with isocodal vertices; isocodal vertices are represented as black enlarged circles.

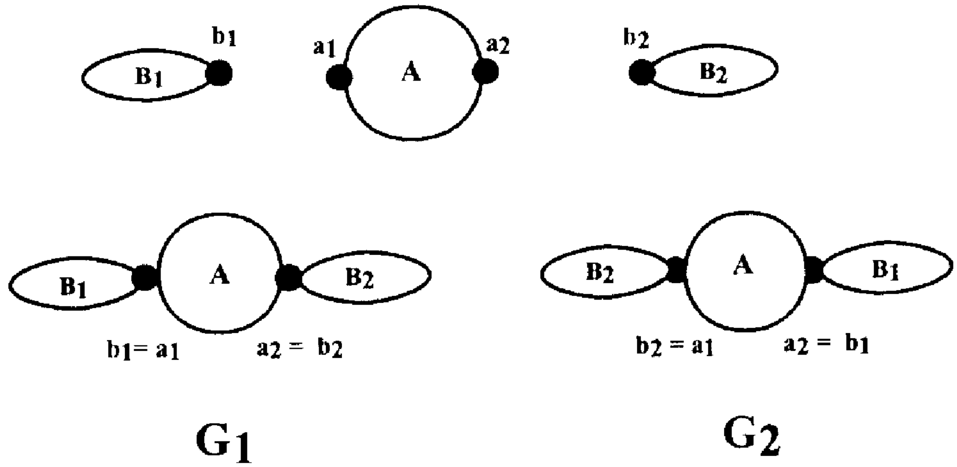
with adjacent endospectral vertices from the collection of endospectral trees.[31] This is a simple method for constructing pairs of endospectral graphs. A systematic search for isocodal vertices in trees up to 16 vertices revealed the existence of a great number of pairs of nonisomorphic trees with isocodal vertices.[23]

ISOCODAL GRAPHS

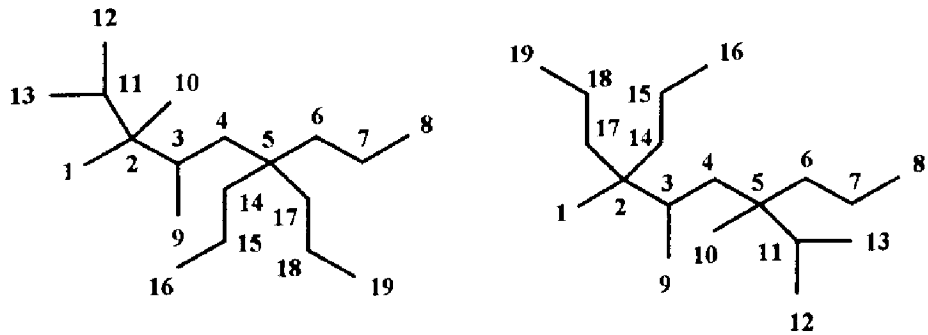
As stated above, the characteristic polynomial of a molecular graph is not a unique structural descriptor. The analysis of the structural causes of its degeneracy led to the characterization of molecular structures using SRWAC[12] and OSC.[22] Recently, a graphical procedure for obtaining pairs of isocodal graphs, *i.e.* graphs with identical atomic codes, was presented.[32] Using the graphical procedure, a pair of 5-trees (graphs with the highest vertex degree 5) with 22 vertices was obtained, which is the smallest pair of isocodal trees generated. A pair of isocodal 3-trees with 26 vertices was also generated. This is a remarkable fact from the organic chemical viewpoint because the molecular graphs of organic compounds have degrees of at most four.

The negative answer to the conjecture that atomic codes are a complete set of invariants is not the end of interest in SRWAC. First, because its relative low degeneracy makes it fit for practical purposes, and deserves further development for unsaturated and heteroatom containing molecules. Second, further studies, revealing the structural conditions of the apparition of degenerate atomic codes may lead to the development of new, more selective graph-theoretical invariants.

In the present paper, we report some more general results concerning isocodal graphs. Two theorems concerning the graphical construction of isocodal graphs are presented and exemplified for trees and cyclic graphs.



Scheme III



T_{12} T_{13}

Scheme IV

Theorem 1. – Let A be a graph with two endospectral vertices a_1 and a_2. Let b_1 be a vertex in a graph B_1 and b_2 a vertex in a graph B_2 such that vertices b_1 and b_2 have the same walk-based atomic codes, *i.e.* the same numbers of self-returning walks for each length of walk (Scheme III).

If G_1 is the graph constructed from A, B_1 and B_2 by identifying vertices a_1 with b_1 and identifying a_2 with b_2 and G_2 is the graph constructed from A, B_1 and B_2 by identifying vertices a_1 with b_2 and identifying a_2 with b_1, then there exists a one-to-one correspondence of the self-returning walk atomic code for vertices from G_1 and G_2.

A similar constructive rule was used to obtain pairs of graphs with an identical distance degree sequence and distance sum sequence.[33]

Theorem 1 enables one to generate a pair of isocodal 4-trees with 19 vertices, namely T_{12} and T_{13} (Scheme IV), by connecting the isocodal vertices

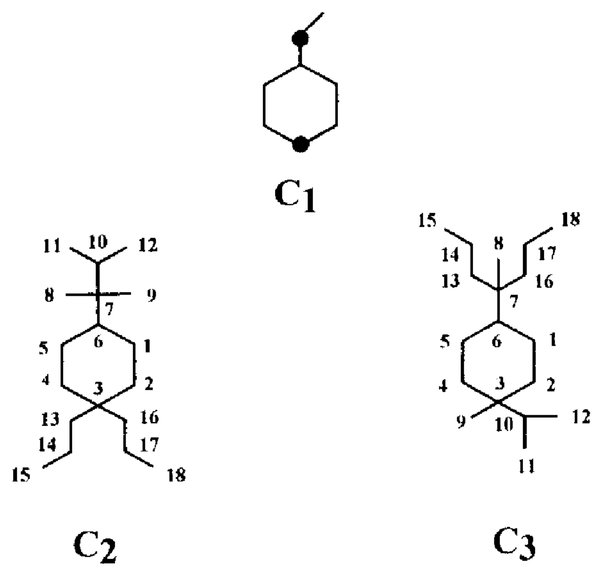
TABLE II

Self-returning walk atomic codes of the isocodal trees T_{12} and T_{13}

	Walk length								
Vertex	2	4	6	8	10	12	14	16	18
1 10	1	4	20	105	560	3016	16377	89580	493196
2 5	4	20	105	560	3016	16377	89580	493196	2731049
3	3	13	66	357	1981	11114	62689	354705	2011226
4	2	9	47	259	1455	8235	46763	266015	1514939
6 14 17	2	8	37	185	962	5109	27493	149378	817953
7 15 18	2	5	16	64	297	1492	7796	41593	224768
8 16 19	1	2	5	16	64	297	1492	7796	41593
9	1	3	13	66	357	1981	11114	62689	354705
11	3	12	56	281	1460	7732	41465	224512	1225384
12 13	1	3	12	56	281	1460	7732	41465	224512

of trees T_4 and T_5, respectively, to the two endospectral vertices of tree T_1. The atomic codes of the vertices in the isocodal trees T_{12} and T_{13} are presented in Table II.

When the pair of trees with isocodal vertices T_4 and T_5 are connected to the two endospectral vertices of graph C_1, a pair of isocodal cyclic graphs, C_2 and C_3, are generated (Scheme V). The atomic codes of the vertices in the isocodal graphs C_2 and C_3 are presented in Table III.

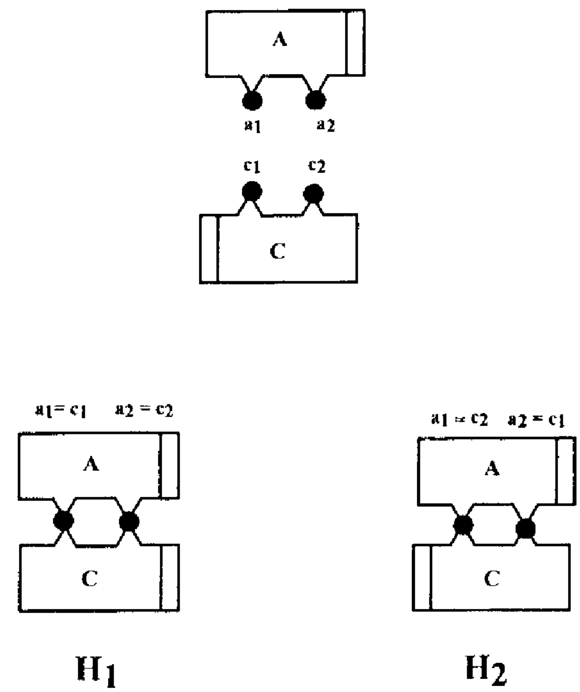


Scheme V

TABLE III

Self-returning walk atomic codes of the isocodal monocyclic graphs C_2 and C_3

Vertex		Walk length 2	4	6	8	10	12	14	16	18
1	5	2	7	34	187	1074	6267	36794	216547	1275714
2	4	2	8	41	226	1275	7270	41735	240806	1395031
3	7	4	20	106	574	3150	17474	97846	552410	3141126
6		3	14	74	412	2348	13536	78528	457352	2670744
8	9	1	4	20	106	574	3150	17474	97846	552410
10		3	12	56	282	1478	7934	43298	239302	1336426
11	12	1	3	12	56	282	1478	7934	43298	239302
13	16	2	8	37	186	978	5276	28940	160704	901248
14	17	2	5	16	64	298	1510	7998	43426	239558
15	18	1	2	5	16	64	298	1510	7998	43426

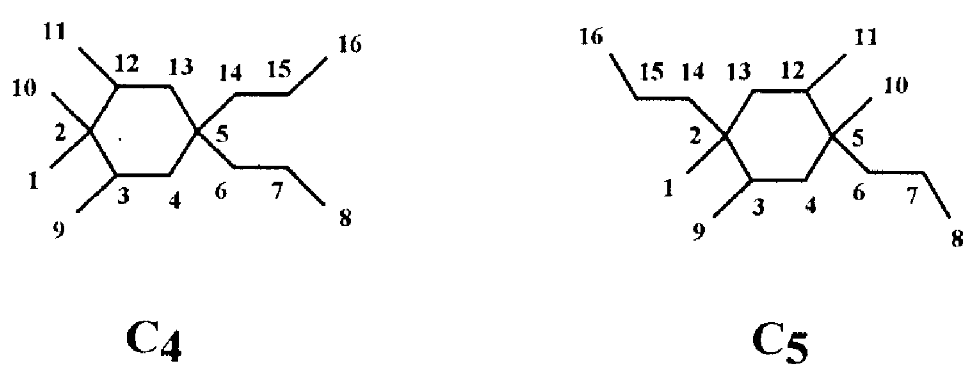


Scheme VI

Theorem 2. – Let A be a graph with two endospectral vertices a_1 and a_2. Let C be a graph with two endospectral vertices c_1 and c_2 (Scheme VI).

If H_1 is the graph constructed from A and C by identifying vertices a_1 with c_1 and identifying a_2 with c_2 and H_2 is the graph constructed from A and C by identifying vertices a_1 with c_2 and identifying a_2 with c_1, then there exists one-to-one correspondence of the self-returning walk atomic code for vertices from H_1 and H_2.

A pair of isocodal monocyclic graphs with 16 vertices, C_4 and C_5, is obtained when the procedure stated by Theorem 2 is applied to two trees T_1 (Scheme VII). The atomic codes of the vertices in the two isocodal graphs C_4 and C_5 are presented in Table IV. Graphs C_4 and C_5 represent the smallest pair of known isocodal graphs.



Scheme VII

TABLE IV

Self-returning walk atomic codes of the isocodal monocyclic graphs C_4 and C_5

Vertex		Walk length							
		2	4	6	8	10	12	14	16
1	10	1	4	20	108	608	3520	20784	124416
2	5	4	20	108	608	3520	20784	124416	751808
3	11	3	13	68	388	2300	13872	84384	515696
4	12	2	9	49	288	1740	10620	65088	399680
6	14	2	8	37	188	1016	5724	33184	196208
7	15	2	5	16	64	300	1552	8528	48688
8	16	1	2	5	16	64	300	1552	8528
9	13	1	3	13	68	388	2300	13872	84384

CONCLUDING REMARKS

Two theorems, representing methods of constructing isocodal graphs, are presented, along with some examples of pairs of isocodal graphs.

A pair of isocodal 4-trees with 19 vertices and a pair of isocodal monocyclic graphs with 16 vertices were obtained. They represent the smalest known isocodal graphs representing trees and cyclic graphs, respectively. Since no exhaustive search for isocodal graphs was made, we do not claim that there are no smaller pairs of isocodal trees and cyclic graphs, respectively.

Acknowledgement. – We thank the Ministry of Research and Technology for financial support of this research under Grants 572B TA4 and TB11.

REFERENCES

1. N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, Florida, 1983.
2. A. T. Balaban (Ed.), *Chemical Applications of Graph Theory*, Academic Press, London, 1976.
3. J. W. Kennedy and L. V. Quintas (Eds.), *Applications of Graphs in Chemistry and Physics*, North Holland, Amsterdam, 1988.
4. A. T. Balaban, *J. Chem. Inf. Comput. Sci.* **25** (1985) 334;
   A. T. Balaban, *J. Mol. Struct. (Theochem)* **120** (1985) 117;
   A. T. Balaban, *J. Mol. Struct. (Theochem)* **165** (1988) 243.
5. N. Trinajstić, D. J. Klein, and M. Randić, *Int. J. Quantum Chem.: Quantum Chem. Symp.* **20** (1986) 699.
6. K. Balasubramanian, *Chem. Rev.* **85** (1985) 599.

7. R. B. King (Ed.), *Chemical Applications of Topology and Graph Theory, Stud. Phys. Theor. Chem.* Vol. 28, Elsevier, Amsterdam, 1983.
8. R. B. King and D. H. Rouvray (Eds.), *Graph Theory and Topology in Chemistry, Stud. Phys. Theor. Chem.* Vol 51, Elsevier, Amsterdam, 1987.
9. D. H. Rouvray, *Acta Pharm. Jugosl.* **36** (1986) 239;
   D. H. Rouvray, *J. Comput. Chem.* **8** (1987) 470;
   D. H. Rouvray, *J. Mol. Struct. (Theochem)* **185** (1989) 187.
10. F. Harary, *Graph Theory*, Adison-Wesley, Reading, Mass., 1969.
11. N. Deo, *Graph Theory with Applications to Engineering and Computer Science*, Prentice-Hall, Englewood Cliffs, New Jersey, 1974.
12. M. Randić, *J. Comput. Chem.* **1** (1980) 386.
13. N. Trinajstić, *J. Math. Chem.* **2** (1988) 197.
14. L. Collatz and U. Sinogowitz, *Abh. Math. Sem. Univ. Hamburg* **21** (1957) 63.
15. C. A. Baker, *J. Math. Phys.* **7** (1966) 2238.
16. M. E. Fisher, *J. Comb. Theory* **1** (1966) 105.
17. A. T. Balaban and F. Harary, *J. Chem Doc.* **11** (1971) 258.
18. F. Harary, C. King, A. Mowshowitz, and R. C. Read, *Bull. London Math. Soc.* **3** (1971) 321.
19. W. C. Herndon and M. E. Ellzey Jr., *Tetrahedron* **31** (1975) 99.
20. M. Randić, N. Trinajstić, and T. Živković, *J. Chem. Soc., Faraday Trans. 2* **72** (1976) 244.
21. F. R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1964, Vol. 1, p. 83.
22. M. Barysz and N. Trinajstić, *Int. J. Quantum Chem.: Quantum Chem. Symp.* **18** (1984) 661;
   M. Barysz, J. V. Knop, S. Pejaković, and N. Trinajstić, *Polish J. Chem.* **59** (1985) 405.
23. J. V. Knop, W. R. Müller, K. Szymanski, M. Randić, and N. Trinajstić, *Croat. Chem. Acta* **56** (1983) 405.
24. M. Randić, W. L. Woodworth, and A. Graovac, *Int. J. Quantum Chem.* **24** (1983) 435.
25. G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci.* **30** (1990) 187;
   G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci.* **31** (1991) 123;
   G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci.* **31** (1991) 422.
26. X. Liu and D. J. Klein, *J. Comput. Chem.* **12** (1991) 1243.
27. W. C. Herndon, *Tetrahedron Lett.* (1974) 671.
28. T. Živković, N. Trinajstić, and M. Randić, *Mol. Phys.* **30** (1975) 517.
29. A. J. Schwenk, in: F. Harary (Ed.), *New Directions in the Theory of Graphs*, Academic Press, New York, 1973, pp. 275-307.
30. Y. Jiang, *Sci. Sin.* **27** (1984) 326.
31. J. V. Knop, W. R. Müller, K. Szymanski, N. Trinajstić, A. F. Kleiner, and M. Randić, *J. Math. Phys.* **27** (1986) 2601.
32. O. Ivanciuc and A. T. Balaban, *J. Math. Chem.* **11** (1992) 155.
33. A. T. Balaban and L. V. Quintas, *MATCH (Commun. Math. Chem.)* **14** (1983) 213.

SAŽETAK

Karakterizacija kemijskih struktura čvornim brojevima zatvorenih šetnji: O konstrukciji izokodalnih grafova

Ovidiu Ivanciuc i Alexandru T. Balaban

Poznato je da opis kemijske strukture brojanjem zatvorenih šetnji u molekularnom grafu daje degenerirane rezultate za određene parove izospektralnih grafova. Na osnovi endospektralnih grafova prikazana su dva teorema (bez dokaza) za konstrukciju parova neizomorfnih grafova s identičnim čvornim brojevima zatvorenih šetnji.