Labeling of Benzenoid Systems which Reflects the Vertex-Distance Relations

Sandi Klavžar*

Department of Mathematics, PF, University of Maribor, Koroška cesta 160, 62000 Maribor, Slovenia

Ivan Gutman

Institute of Physical Chemistry, Attila Jozsef University, P.O. Box 105, H-6701 Szeged, Hungary

Bojan Mohar

Department of Mathematics, University of Ljubljana, 61111 Ljubljana, Slovenia

Received February 1, 1995[⊗]

It is shown that the vertices of benzenoid systems admit a labeling which reflects their distance relations. To every vertex of a molecular graph of a benzenoid hydrocarbon a sequence of zeros and ones (a binary number) can be associated, such that the number of positions in which these sequences differ is equal to the graph-theoretic vertex distance. It is shown by an example that such labelings can be used not only for nomenclature purposes but also for fast evaluation of molecular parameters based on the graph distance.

1. INTRODUCTION

In chemical graph theory[1] there is a large number of molecular topological indices based on vertex distances. The best known among them are the classical Wiener[2] and Balaban[3] indices, but quite a few others are also encountered. Of them we mention the family of molecular topological indices introduced by Schultz, Schultz, and Schultz,[4−12] the Harary index,[13,14] the family of hyper-Wiener indices[15−17] and the recently proposed Szeged index.[18] For review and comparative studies of vertex-distance-based molecular structure descriptors see the articles published by Mihalić and others.[13,19,20]

A seemingly unrelated field of activity in contemporary chemical graph theory is the search for methods to canonically label the atoms of chemical compounds so that the labeling is as much as possible structure-based (and not conventional) and as much as possible convenient for computer-aided manipulation with structural information. Of the numerous works in this direction we mention S. B. Elk's pioneering efforts aimed at polycyclic molecules, especially at benzenoid systems.[21−23]

In this paper we offer results that are related to both the canonical labeling of the vertices of benzenoid systems and to vertex distances. In order to be able to formulate them we need some preparation. We also find it useful to make the reader familiar with the basic features of the (mathematical) theory of Hamming graphs.

The molecular graphs of benzenoid hydrocarbons will be referred to as benzenoid systems; their properties are discussed in due detail elsewhere.[1,24] In mathematical literature the very same objects are called "hexagonal systems" or "hexagonal animals".[24] Notice that our definition of a benzenoid system (the same as given in ref 24) excludes coronacondensed systems, benzene modules, helicenes, and other related polycyclic moieties. Hence, the benzenoid systems considered here are either cata- or pericondensed.

A *benzenoid system* G is a finite connected plane graph with no cut vertices in which every interior region is bounded by a regular hexagon of side length 1. (Recall that a *plane graph* is the graph together with its drawing in the plane.) A straight line segment **C** in the plane with end points P_1 and P_2 is called a *cut segment* if **C** is orthogonal to one of the three edge directions, each P_1 and P_2 is the center of an edge, and the graph obtained from G by deleting all edges intersected by **C** has exactly two connected components. An *elementary cut* C of a cut segment **C** is the set of all edges intersected by **C**. By C_{uv} we will denote the elementary cut containing an edge uv. In Figure 1, a cut segment **C** is indicated, and e_1, e_2, and e_3 are the edges of the corresponding elementary cut. (We refer to ref 25 for more information on the defined terms.)

Let Σ be a finite alphabet and let w_1 and w_2 be words of equal length over Σ. Then the *Hamming distance* between w_1 and w_2, $H(w_1,w_2)$, is the number of positions in which w_1 and w_2 differ. A graph G is a *Hamming graph*, if each vertex $v \in V(G)$ can be labeled by a word $l(v)$ of fixed length, such that $H(l(u),l(v)) = d_G(u, v)$ for all $u, v \in V(G)$. Any such labeling is called a *canonical labeling*. Here, $d_G(u, v)$ denotes the usual shortest path distance in G between u and v. A path between two vertices of a graph G is also called a G-path. In particular, if $\Sigma = \{0, 1\}$, we call G a *binary Hamming graph*. An example of a binary Hamming graph and an appropriate labeling is shown in Figure 2. We refer to refs 26 and 27 for the original interest for these graphs in communication theory.

Binary Hamming graphs were first characterized by Djoković,[28] and several other characterizations were obtained later. To recall the characterization from ref 28 we need two more definitions.

A subgraph H of a graph G is said to be *convex* if it is connected and if any shortest path in the graph G between two vertices of H lies entirely in H. For an edge uv of a graph G let V_{uv} be the set of vertices in G which are closer to u than to v, i.e.

$$V_{uv} = \{w | w \in V(G), d_G(w, u) < d_G(w, v)\}$$

[⊗] Abstract published in *Advance ACS Abstracts*, April 15, 1995.

0095-2338/95/1635-0590$09.00/0

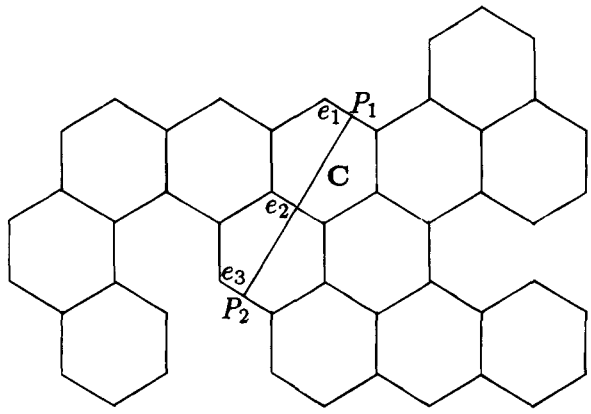

Figure 1. A cut segment.

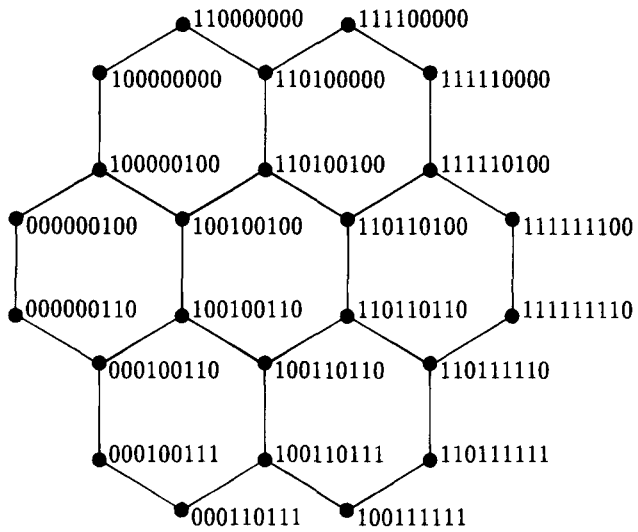

Figure 2. A binary Hamming graph. Observe that the (binary) words by which its vertices are labeled are of length nine—this is related to the existence of nine elementary cuts in coronene (for details see text).

Note that in a bipartite graph G, V_{uv} and V_{vu} form a partition of $V(G)$. Now we can state the aforementioned result of Djoković:[28]

Theorem 1.1. *A graph G is a binary Hamming graph if and only if G is bipartite and for every edge uv of G, V_{uv} and V_{vu} induce convex subgraphs of G.*

2. HEXAGONAL SYSTEMS AS BINARY HAMMING GRAPHS

The following result is the principal observation of this paper.

Theorem 2.1. *A benzenoid system is a binary Hamming graph.*

The actual method by which the canonical labeling of the vertices of a benzenoid system is achieved is explained in detail later in this paper (before Lemma 2.2) and is illustrated in Figure 3.

Proof. Let G be a benzenoid system. It is well-known that G is bipartite.[24] Thus by theorem 1.1 it remains to prove that for every edge uv of G, V_{uv} and V_{vu} induce convex subgraphs of G.

Consider the elementary cut C_{uv}. The graph $G\backslash C_{uv}$ consists of two connected components, and it is straightforward to see that they are precisely the graphs induced by the sets V_{uv} and V_{vu}.

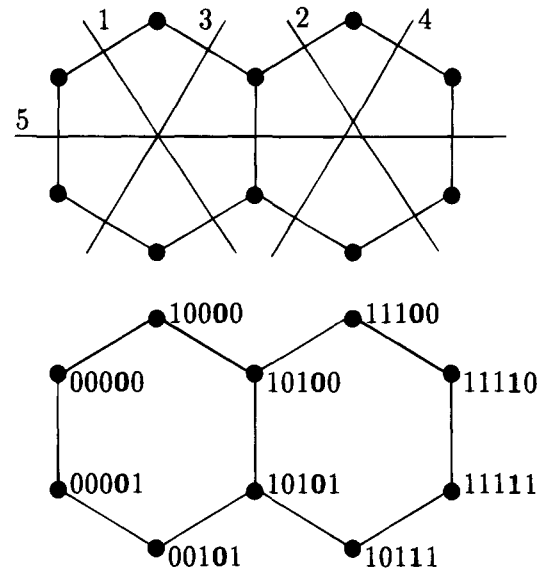

Figure 3. A simple example.

Finally, we want to show that V_{uv} induces a convex subgraph of G. Suppose not. Then there are vertices $x, y \in V_{uv}$ and a shortest path P between x and y which is not completely in V_{uv}. It follows that P is of the form $Qu'v'Q'v''u''Q''$, where $u'v'$ and $v''u''$ are edges of the elementary cut C_{uv}, Q is a shortest path between x and u' in V_{uv}, Q' a shortest path between v' and v'' in V_{vu}, and Q'' a shortest path between u'' and y in V_{uv}. However, there is a unique shortest path between u' and u'' (which lies completely in V_{uv}) that passes along the elementary cut C_{uv}. It follows that P is not a shortest path, a contradiction.

Since a benzenoid system is a binary Hamming graph, one would like to obtain a corresponding labeling fast. In what follows we present such an algorithm which can be very easily carried out by hand.

Let $C_1, C_2, \ldots, C_k$ be the elementary cuts of a benzenoid system G. Let G_i^0 and G_i^1 be the connected components of the graph $G\backslash C_i$, $1 \le i \le k$. Define a labeling

$$l: V(G) \rightarrow \{0, 1\}^k$$

in the following way. For $u \in V(G)$ let the ith component of $l(u)$, which we denote by $l_i(u)$, be defined by

$$l_i(u) = \begin{cases} 0, & \text{if } u \in G_i^0 \\ 1, & \text{if } u \in G_i^1 \end{cases}$$

Hence, $l(u) = (l_1(u), l_2(u), \ldots, l_k(u))$. We claim that l represents a canonical labeling of G. For this purpose we first need the following lemma. Its proof is analogous to the last part of the proof of theorem 2.1, thus we will not repeat it here.

Lemma 2.2. *Let G be a benzenoid system, let C be an elementary cut of G, and let u and v be vertices of G. If P is a shortest u−v path in G, then $|P \cap C| \le 1$.*

We now have the following:

Theorem 2.3. *For arbitrary vertices u and v of a benzenoid system G,*

$$H(l(u),l(v)) = d_G(u, v)$$

Proof. Let P be any shortest path between u and v.

By lemma 2.2 each edge of P belongs to a different elementary cut of G. It follows that $l(u)$ and $l(v)$ differ in at least as many coordinates as the number of edges on P. In other words, $H(l(u),l(v)) \geq d_G(u, v)$.

To prove the converse, consider any elementary cut C_i of G which has no edge in common with P. Since $G\backslash C_i$ consists of two connected components, P lies entirely in one of them. Therefore, $l_i(u) = l_i(v)$, and hence $l(u)$ and $l(v)$ differ in at most as many coordinates as the number of edges on P, i.e., $H(l(u),l(v)) \leq d_G(u, v)$.

Combining both inequalities completes the proof.

For a simple example of the above labeling procedure consider the benzenoid system in Figure 3; it has five elementary cuts that are indicated above, and the corresponding labeling is presented below. We have assumed that the graph G_i^0, $1 \leq i \leq 5$, is the left (respectively the bottom) connected component of the graph $G\backslash C_i$. For instance, the fourth coordinate is indicated bold.

Note also that the benzenoid system from Figure 2 has nine elementary cuts, and thus the corresponding labeling is a binary number of length 9. The order of cuts in Figure 2 is selected analogously as in Figure 3 (but, of course, any order would do). The first set of cuts is at $-60°$, the second set is at 60°, and the third set is horizontal. Each of these sets define three bits.

The results of this section can be directly extended to generalized benzenoid systems (as defined in ref 25) and to coronoid hydrocarbons[29] as well as to various other classes of polycyclic graphs of chemical interest. We would also like to mention that the above labeling procedure could also be deduced from the algorithm given in ref 30 which, in turn, is based on the theory of isometric embeddings of graphs into Cartesian products developed by Graham and Winkler.[31]

3. SOME APPLICATIONS

Besides the obvious possibility of applying the Hamming labeling of a hexagonal system for nomenclature purposes we show in this section how it can be used for fast calculation of molecular parameters based on the graph distance. As an example we present the simplest case of the Wiener index $W(G)$ which is equal to the sum of distances $d_G(u, v)$ taken over all pairs of vertices u, v of G.

Let l: $V(G) \rightarrow \{0, 1\}^q$ be a Hamming labeling of length q of a benzenoid system G with $n = |V(G)|$ vertices. (Indeed, G can be any binary Hamming graph.) For $i = 1, ..., q$, let n_i be the number of vertices u of G such that the ith component of $l(u)$ is equal to 1.

Proposition 3.1. *The Wiener index of G is equal to*

$$W(G) = \sum_{i=1}^{q} n_i(n - n_i) \qquad (1)$$

Proof. Let $V = V(G)$. For $u, v \in V$, let $\delta_i(u, v)$ be 0 if ith components of $l(u)$ and $l(v)$ agree, and 1 otherwise. Since l is a Hamming labeling, we have

$$\begin{aligned} W(G) &= \frac{1}{2}\sum_{u\in V}\sum_{v\in V} d_G(u, v) \\ &= \frac{1}{2}\sum_{u\in V}\sum_{v\in V} H(l(u),l(v)) \\ &= \frac{1}{2}\sum_{u\in V}\sum_{v\in V}\sum_{i=1}^{q} \delta_i(u, v) \\ &= \sum_{i=1}^{q}\left(\frac{1}{2}\sum_{u\in V}\sum_{v\in V} \delta_i(u, v)\right) \\ &= \sum_{i=1}^{q} n_i(n - n_i) \end{aligned}$$

Proposition 3.1 could also be derived directly. Let C_1, ..., C_q be the elementary cuts (as in the previous section), and for $i = 1, ..., q$ let G_i^0 and G_i^1 be the corresponding components of $G\backslash C_i$. Then $n_i = |V(G_i)|$. By using the counting method of Wiener[2] (cf. also papers of Lukovits and Gutman[32] and Mohar and Pisanski[33]), one can derive (1) directly. The advantage of (1) is that one can store information about the graph G just by representing $l(u)$, $u \in V(G)$, and then being able to compute various distance-based parameters without really reconstructing the graph.

ACKNOWLEDGMENT

This work was supported in part by the Ministry of Science and Technology of Slovenia.

REFERENCES AND NOTES

(1) Trinajstić, N. *Chemical Graph Theory*, 2nd revised ed.; CRC Press: Boca Raton, FL, 1992.
(2) Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **1947**, *69*, 17−20.
(3) Balaban, A. T. Highly discriminating distance-based topological index. *Chem. Phys. Lett.* **1982**, *89*, 399−404.
(4) Schultz, H. P. Topological organic chemistry. I. Graph theory and topological indices of alkanes. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 227−228.
(5) Schultz, H. P.; Schultz, E. B.; Schultz, T. P. Topological organic chemistry. 2. Graph theory, matrix determinants and eigenvalues, and topological indices of alkanes. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 27−29.
(6) Schultz, H. P.; Schultz, T. P. Topological organic chemistry. 3. Graph theory, binary and decimal adjacency matrices, and topological indices of alkanes. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 144−147.
(7) Schultz, H. P.; Schultz, E. B.; Schultz, T. P. Topological organic chemistry. 4. Graph theory, matrix permanents and topological indices of alkanes. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 69−72.
(8) Schultz, H. P.; Schultz, T. P. Topological organic chemistry. 5. Graph theory, matrix Hafnians and Pfaffians, and topological indices of alkanes. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 364−366.
(9) Müller, W. R.; Szymanski, K.; Knop, J. V.; Trinajstić, N. Molecular topological index. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 160−163.
(10) Knop, J. V.; Müller, W. R.; Szymanski, K.; Trinajstić, N. On the determinant of the adjacency-plus-distance matrix as a topological index for characterizing alkanes. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 83−84.
(11) Klein, D. J.; Mihalić, Z.; Plavšić, D.; Trinajstić, N. Molecular topological index. A relation with the Wiener index. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 304−305.
(12) Gutman, I. Selected properties of the Schultz molecular topological index. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1087−1089.
(13) Mihalić, Z.; Nikolić, S.; Trinajstić, N. Comparative studies of the descriptors derived from the distance matrix. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 28−37.
(14) Plavšić, D.; Nikolić, S.; Trinajstić, N.; Mihalić, Z. On the Harary index for the characterization of chemical graphs. *J. Math. Chem.* **1993**, *12*, 235−250.
(15) Randić, M. Novel molecular descriptor for structure-property studies. *Chem. Phys. Lett.* **1993**, *211*, 478−483.

(16) Randić, M.; Guo, X.; Oxley, T.; Krishnapriyan, H. Wiener matrix: Source of novel graph invariants. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 709−716.
(17) Randić, M.; Guo, X.; Oxley, T.; Krishnapriyan, H.; Naylor, L. Wiener matrix invariants. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 361−367.
(18) Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Dömötör, G. The Szeged index and some of its analogies with the Wiener index. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547−550.
(19) Mihalić, Z.; Trinajstić, N. A graph-theoretical approach to structure-property relationships. *J. Chem. Educ.* **1992**, *69*, 701−712.
(20) Mihalić, Z.; Veljan, D.; Amić, D.; Nikolić, S.; Plavšić, D.; Trinajstić, N. The distance matrix in chemistry. *J. Math. Chem.* **1992**, *11*, 223−258.
(21) Elk, S. B. Canonical orderings of hexagons that; (1) nearly meet the intent of Patterson's nomenclature rules and (2) orient a molecule for purposes of assigning a IUPAC name. *Polycyclic Aromat. Comp.* **1990**, *1*, 109−121.
(22) Elk, S. B. A canonical ordering of polybenzenes and polymantanes using prime number factorization technique. *J. Math. Chem.* **1990**, *4*, 55−68.
(23) Elk, S. B. Graph theoretical algorithm to canonically name the isomers of the regular polyhedranes. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 14−22.
(24) Gutman, I.; Cyvin, S. J. Introduction to the Theory of Benzenoid Hydrocarbons; Springer-Verlag, Berlin, 1989.
(25) Sachs, H. Perfect matchings in hexagonal systems. *Combinatorica* **1984**, *4*, 89−99.
(26) Graham, R.; Pollak, H. On the addressing problem for loop switching. *Bell Syst. Tech. J.* **1971**, *50*, 2495−2519.
(27) Graham, R.; Pollak, H. On embedding graphs in squashed cubes. *Lect. Notes Math.* **1972**, *303*, 99−110.
(28) Djoković, D. Distance preserving subgraphs of hypercubes. *J. Combin. Theory Ser. B* **1973**, *14*, 263−267.
(29) Cyvin, S. J.; Brunvoll, J.; Cyvin, B. N. Theory of Coronoid Hydrocarbons; Springer-Verlag: Berlin, 1991.
(30) Imrich, W.; Klavžar, S. A simple *O*(*mn*) algorithm for recognizing Hamming graphs. *Bull. Inst. Comb. Appl.* **1993**, *9*, 45−56.
(31) Graham, R.; Winkler, P. M. On isometric embeddings of graphs. *Trans. Amer. Math. Soc.* **1985**, *288*, 527−536.
(32) Lukovits, I.; Gutman, I. Edge-decomposition of the Wiener number. *Commun. Math. Chem.* (MATCH) **1994**, *31*, 133−144.
(33) Mohar, B.; Pisanski, T. How to compute the Wiener index of a graph. *J. Math. Chem.* **1988**, *2*, 267−277.

CI9500142