match no. 28 pp. 209-218 1992

TWO CODES FOR HEXAGONAL SYSTEMS

Peter E. John
Technische Hochschule Ilmenau
Institut für Mathematik
PSF 327
O - 6300 Ilmenau

(received: May 1990)

Abstract

The nomenclature of structures plays an important role in chemistry. Therefore we give two simple codes for benzenoid hydrocarbons (BHs).

1 INTRODUCTION

The graph theory enables to develop diverse nomenclature systems for BHs. The following systems based on topological viewpoint have been reported in the literature: 2D- and 3D- codes (A.T. Balaban /1,2/, LA-sequence (I. Gutman /3,4/), the graph-center model (D. Bonchev, A.T. Balaban and M. Randić /4/), the boundary code methods (N. Trinajstić et al /6,7/ R. Tošić, R. Doroslovački and I. Gutman /8/, W.C. Herndon and A.J. Bruce /9/) and the compact name system (J. Cioslowski and A.M. Turek /10,11/). Let us recall the papers of S.B. Elk /12/ on nomenclature of polybenzenes. Note that Balabans 2D-code is identical to Gutmans LA-sequences.

In this paper we give two simple codes (boundary path code and rectangle block code) for a given BH. The boundary path code which results from the numbers of edges is very similar to the boundary code introduced by W.C. Herndon and A.J. Bruce and depending on the numbers of vertices. The rectangle block code is not only suitable for the computer input of graphs which are from interest, here. It is also from interest for the implementation of algorithms of such graphs which can be realisely simple by using the cartesian co-ordinate system.

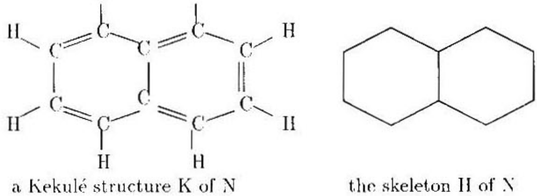
A hexagonal cell (briefly: h-cell) is a closed plane region bounded by a regular hexagon of unit side length.

A hexagonal system (HS) is a finite 2-connected plane graph in which every finite region is an h-cell.

A rectangle cell (briefly: r-cell) is a closed plane region bounded by a rectangle of side lengths 1 and 2.

A rectangle system (RS) is a finite 2-connected plane graphs in which every finite region is an r-cell.

The structural formula of a BH consists of a hexagonal system H spanned by the carbon atoms and some hanging edges such that each carbon atom which belongs to only one cell (to more than one cells) of H is connected to exactly one hydrogen atom (or to no hydrogen atom, respectively) (FIG. 1); let us call H the skeleton of the BH.



a Kekulé structure K of N — the skeleton H of N

naphthalene (N) — Figure 1

2 THE BOUNDARY PATH CODE

Let H be a HS. The boundary $B = B(H)$ of the exterior region is a circuit of even length; we shall call B "the boundary of H". Whenever an orientation is needed, then we shall suppose that B is oriented in the positive sense (with the exterior region on its right-hand side) (FIG. 2).

Let C be any cell (hexagonal region) of H and let $B(C)$ be its boundary ($B(C)$ is a circuit of length 6); put $B'(C) = B \cap B(C)$. C is called an interior cell if $B'(C) = \emptyset$ and C is called boundary cell if $B'(C) \neq \emptyset$.

If $B'(C)$ is a circuit then C is the only cell of H. Suppose that H has more than one cell and that C is a boundary cell of H: then $B'(C)$ is the disjoint union of one, two, or three paths wich we shall call b-paths (boundary-paths) of H. p is a b-path of H if there is a (boundary) cell C in H such that p is a maximal path (of H) in $B'(C)$. Clearly, B uniquely determinds a cyclic sequence of b-paths, $(p_1, p_2, ..., p_q)$, where p_j and p_{j+1} (the subscript is taken mod q) belong to the boundaries of two adjacent boundary cells and where

$$B = \bigcup_{i=1}^{q} p_i.$$

Let $l(p)$, $l(B)$ denote the length of p, or B, i.e. the number of edges contained in p, or B, respectively. Let $S = S(H)$ be the cyclic sequence $(l(p_1), l(p_2), ..., l(p_q))$ of H, then S is called the *boundary path code* of H (see also /13/).

Example: For the hexagonal system H_0 (FIG. 2) we find

$$S_0 = S(H_0) = (3, 4, 3, 1, 1, 5, 3, 2).$$

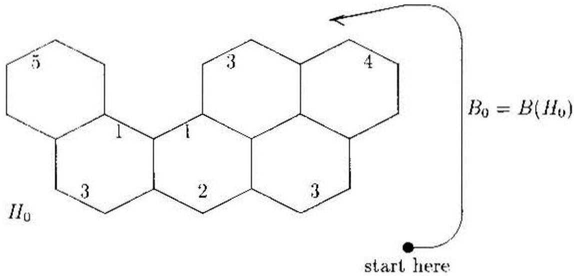


Figure 2

Remark 1: Let a_i $(i \in \{1,2,3,4,5\})$ denote the number of elements in S which have the value i. In our example (H_0) :

$$a_1 = 2, \quad a_2 = a_4 = a_5 = 1, \quad a_3 = 3.$$

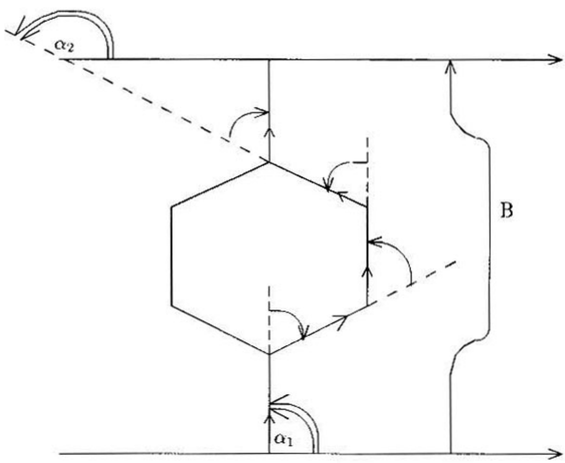
Clearly,

$$\sum_{i=1}^{5} i \cdot a_i = \sum_S l(p) = l(B).$$

Observation 1:

$$\sum_{i=1}^{5} (i-2) \cdot a_i = \sum_S (l(p) - 2) = 6.$$

Proof: When traversing a b-path p, the change of angle $d_p(\alpha)$ is $(l(p)-2) \cdot \frac{\pi}{3}$, see FIG. 3.



$d_p(\alpha) = \alpha_2 - \alpha_1 = (-\frac{\pi}{3}) + \frac{\pi}{3} + \frac{\pi}{3} = \frac{\pi}{3}$

Figure 3

When traversing B, the total change of angle is 2π, thus we have

$$\sum_S (l(p) - 2) \cdot \frac{\pi}{3} = 2\pi.$$

In our example:

$$\sum_{i=1}^{5} (i-2) \cdot a_i = (-1) \cdot 2 + 0 \cdot 1 + 1 \cdot 3 + 2 \cdot 1 + 3 \cdot 1 = 6.$$

3 THE RECTANGLE BLOCK CODE

Let H be a HS. Assume H to be drawn in the plane in such a way that some of its edges are horizontal, this can be done in six ways (where some of the drawings may be equal); cf. FIG 4.

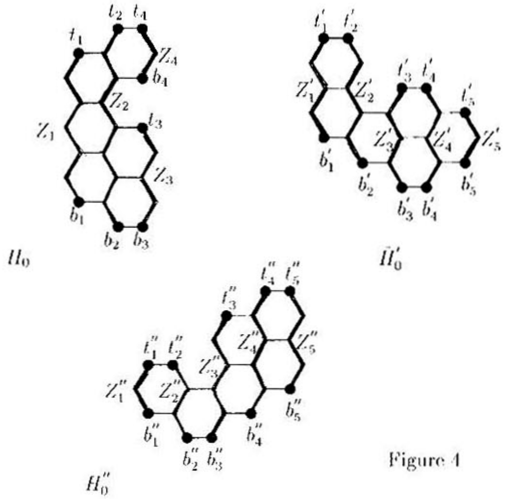


Figure 4

In such a drawing $\overline{H}$ of H we find a set "zigzag lines", say $Z_1, Z_2, ..., Z_k$, where Z_j means a maximal monotonic path (non-interrupted zigzag line) in $\overline{H}$ connecting a top point t_j with a bottom point b_j, as indicated in FIG. 4. (An example of an "interrupted" zigzag line, consisting of two different Z_j's is contained in $\overline{H}_0$). Among the six drawings of H we select one ($\overline{H}$) in which k, the number of maximal noninterrupted zigzag lines, is minimum; so, considering the example of FIG. 2, we select $\overline{H}_0$ (or the drawing obtained from $\overline{H}_0$ by turning it 180^o).

We stretch $\overline{H}$ in vertical direction such that a rectangle system is formed which we denote by $\overline{R}$, cf. FIG. 5.

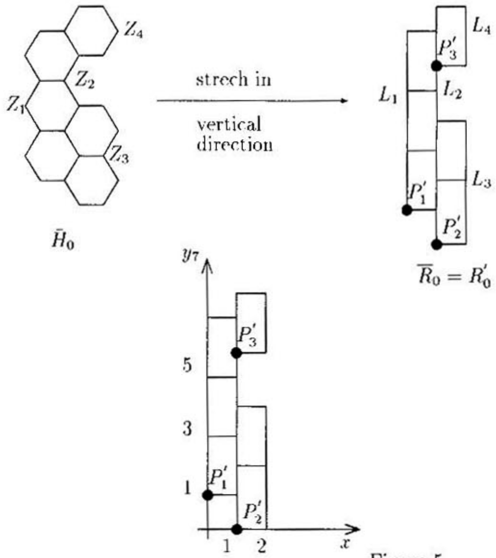


Figure 5

By this procedure the zigzag line Z_j of $\overline{H}$ is transformed in a (vertical) straight line segment L_j of $\overline{R}$. The (vertical) straight line segments L', L'' of $\overline{R}$ are called adjacent if there exists an (horizontal) edge e of $\overline{R}$ such that e is incident with a vertex v' of L' and a vertex v'' of L''. A vertical rectangle block (VRB) of $\overline{R}$ is a maximum RS which lies between adjacent straight line segments of $\overline{R}$. Let R' a VRB of R and P' be the left below vertex of R'

(see FIG. 5) which is called indexpoint of R'. The VRBs of $\overline{R}$ are running from bottom to top and from left to right indicated by $R'_1, R'_2, ..., R'_{k-1}$ and therfore the indexpoints of $\overline{R}$ are $P'_1, P'_2, ..., P'_{k-1}$.

In our example $\overline{R}_0$ has exactly three VRBs R'_1, R'_2, R'_3 with indexpoints P'_1, P'_2, P'_3 (see FIG. 5).

To the indexpoint P'_s, $s = 1, 2, ..., k-1$, and also for VRBs R'_s of $\overline{R}$ is assigned a tripel x_s, y_s, z_s in the following way: We place $\overline{R}$ in such a way in a cartesian x-y-co-ordinate system such that P'_s has the co-ordinates x_s, y_s where

(i) $x_s, y_s \geq 0$,

(ii) $min\{x_s, \quad s = 1, 2, ..., k-1\} = min\{y_s, \quad s = 1, 2, ..., k-1\} = 0$ and

(iii) z_s is the number of r-cells of R'_s.

The *rectangle block code* $\mathbf{P}' = \mathbf{P}'(H)$ of H is the set of points

$$P'_s(x_s, y_s, z_s), \quad s = 1, 2, ..., k-1,$$

$$\mathbf{P}' = \{P'_1, P'_2, ..., P'_{k-1}\}.$$

In our example, the assigment gives

$$\mathbf{P}'_0 = \mathbf{P}'(H_0) = \{P'_1 \quad (0; 1; 3); P'_2 \quad (1; 0; 2); P'_3 \quad (1; 6; 1)\}.$$

4 CONCLUDING REMARKS

4.1. It is worth mentioning that the codes described here can be applied - in a suitable extended form - to further classes of planar graphs, e.g., systems with holes (FIG. 6).

4.2. If we delete in $S = S(H)$ of H all 2's than we can find a reduced cyclic sequence $S' = S'(H) = (l(p'_1), l(p'_2), ..., l(p'_{q'}))$ of H

$(l(p_i') \in \{1;3;4;5\}, \quad i = 1,2,...,q')$. Note that S' characterises a class $\mathbf{H}'$ have the same sequence S'. For the HS depicted in FIG. 7 we obtain $S_0' = S'(H_0) = S'(H_1) = (3;4;3;1;1;5;3)$, therefore H_0 and H_1 belongs to the same class $\mathbf{H}_0'$ of HSs.

Acknowledgement: This research was supported by the Deutsche Forschungsgemeinschaft (Sa 509/1-1).

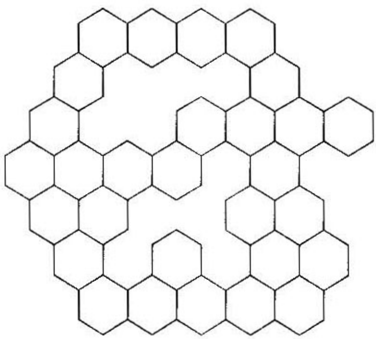
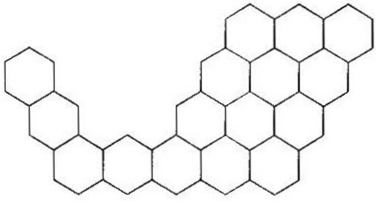


Figure 6



H_1

Figure 7

References

[1] Balaban, A.T., Tetrahedron, **25** (1969), 2949-2956

[2] Balaban, A.T., Rev. Roum. Chim., **15** (1970), 1243-1250 and **22** (1977), 45-47

[3] Gutman, I., Theor. Chim. Acta (Berlin), **45** (1977), 309-315

[4] Gutman, I., Z. Naturforsch. **37a** (1982), 69-73

[5] Bochev, D., A.T. Balaban and M. Randić, Int. J. Quant. Chem. **19** (1981) 61-82

[6] Knop,J.V., K. Szymanski, G. Jashari and N. Trinajstić, Croat. Chem. Acta **56** (3) (1983), 443-450

[7] Knop,J.V., K. Szymanski and N. Trinajstić, Match (Mülheim) **16** (1948), 103-118

[8] Tošić, R., R. Doroslovački and I. Gutman, Match (Mülheim), **19** (1986), 219-228

[9] Herndron, W.C. and A.J. Bruce, In: Graph Theory and Topology in Chemistry (R.B. King and D.H. Rouvray - Edts.), Studies in Physical and Theoretical Chemistry, No 51, Elsevier, Amsterdam - Oxford - New York - Tokyo, (1987), 491-513

[10] Cioslowski, J. and A.M. Turek, Computers & Chemistry **9** (1985), 301-308

[11] Cioslowski. J. and M.D. Wala, Match (Mülheim) **20** (1986), 209-227

[12] Elk. S.B., Match (Mülheim), **8** (1980), 121-158, **13** (1982), 239-269 and **17** (1985), 255-268

[13] John, P., Wiss. Z. d. Friedrich-Schiller-Universität Jena, **39** (2/3) (1990), 192-200