MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **72** (2014) 27-38

ISSN 0340 - 6253

How to Obtain The Number of Hexagons in a Benzenoid System from Its Boundary Edges Code

Jurij Kovič

University of Ljubljana, IMFM, Jadranska 19, 1000 Ljubljana, Slovenia, and Institute Andrej Marušič, University of Primorska, Muzejski trg 1, 6000 Koper, Slovenia
jurij.kovic@siol.net

(Received October 16, 2013)

Abstract. We calculate the number $h(\mathcal{B})$ of hexagons in a benzenoid system $\mathcal{B}$ from the shape of its boundary, represented by its "boundary edges code" $e(\mathcal{B})$ [*J. Mol. Struct. (Theochem)* **363** (1996) 237-247], which counts the boundary edges of $\mathcal{B}$ belonging to consecutive hexagons. We classify hexagons in a benzenoid system with regard to two parameters: the number of their adjacent hexagons and the number of their entries in the $e(\mathcal{B})$ and present a simple test for finding boundary hexagons of $\mathcal{B}$ which have at least two entries in $e(\mathcal{B})$.

1. Introduction

We assume the reader is familiar with the basic concepts of *benzenoid* and *coronoid hydrocarbons* (in this paper denoted $\mathcal{B}$ and $\mathcal{C}$), as well as with their mathematical model – planar polyhexes, composed from finite number of hexagons joined face to face, called *benzenoid systems* and *coronoid systems*, as explained e.g. by Gutman and Cyvin [9] . "Faces" in polyhexes represent rings of carbon atoms C, while the hydrogen atoms H are omitted; likewise, the positions of double bonds between carbon atoms in this simple mathematical model are ignored. Thus benzenoid systems (or shortly, *benzenoids*) are planar polyhexes with one component of boundary edges, and coronoid systems (or shortly, *coronoids*) are planar polyhexes with at least two components of boundary edges (two boundary edges e_1 and e_k are in the same component if there are boundary edges $e_2,\ldots,e_{k-1}$ such that any two edges e_{i+1} and e_i share a boundary vertex v_i).

A benzenoid $\mathcal{B}$ is uniquely determined by the shape of its boundary, represented by the *boundary edges code* (BEC) of Hansen, Lebbateaux and Zheng [10] (it is identical to the perimeter code *PC*-2 of Herndon and Bruce [11]). Also, Knop et al. [13,14] and

independently Doroslovački and Tošić [6] characterized a benzenoid system by associating to its perimeter a sequence of symbols of six letter alphabet {0,1,2,3,4,5} (see also [15]), corresponding to unit vectors in six possible directions of edges in a hexagonal net.

Remark. General hexagonal patches that are not benzenoids are not uniquely determined by their BEC; the first example of 25-hexagon "twins" with the same boundary was found by Guo, Hansen and Zheng [8], who proved also that the number of hexagons in such twins is the same, and that the boundary uniquely determines hexagonal patches with no more than 24 hexagons. Graver [7] shows that polyhexes having twins must cover at least one point three or more times when embedded in a plane. Brinkmann, Delgado-Friedrichs and von Nathusius [4] showed that the boundary encoding uniquely determines the number of faces in an (m,k)-patch (i.e. an embedded 2-connected planar graph with at least 3 vertices, where every bounded face is a k-gon, the vertex degree is m for inner vertices and at most k for vertices on the boundary) if (m,k) is not (6,3), (4,4) or (3,6).

This boundary edges code (BEC), denoted by $e(\mathfrak{B}) = (k_1, k_2,\ldots, k_r)$, counts the numbers k_i of edges in consecutive boundary hexagons (we travel around the boundary in the clockwise direction, starting at any hexagon, and having the interior of $\mathfrak{B}$ always at our right). BEC can be defined also for coronoids $\mathfrak{C}$. The entries k_i in the $e(\mathfrak{B})$ may be only numbers 1, 2, 3, 4, 5; the only exception to this rule is *benzen*, represented by a polyhex with $h(\mathfrak{B}) = 1$, whose BEC is $e(\mathfrak{B}) = (6)$. However, this code is not unique: even if we demand that the code be *standard* (i.e. the counting of boundary edges starts at the bottom left hexagon of a *standard drawing* of a benzenoid), there are benzenoids $\mathfrak{B}$ having 12 different BEC (corresponding to 12 different embeddings of $\mathfrak{B}$ into a hexagonal net), as shown by an example given in the Appendix of the paper [12]. It is shown in Figure 1; the reader may find it interesting to determine the other 11 standard codes of it.

Figure 1: A benzenoid with a boundary code $e(\mathfrak{B}) = (4,1,4,2,1,3,3,5,1,1,3,1,1,1,5,2)$.

In [12] it is shown also how the *symmetry scheme* of a benzenoid $\mathfrak{B}$ may be found from $e(\mathfrak{B})$, but only if we know also $h(\mathfrak{B})$, the number of hexagons of $\mathfrak{B}$; for example, the "face-

centered" and "vertex-centered" benzenoids with 3-fold symmetry satisfy the relations $h(\mathfrak{B}) \equiv 1 \pmod 3$ and $h(\mathfrak{B}) \equiv 1 \pmod 3$, respectively. The 14 possible symmetry schemes are just visual representations of the 14 different *symmetry groups* of benzenoids, as given in the classiffication of Gutman and Cyvin [9] which takes into account also the *point groups* of benzenoids, and consequently distinguishes between "horizontal" and "vertical" reflection lines; likewise it treats the cases in which the eventual center of rotational symmetries of $\mathfrak{B}$ is in the center of a face, or in the center of an edge, or in a vertex of a hexagonal net into which $\mathfrak{B}$ is embedded, as different. Also, an algorithm was proposed in [12] for obtaining $h(\mathfrak{B})$ from $e(\mathfrak{B})$, but it was accompanied with a remark that some technical details of it were somewhat complicated and would be given elsewhere. To keep this paper reasonably self-contained we repeat here the main idea of that algorithm: starting with $\mathfrak{B} = \mathfrak{B}_1$ and eliminating (properly chosen) boundary hexagons we produce a sequence of benzenoids $\mathfrak{B}_1, \mathfrak{B}_2, \ldots, \mathfrak{B}_{h(\mathfrak{B})}$. Then the number $h(\mathfrak{B})$ may be obtained from the recursive formula: $h(\mathfrak{B}_i) = h(\mathfrak{B}_{i+1}) + 1$, where $\mathfrak{B}_{i+1}$ is a benzenoid system obtained from $\mathfrak{B}_i$ by taking one hexagon away. Thus to obtain $h(\mathfrak{B}_i)$ it suffices to solve a "smaller" problem: to obtain $h(\mathfrak{B}_{i+1})$ from $e(\mathfrak{B}_{i+1})$. For the elimination of boundary hexagons with entries 5, 4 and 3 in the BEC (see Figure 2) simple formulas may be found for expressing $e(\mathfrak{B}_{i+1})$ in terms of $e(\mathfrak{B}_i)$. Repeating this recursive process we finally get a benzenoid system $\mathfrak{B}_{h(\mathfrak{B})}$ with only one hexagon. Then $h(\mathfrak{B})$ = (the number of these recursive steps) + 1. While the elimination of hexagons having entries 5 and 4 in the BEC is easy, the elimination of hexagons having entries 3 turned out to be a tricky problem: if we carelessly eliminate in $\mathfrak{B}_i$ a hexagon having two entries (3 and 1) in $e(\mathfrak{B}_i)$, then $\mathfrak{B}_{i+1}$ consists of two components, therefore $\mathfrak{B}_{i+1}$ is not a benzenoid system and the algorithm can not be applied to $\mathfrak{B}_{i+1}$ (see Figure 2 bottom right)! To explain in detail how this obstacle can be overcame is the purpose of this paper.

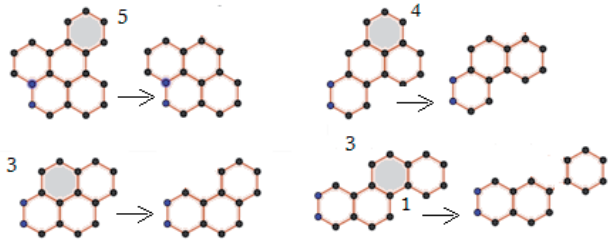


Figure 2: Examples of deleting (dark) hexagons with 5, 4, 3 and 3 + 1 boundary edges.

2. How to calculate the positions of boundary faces from $e(\mathfrak{B})$

Since an elimination of a hexagon with entries 3 and 1 in $e(\mathfrak{B}_i)$ is forbidden, it is clear what remains to be done to guarantee that the proposed algoritem would work:

Step 1) we have to prove that after eliminating hexagons with entries 5 and 4 there is always a boundary hexagon having an entry 3 in $e(\mathfrak{B}_i)$, and

Step 2) we have to construct a test telling us for each boundary hexagon of $\mathfrak{B}_i$ whether it has at least two entries in the BEC or not.

If we can do both 1) and 2), then we can also find a hexagon having only 3 consecutive edges at the boundary of $\mathfrak{B}_i$, since the set of such hexagons is nonempty and finite!

The key idea in carrying out the above plan is: we can use vectors to express (and calculate) positions of boundary vertices, edges and hexagonal »faces« of $\mathfrak{B}$. This plan can be carried out in various ways. But since the numbers n_k in $e(\mathfrak{B}) = (k_1,k_2,\ldots,k_r)$ correspond to boundary faces, the best way to do this is to focus on the vectors $\boldsymbol{f}_1,\ldots,\boldsymbol{f}_r$ with endpoints in the centers of the consecutive boundary hexagons $H_1,\ldots,H_r$. In order to to this well some preparations are needed.

The following lemmas are very easy to prove (therefore we omit the proofs):

Lemma 1. Let $\boldsymbol{a}_0$ and $\boldsymbol{a}_1$ be any two chosen "base" vectors starting at the center O of the same face in the hexagonal net (into which $\mathfrak{B}$ is embedded) and with the endpoints in an adjacent face such that the turn from $\mathbf{a}_0$ to $\mathbf{a}_1$ is 60° in the clockwise direction. Let $\boldsymbol{a}_2 = \boldsymbol{a}_1 - \boldsymbol{a}_0$, $\mathbf{a}_3 = -\mathbf{a}_0$, $\mathbf{a}_4 = -\mathbf{a}_1$, $\mathbf{a}_5 = \mathbf{a}_0 - \mathbf{a}_1$. Denote the axes of a coordinate system starting at O and pointing in the directions of $\boldsymbol{a}_0$ and $\boldsymbol{a}_1$ by x and y. Then these 6 vectors can be represented as follows: $\boldsymbol{a}_0 = (1,0)$, $\boldsymbol{a}_1 = (0,1)$, $\boldsymbol{a}_2 = (-1, 1)$, $\mathbf{a}_3 = (-1, 0)$, $\mathbf{a}_4 = (0, -1)$, $\mathbf{a}_5 = (1, -1)$.

Lemma 2. Any walk through the centers of a chain of adjacent hexagons $H_1, H_2, H_3,\ldots,H_r$ can be expressed as a sequence $\Delta(\mathfrak{B}) = (\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2, \boldsymbol{\Delta}_3,\ldots, \boldsymbol{\Delta}_r)$ of "difference vectors" $\boldsymbol{\Delta}_j = \boldsymbol{f}_{j+1} - \boldsymbol{f}_j$, where $\boldsymbol{f}_j$ is the vector starting at the coordinate origin O and ending in the center of H_j, $j = 1,2,\ldots, r$. Hence this holds also in the case that $H_1, H_2, H_3,\ldots,H_r$ are consequent boundary faces of a benzenoid $\mathfrak{B}$ with entries $k_1, k_2, k_3,\ldots, k_r$ in the boundary code $e(\mathfrak{B}) = (k_1, k_2,\ldots, k_r)$.

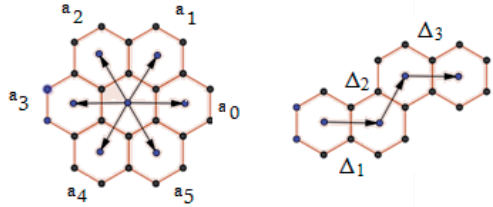


Figure 3: (left) the vectors $\boldsymbol{a}_0 = (1,0)$, $\boldsymbol{a}_1 = (0,1)$, $\boldsymbol{a}_2 = (-1, 1)$, $\mathbf{a}_3 = (-1, 0)$, $\mathbf{a}_4 = (0, -1)$, $\mathbf{a}_5 = (1, -1)$, (right) a walk expressed with difference vectors $\boldsymbol{\Delta}_j$.

Lemma 3. Since $\boldsymbol{f}_{j+1} = \boldsymbol{f}_j + \boldsymbol{\Delta}_j = \boldsymbol{f}_1 + \boldsymbol{\Delta}_1 + \boldsymbol{\Delta}_2 + \boldsymbol{\Delta}_3 + \ldots + \boldsymbol{\Delta}_j$, such a walk is closed (i.e. $\boldsymbol{f}_1 = \boldsymbol{f}_j$) if and only if $\boldsymbol{\Delta}_1 + \boldsymbol{\Delta}_2 + \boldsymbol{\Delta}_3 + \ldots + \boldsymbol{\Delta}_j = 0$.

Lemma 4. Consequently, $H_1 = H_j$ is a boundary hexagon of $\mathfrak{B}$ having at least two entries in $\boldsymbol{e}(\mathfrak{B}) = (k_1, k_2, \ldots, k_r)$ if there is such an $j < r$ for which $\boldsymbol{f}_1 = \boldsymbol{f}_j$, hence $\boldsymbol{\Delta}_1 + \boldsymbol{\Delta}_2 + \boldsymbol{\Delta}_3 + \ldots + \boldsymbol{\Delta}_j = 0$. Since each difference vector $\boldsymbol{\Delta}_j = (x_j, y_j)$ can be expressed as one of the vectors $\boldsymbol{a_0} = (1,0)$, $\boldsymbol{a_1} = (0,1)$, $\boldsymbol{a_2} = (-1, 1)$, $\mathbf{a_3} = (-1, 0)$, $\mathbf{a_4} = (0, -1)$, $\mathbf{a_5} = (1, -1)$, this is equivalent to the conditions $x_1 + x_2 + x_3 + \ldots + x_j = 0$ and $y_1 + y_2 + y_3 + \ldots + y_j = 0$, that are very easy to check.

Lemma 5. So we can write $\boldsymbol{\Delta}_j = \boldsymbol{a}_{d(j)}$, where $d(j)$ denotes one of the six possible directions 0, 1, 2, 3, 4, or 5 of difference vectors.

But we can actually calculate $\boldsymbol{f}_j$ and $\boldsymbol{\Delta}_j$ from $e(\mathfrak{B})$!

Theorem 1. $d(j + 1) = d(j) + k_{j+1} - 2 \pmod 6$.

Proof. The numbers $d(j + 1) = d(j)$ denote the directions of difference vectors $\boldsymbol{\Delta}_{j+1}$ and $\boldsymbol{\Delta}_j$.

It is easy to see that if $k_{j+1} = 2$ then $d(j + 1) = d(j)$ (see Figure 4). Likewise:

if $k_{j+1} = 1$ then $d(j + 1) = d(j) - 1 \pmod 6$,

if $k_{j+1} = 3$ then $d(j + 1) = d(j) + 1 \pmod 6$,

if $k_{j+1} = 4$ then $d(j + 1) = d(j) + 2 \pmod 6$,

if $k_{j+1} = 5$ then $d(j + 1) = d(j) + 3 \pmod 6$. □

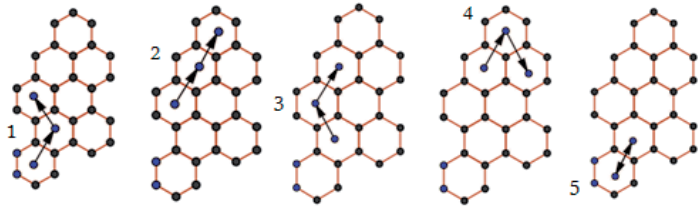


Figure 4: Proof of the formula $d(j + 1) = d(j) + k_{j+1} - 2 \pmod 6$.

Theorem 2. Embed $\mathfrak{B}$ in the hexagonal net so that $\boldsymbol{f}_1 = 0$ (i.e. the coordinate origin is placed in the center of the face H_1) and $\boldsymbol{f}_2 = \boldsymbol{a}_0$. Then $\boldsymbol{f}_j$ may be calculated recursively from $e(\mathfrak{B})$ by the formulas:

$\boldsymbol{f}_1 = 0$, $\quad d(1) = 0$,

$\boldsymbol{f}_2 = \boldsymbol{f}_1 + \boldsymbol{a}_{d(1)}$, $\quad d(2) = d(1) + k_2 - 2 \pmod 6$

…..

$\boldsymbol{f}_{j+1} = \boldsymbol{f}_j + \boldsymbol{a}_{d(j)}$, $\quad d(j+1) = d(j) + k_{j+1} - 2 \pmod 6$

.....

$\boldsymbol{f}_r = \boldsymbol{f}_{r-1} + \boldsymbol{a}_{d(r-1)}$, $\quad d(r+1) = d(r) + k_1 - 2 \pmod 6$.

$\boldsymbol{f}_1 = \boldsymbol{f}_r + \boldsymbol{a}_{d(r)}$ (this is a test if $e(\mathfrak{B})$ really defines a benzenoid!).

Proof. This follows immediately from Theorem 1.□

The formulas of Theorem 1 may be better understood if reinterpreted in the context of the "turtle geometry" of Abelson and diSessa [1]. In this geometry a "turtle" moves along a piecewise linear path. The sum of its "local turnings" (left or right for an angle α_i) is called the "total turning". Hence the formulas of Theorem 1 simply say:

Theorem 3. The local turnings $d(j + 1) - d(j)$ of a turtle crawling along the path through the centers $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_r$ of boundary faces $H_1, H_2, \ldots, H_r$ of a benzenoid $\mathfrak{B}$ can be expressed with the entries k_{j+1} in $e(\mathfrak{B}) = (k_1, k_2, \ldots, k_r)$ as follows: $d(j + 1) - d(j) = k_{j+1} - 2 \pmod 6$.

Using this geometry it is easy to prove the following deep theorem ([1], p. 24):

Theorem 4. The total turning $t(P) = \alpha_1 + \ldots + \alpha_r$ (a turtle makes) along a closed path P is an integer multiple of 360º. The path is simple if and only if $t(P) = \pm 360º$.

Proof. This follows from the "deformation theorem" ([1] , pp. 182-188) which says that the total turning along a closed curve is not changed under continuous deformations. Since a simple closed curve can be deformed into a square, and since a square path obviously has the total turning $t(P) = 4$ right angles $= \pm 360^o$, the theorem follows. □

Suppose a turtle travels with a constant (scalar value of) velocity along the polygonal line P between adjacent boundary faces. Then the difference vectors $\mathbf{\Delta}_j$ may be interpreted as velocities (vectors!), and the changes in them $\mathbf{\Delta}_{j+1} - \mathbf{\Delta}_j$ as accelerations (expressed with local turnings $d(j+1) - d(j)$). Now the Theorem 4 can be reformulated and applied to benzenoids:

Definition 1. Let n_i denote the number of entries i in $e(\mathfrak{B})$.

Theorem 5. $(k_1 - 2) + \ldots + (k_r - 2) = 6 + 6n_5$ for any benzenoid $\mathfrak{B}$.

Proof. Suppose first the closed path through the centers $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_r$ of consecutive boundary faces $H_1, H_2, \ldots, H_r$ of a benzenoid $\mathfrak{B}$ is simple. Then $t(P) = +360^o$, hence $d(1) + \ldots + d(r) = 6$. If this path is not simple, it can happen only if some $k_{j+1} = 5$. But then the parts of this path traversed from $\boldsymbol{f}_j$ to $\boldsymbol{f}_{j+1}$ and back to $\boldsymbol{f}_j$ contributes another 6 to the $t(P)$. □

Theorem 6. If $n_4 = n_5 = n_6 = 0$, then $n_3 = n_1 + 6$, hence $n_3 \geq 2$.

Proof. The entries 1,2,3 contribute – 1,0,1, respectively, to the sum $(k_1 - 2) + \ldots + k_r - 2) = 6$.

Remark. This theorem completes the step 1) of our plan. Now we procced to the step 2). The test if a boundary hexagon H has at least two entries in $e(\mathfrak{B})$ is given in the following theorem:

Theorem 7. Choose a boundary hexagon H, name it H_1, and calculate $\boldsymbol{f}_j$ from $e(\mathfrak{B}) = (k_1, k_2, \ldots, k_r)$ by the formulas of Theorem 2. If there is a $j < r$ such that $\boldsymbol{f}_1 = \boldsymbol{f}_j$ then H has at least two entries in $e(\mathfrak{B})$.

Remark. Theorem 7 gurantees that the step 2) of our plan can be executed indeed: we just have to check which of the hexagons with entries 3 (existence of at least one such hexagon is guaranteed by Theorem 6) has only one entry in the $e(\mathfrak{B})$, if all the entries of it are 1, 2 or 3.

Now we just put all the pieces together:

Theorem 8. To calculate $h(\mathfrak{B})$ from $e(\mathfrak{B})$ we may use the following procedure:

i) embed $\mathfrak{B}$ into a hexagonal net so that $f_1 = 0$ and $f_2 = a_0$ (as in Theorem 2) and calculate $f_3,\ldots,f_r$ by recursive formulas of Theorem 2;

ii) produce a sequence of benzenoids $\mathfrak{B} = \mathfrak{B}_1, \mathfrak{B}_2,\ldots, \mathfrak{B}_{h(\mathfrak{B})}$, where $\mathfrak{B}_{j+1}$ is obtained from $\mathfrak{B}_j$ by taking one properly chosen boundary hexagon away as follows:

a) if there is a hexagon H_i in $\mathfrak{B}_j$ such that $k_i = 5$ in $e(\mathfrak{B}_j) = (\ldots, a, 5, b, \ldots)$, then eliminate H_i to obtain $\mathfrak{B}_{j+1}$. Then $e(\mathfrak{B}_{j+1}) = (\ldots, a + b + 1, \ldots)$;

b) if all the entries in $e(\mathfrak{B}_j)$ are less than 5 and if there is a hexagon H_i in $\mathfrak{B}_j$ such that $k_i = 4$ in $e(\mathfrak{B}_j) = (\ldots, a, 4, b, \ldots)$, eliminate H_i to obtain $\mathfrak{B}_{j+1}$. Then $e(\mathfrak{B}_{j+1}) = (\ldots, a + 1, 1 + b, \ldots)$;

c) if all the entries in in $e(\mathfrak{B}_j)$ are less than 4 there is (by Theorems 7 and 8) a hexagon H_i in $\mathfrak{B}_j$ such that $k_i = 3$ and having no other entries in $e(\mathfrak{B}_j) = (\ldots, a, 3, b, \ldots)$. Eliminate H_i to obtain $\mathfrak{B}_{j+1}$. Then $e(\mathfrak{B}_{j+1}) = (\ldots, a + 1, 1, 1 + b, \ldots)$;

iii) Then the number of hexagons of $\mathfrak{B}$ is $h(\mathfrak{B})$ = (the number of these recursive steps) + 1.

Proof. To check the formulas for $e(\mathfrak{B}_{j+1})$ see Figure 2. □

Definition 2. Let $e_b(\mathfrak{B})$ be the number of boundary edges in a benzenoid $\mathfrak{B}$.

Theorem 9. If we delete from $\mathfrak{B}_j$ a hexagon, having entry 5 in $e(\mathfrak{B}_j)$, then $e_b(\mathfrak{B}_{j+1}) = e_b(\mathfrak{B}_{j+1}) - 4$. If we delete from $\mathfrak{B}_j$ a hexagon, having entry 4 in $e(\mathfrak{B}_j)$, then $e_b(\mathfrak{B}_{j+1}) = e_b(\mathfrak{B}_{j+1}) - 2$. If we delete from $\mathfrak{B}_j$ a hexagon, having only entry 3 in $e(\mathfrak{B}_j)$, then $e_b(\mathfrak{B}_{j+1}) = e_b(\mathfrak{B}_{j+1})$.

Theorem 10. Each benzenoid $\mathfrak{B}$ can be constructed by a recursive process, producing a sequence of benzenoids $\mathfrak{B}_1, \mathfrak{B}_2,\ldots, \mathfrak{B}_m = \mathfrak{B}$, starting with a single hexagon $\mathfrak{B}_1$ and adding hexagons to $\mathfrak{B}_i$ having only one entry 5, 4 or 3 in $e(\mathfrak{B}_{i+1})$. Let a, b, c be the numbers of deleted hexagons of type a), b) and c) from Theorem 10. Then $h(\mathfrak{B}) = a + b + c$ and $e_b(\mathfrak{B}) = 6 + 2b + 4c$. If $a = 0$ then $b = e_b(\mathfrak{B})/2 + 2h(\mathfrak{B}) - 3$ and $c = b(\mathfrak{B})/2 - h(\mathfrak{B}) - 3$.

Proof. We just reverse the process of Theorem 10. Obviously $h(\mathfrak{B}) = a + b + c$. Since $e_b(\mathfrak{B}_1) = 6$, and using formulas of Theorem 11, we have $e_b(\mathfrak{B}) = 6 + 2b + 4c$. The rest are just easy calculations. □

3. Summary and generalizations

We have given the missing details in the algorithm, proposed in [12], for counting the number of faces $h(\mathfrak{B})$ in a benzenoid $\mathfrak{B}$ from its boundary edges code $e(\mathfrak{B})$. We can use the same method for the counting of faces of planar pentagonal systems, as well as for the »spherical pentagonal systems«, obtained from the dodecahedron by deleting some of its faces (studied e.g. in Deza, Fowler, Grishukin [5]).

Definition 3. Let $\mathfrak{p}$ be a planar molecule, composed of regular pentagons connected face to face. For pentagonal systems without holes and overlapping faces we define a boundary edges code $e(\mathfrak{p})$ as follows: the entries k_i of $e(\mathfrak{p}) = (k_1, k_2,\ldots, k_r)$ count the numbers k_i of edges in consecutive boundary pentagons (we travel around the boundary in the clockwise direction, starting at any pentagon, and having the interior of $\mathfrak{p}$ always at our right). The entries k_i in the $e(\mathfrak{p})$ may be only numbers 1, 2, 3, 4; the only exception to this rule is a single pentagon, whose boundary edges code is $e(\mathfrak{p}) = (5)$.

Figure 5: A pentagonal system $\mathfrak{p}$ with a boundary edges code $e(\mathfrak{p}) = (4,2,4,3,1)$.

Theorem 11. Let $\mathfrak{p}$ be a planar pentagonal system without holes and overlapping faces. To calculate the number of faces $f(\mathfrak{p})$ from $e(\mathfrak{p})$ we may use the following procedure:

Produce a sequence of pentagonal systems $\mathfrak{p} = \mathfrak{p}_1, \mathfrak{p}_2,\ldots, \mathfrak{p}_{f(\mathfrak{p})}$, where $\mathfrak{p}_{j+1}$ is obtained from $\mathfrak{p}_j$ by taking one properly chosen boundary pentagon away as follows:

i) if there is a hexagon P_i in $\mathfrak{p}_j$ such that $k_i = 4$ in $e(\mathfrak{p}_j) = (\ldots, a, 4, b, \ldots)$, then eliminate H_i to obtain $\mathfrak{p}_{j+1}$. Then $e(\mathfrak{p}_{j+1}) = (\ldots, a, b + 1, \ldots)$;

ii) if all the entries in $e(\mathfrak{p}_j)$ are less than 4 and if there is a pentagon P_i in $\mathfrak{p}_j$ such that $k_i = 3$ in $e(\mathfrak{B}_j) = (\ldots, a, 3, b, \ldots)$, eliminate P_i to obtain $\mathfrak{p}_{j+1}$. Then $e(\mathfrak{p}_{j+1}) = (\ldots, a, 1, 1 + b, \ldots)$;

iii) if all the entries in $e(\mathfrak{p}_j)$ are 1 or 2, eliminate a pentagon P_i in $\mathfrak{p}_j$ having only the entry $k_i = 2$ in $e(\mathfrak{B}_j) = (\ldots, a, 2, b, \ldots)$, to obtain $\mathfrak{p}_{j+1}$. Then $e(\mathfrak{p}_{j+1}) = (\ldots, a + 1, 1, 1 + b, \ldots)$.

Then the number of hexagons of $\mathfrak{P}$ is $h(\mathfrak{P})$ = (the number of these recursive steps) + 1.

Proof. To check the formulas for $e(\mathfrak{P}_{j+1})$ in the planar case see Figure 6. In steps ii) and iii) we obtain a structure, that is no longer face-to face connected; nevertheless, a boundary edges code can be defined also for such structures (just as in Definition 3)!

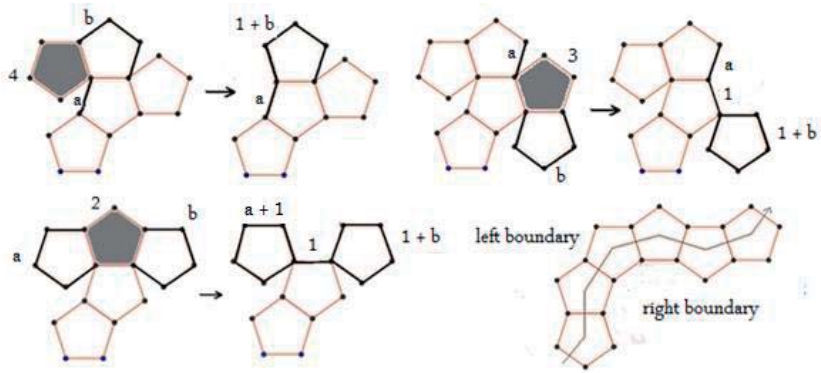


Figure 6: Eliminations (…,*a*,4,*b*,…) → (…,*a*,1+*b*,…) and (…, *a*,3,*b*,…) → (…,*a*,1,1+*b*,…) and (…,*a*,2,*b*,…) → (…,*a*+1, 1, 1+*b*,…) and a chain of pentagons with entries 2 and 1.

To see that the steps i), ii) and iii) suffice to eliminate all pentagons except the last one, suppose there are only entries 1 and 2 in $e(\mathfrak{P}_j)$, and that all the boundary pentagons have both entries 2 and 1 in $e(\mathfrak{P}_j)$ (it is obviously not possible that all the entries were 1). But this leads into a contradiction, since in that case $\mathfrak{P}_j$ would have two components of boundary edges (this is true for each pentagon, consequently the »left« boundary would never meet the »right« one – see Figure 8), therefore $\mathfrak{P}_j$ would have holes; but since $\mathfrak{P}_1$ is without holes, then all the $\mathfrak{P}_j$ are without holes, too, since the steps i) and ii) can not produce them.

In the case of »dodecahedral pentagonal system without holes« $\mathfrak{P}$, the transformation formulas $e(\mathfrak{B}_j) \rightarrow e(\mathfrak{B}_{j+1})$ of the steps i), ii) and iii) are the same, but in the steps ii) and iii) we do not have a situation of the boundary intersecting itself as in the planar case.□

Appendix: Types of faces in benzenoids and coronoids

Balaban and Nenitezcu [2,3] classified the types of faces in benzenoids $\mathfrak{B}$ and coronoids $\mathfrak{C}$ by the number of of their adjacent hexagons. In fact, they forgot the third one of the four-contact cases (see Table 1). We can refine that classification as follows:

Theorem 12. There are 11 types of boundary faces of $\mathfrak{B}$ or $\mathfrak{C}$. They may have 1, 2 or 3 entries in the BEC. The type of any boundary hexagon H in $\mathfrak{B}$ can be deduced from $e(\mathfrak{B})$.

Proof. Different types of boundary hexagons all have different sets of entries in $e(\mathfrak{B})$ except the two 1+1 types (see Table 1), which can be distinguished using the information hidden in $e(\mathfrak{B})$ as follows: let $H_1 = H_j$ be a (1 + 1)-type hexagon. Then $d(j) = -d(1)$ or $d(j) = d(1) - 2 \pmod 6$. And we know from Theorem 2 that $d(1),\ldots, d(r)$ can be calculated from $e(\mathfrak{B})$. □

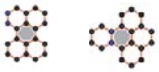
	0	1	2	3
One-contact		5 		
Two-contact		4 	3+1 	
Three-contact		3 	2+1 	
Four-contact		2 	1+1 	
Five-contact		1 		
Six-contact				

Table 1: Classification of hexagons by the number of their contacts (1, 2, 3, 4, 5 or 6) with other hexagons and by the number of components (0, 1, 2 or 3) of their edges at the boundary.

Acknowledgment: The author acknowledges partial support from ARRS grants P1-0294.

References

[1] H. Abelson, A. di Sessa, *Turtle Geometry: The Computer as a Medium for Exploring Mathematics*, MIT Press, Cambridge, 1981.

[2] A. T. Balaban, Chemical graphs. VII. Proposed nomenclature of branched cata-condensed benzenoid polycyclic hydrocarbons, *Tetrahedron* **25** (1969) 2949–2956.

[3] A. T. Balaban, C. D. Nenitzescu, Dehydrogenating condensations of aromatics (Scholl and related reactions), in: G. A. Olah (Ed.), *Friedel–Crafts and Related Reactions*, Wiley-Interscience, New York, 1964, pp. 979–1047.

[4] G. Brinkmann, O. Delgado–Friedrichs, U. von Nathusius, Numbers of faces and boundary encodings of patches, in: S. Fajtlowicz, P. W. Fowler, P. Hansen, M. F. Janowitz, F. S. Roberts (Eds.), *Graphs and Discovery*, Am. Math. Soc., Providence, 2005, pp. 27–38.

[5] M. Deza, P. W. Fowler, V. Grishukin, Allowed boundary sequences for fused polycyclic paths and related algorithmic problems, *J. Chem. Inf. Comput. Sci*. **41** (2001) 300–308.

[6] R. Doroslovački, R. Tošić, A characterization of hexagonal systems, *Rev. Res. Fac. Sci. Univ. Novi Sad Ser. Math*. **14** (1984) 201–209.

[7] J. E. Graver, The (*m,k*)-patch boundary code problem, *MATCH Commun. Math. Comput. Chem.* **48** (2003) 189–196.

[8] X. Guo, P. Hansen, M. Zheng, Boundary uniqueness of fusenes, *Discr. Appl. Math.* **118** (2002) 209–222.

[9] I. Gutman, S. V. Cyvin, *Introduction to the Theory of Benzenoid Hydrocarbons*, Springer, Berlin, 1989.

[10] P. Hansen, C. Lebatteux, M. Zheng, The boundary–edges code for polyhexes, *J. Mol. Struct. (Theochem)* **363** (1996) 237–247.

[11] W. C. Herndon, A. J. Bruce, Perimeter code for benzenoid aromatic hydrocarbons, in: R. B. King, D. H. Rouvray (Eds.), *Graph Theory and Topology in Chemistry*, Elsevier, Amsterdam, 1987, pp. 491–513.

[12] J. Kovič, T. Pisanski, A. T. Balaban, P. W. Fowler, On symmetries of benzenoid systems, *MATCH Commun. Math. Comput. Chem*., in press.

[13] J. V. Knop, K. Szymanski, Ž. Jeričević, N. Trinajstić, Computer enumeration and generation of benzenoid hydrocarbons and identification of bay regions, *J. Comput. Chem.* **4** (1983) 23–32.

[14] J. V. Knop, K. Szymanski, G. Jashari, N. Trinajstić, The connection between the adjacency matrix and the boundary code of benzenoid hydrocarbons, *Croat. Chem. Acta* **56** (1983) 443–450.

[15] R. Tošić, R. Doroslovački, I. Gutman, Topological properties of benzenoid systems, XXXVIII, The boundary code, *MATCH Commun. Math. Comput. Chem*. **19** (1986) 219–228.