match no. 16 pp. 103-118 1984

A NOTE ON TREES WITH POLYHEXAGONAL SUPERVERTICES

J.V.Knop, K.Szymanski, and **N.Trinajstić** *

Computer Centre, The University of Düsseldorf,
4000 Düsseldorf, Federal Republic of Germany

(Received: November 1983)

ABSTRACT

Trees with polyhexagonal supervertices (i.e. supertrees) are defined. An algorithm is developed which can be used to represent such trees uniquely.

*Permanent address: The Rugjer Bošković Institute, P.O.B. 1016, 41001 Zagreb, Croatia, Yugoslavia

In references [1,2] we have developed a method for the unique representation of trees. In [2-4], we published a method for the unique representation of polyhexes, and in [2,5,6] this method was extended to substituted polyhexes. In this paper, we wish to show that both methods are related.

We will first briefly describe the basis of both methods.

REPRESENTATION OF PLANAR POLYHEXES

A planar polyhex [7] is a configuration constructed in the plane by assembling **h** regular hexagons in such a way that

(i) two hexagons have exactly one common edge or are disjoint, and

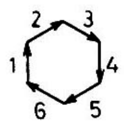
(ii) the covered area in the plane is simply connected [8] .

The number of hexagons **h** , making up a given polyhex, is the degree of the polyhex. To represent a polyhex numerically, we use digits assigned to vectors running along the boundary edges of the structure, and construct a sequence which follows the direction of these vectors. The boundary of a polyhex in the plane is a cycle in a graph theoretical sense, because the covered area is simply connected [9]. The interior of a polyhex [9,10] is reconstructable [11] because of the uniformity of the interior as far as the boundary is known, and so we can therefore represent a polyhex only by its boundary, to which we can ascribe a numerical description.

Considering the geometry of a single hexagon we can denote each edge by a vector in the plane, and each of the six vectors by a digit. The orientation of the vectors is arbitrarily taken to be clockwise. Similarly, the digit 1 is arbitrarily assigned to the left vertical vector of the hexagon. This is shown in Fig.1 .

Fig. 1

Vector description of a hexagon and the labelling of its edges



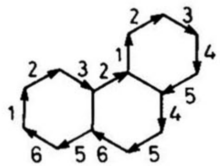
1

There are also other possibilities open for assigning vectors to a hexagon and for labelling its edges [12] .

In Fig. 2 we show how the labelling of the edges of a hexagon is extended to polyhexes.

Fig. 2

Vector description of a polyhex and the labelling of its edges



2

The boundary of the polyhex may be labelled by different sequences of numbers depending on which edge starts the sequence. We choose that which represents the lexicographic maximum as the unique boundary code of the polyhex. The boundary codes for the structures in Fig.1 and Fig.2 are given below :

1 : (6 1 2 3 4 5)

2 : (6 5 6 1 2 3 2 1 2 3 4 5 4 5)

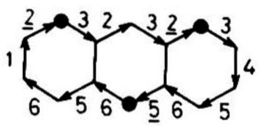
An important point to note is that the polyhex can be reconstructed from the boundary code. A notation closely related to that of the benzene periphery, and generalized to the polyhexes, has been used by Balaban [13] in his work on the configuration of annulenes.

In several publications we have shown how all geometrically planar polyhexes can be generated and ordered [2,3,4,14].

In the case of substituted polyhexes, the positions of substitution are denoted by black dots, and the edges leading to black dots are underlined in the boundary sequences. A **k** - substituted polyhex is a polyhex with **k** black dots (positions of substitution). An example of a substituted polyhex is shown in Fig.3 .

Fig. 3

Vector representation of a substituted polyhex and the labelling of its edges. Positions of substitution are denoted by black dots



3

The boundary sequence of the structure in Fig.3 is given below :

3 : (6 5 6 5 1 2 3 2 3 2 3 4 5)

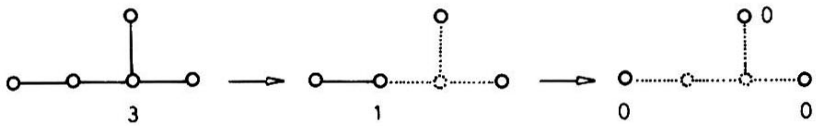
There are also other ways open for labelling the edges of substituted polyhexes [2,5,6].

REPRESENTATION OF TREES

A tree is a connected graph without cycles [15]. Trees with **N** vertices can be represented numerically by N-tuples of non-negative integers less than N [1]. The N-tuples of trees can be produced by mapping the trees onto N-tuples of non-negative integers by induction (the trivial tree with one vertex is represented by 1-tuple). In order to simplify the discussion we introduce a term called the starting vertex , which represents a vertex of a tree at which we start the N-tuple. Thus a given tree with $N>1$ vertices and **M** edges incident to the starting edge produces M subtrees obtained by removing the starting vertex and all incident edges. The subtrees with $L_1, L_2, \ldots, L_M$ vertices (where $L_1 + L_2 + \ldots + L_M = N-1$) are, by induction, provided by L subtuples. We concatenate the 1-subtuple (M) and these N-1 subtuples, and get a tuple of $1+L_1+L_2+ \ldots +L_M =N$ components which we define to be representative for the tree. We note at this point of the discussion that each vertex of the tree is mapped onto one component of the N-tuple and in the case of the starting vertex this component is equal to the valency of the starting vertex, and in all other cases is equal to the initial valency decreased by one. Each vertex of a tree may be used as a starting point of the N-tuple. This means that the tree may have several different N-tuples. In order to select a unique N-tuple representation of a given tree we again use the concept of the lexicographic maximum. This corresponds to the highest possible N-tuple representation of a tree. As an illustrative example we show in Fig.4 the construction of the highest lexicographic N-tuple for 2-methylbutane represented by a branched tree with 5 vertices.

Fig. 4

The construction of the highest lexicographic N-tuple for a tree representing 2-methylbutane. The numbers represent the valencies of the vertices. The N-tuple starts at the vertex with the highest valency



4

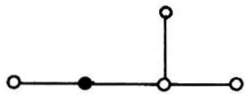
The N-tuple for the tree in Fig.4 is given by:

4 : (3 1 0 0 0)

The same approach may be used for rooted trees. A rooted tree is a tree in which one or more vertices can be distinguished from others [16]. An example of a rooted tree is given in Fig.5 .

Fig. 5

Example of a rooted tree. The root-vertex is in black



5

The starting vertex of the N-tuple representing a rooted tree is the root-vertex. Thus, the N-tuple for the tree in Fig.5 is given by:

5 : (2 2 0 0 0)

An important theorem which states that two non-isomorphic (rooted) trees cannot have the same N-tuple was proved in [2]. In [1] we illustrated how all trees can be accurately generated and ordered by the above approach.

TREES WITH POLYHEXAGONAL SUPERVERTICES

We can now combine both methods by introducing the concept of a supervertex. We will consider graphs that are obtained by replacing a vertex of valency **k** by a k-substituted polyhex. We call such a vertex a polyhexagonal supervertex. In N-tuple representing a tree with polyhexagonal supervertices the component corresponding to a supervertex is replaced by the boundary representation of the substituted polyhex in such a way that each marked edge of the polyhex points to an edge of the tree adjacent to a vertex of the polyhex.

The following algorithm leads to a unique representation of trees with the polyhexagonal supervertices:

(1) Compute the N-tuple representation of a tree

(2) If the starting vertex is replaced by a polyhex, rotate the boundary sequence of the polyhex until the edge of the polyhex, which leads to the vertex of the tree represented by the next component of the N-tuple representation of the tree, is the last edge in the boundary code. Now concatenate the subtrees in the order given by the boundary sequence of the polyhex.

(3) If any vertex of a tree, other than the starting vertex, is substituted by a polyhex, rotate the boundary sequence of the polyhex in such a way that the edge of the polyhex, which leads to the vertex of the tree adjacent to the subtree of which the considered vertex is the starting vertex, is the last edge in the sequence. Now concatenate the subtrees in the

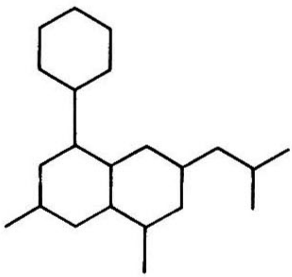
order given by the boundary code of the polyhex.

Since all trees and all geometrically planar substitututed polyhexes may be generated, we can therefore generate all trees with the polyhexagonal supervertices (or polyhexagonal supertrees).

The following example will illustrate the algorithm. The structure in Fig.6 may be considered as a polyhexagonal supertree.

Fig. 6

An example of a polyhexagonal supertree

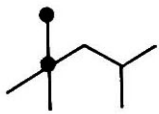


6

The corresponding tree with polyhexagonal supervertices is shown in Fig.7 .

Fig. 7

The tree with polyhexagonal supervertices (denoted by black dots) corresponding to structure 6



7

The N-tuple representation of structure 7 is as follows:

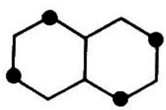
7 : (4 1 2 0 0 0 0 0)

The substituted polyhexes that are replaced by supervertices are given in Fig.8 .

Fig. 8

The substituted polyhexes in structure 6 that are replaced by polyhexagonal supervertices to yield structure 7. The positions of substitution are denoted by black dots

8



9

The boundary sequence representing the structure **8** and **9** are shown below (edges leading for the positions of substitution, i.e. black dots, are underlined in the sequence) :

8 : $(6\ 1\ 2\ 3\ 4\ \underline{5})$

9 : $(6\ 5\ \underline{6}\ 1\ \underline{2}\ 3\ 2\ \underline{3}\ 4\ \underline{5})$

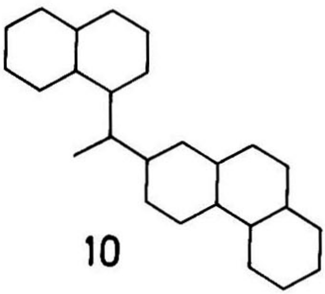
First we substitute the starting vertex of **7** by the rotated boundary sequence of polyhex **9** and obtain:

$$\left[(4\ \underline{5}\ 6\ 5\ \underline{6}\ 1\ \underline{2}\ 3\ 2\ \underline{3})\ 1\ 2\ 0\ 0\ 0\ 0\ 0\ 0 \right]$$

Next we replace the black labeled vertex of **7** by the boundary sequence corresponding to hexagon (a hex) **8** and obtain the representation of the supertree in Fig.7 :

$$\left[(4\ \underline{5}\ 6\ 5\ \underline{6}\ 1\ \underline{2}\ 3\ 2\ \underline{3})\ 1\ 2\ 0\ 0\ 0\ 0\ 0\ 0\ (6\ 1\ 2\ 3\ 4\ \underline{5}) \right]$$

In the case of structures such as one shown below,

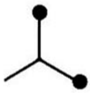


10

which contains a tree subgraph ,

11

The two points of attachement of polyhexes are identical. The supertree corresponding to **10** , is given below ,



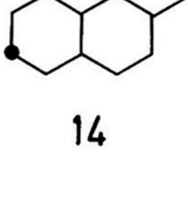
12

with the N-tuple representation ,

12 : (3 0 0 0)

The substituted polyhexes belonging to **10** are depicted below (black dots denoting the positions of attachment),

13



14

Their boundary sequences are as follows:

13 : (6 5 6 1 2 3 2 3 4 $\underline{5}$)

14 : (6 5 $\underline{6}$ 1 2 3 2 1 2 3 4 5 4 5)

In order to ensure the lexicographically highest representation for structure **10**, in the N-tuple representation of **12** first must enter boundary sequence for **14** and then for **13**. This is given below:

12 : [3(6 5 $\underline{6}$ 1 2 3 2 1 3 4 5 4 5)(6 5 6 1 2 3 2 3 4 $\underline{5}$)0] .

ACKNOWLEDGEMENT This work was supported in part by the German-Yugoslav scientific cooperation administrated through Internationales Büro-KFA-Jülich .

REFERENCES

[1] J.V.Knop, W.R.Müller, Ž.Jeričević, and N.Trinajstić: "Computer Enumeration and Generation of Trees and Rooted Trees", J. Chem. Inf. Comput. Sci. **21**, 91-99 (1981).

[2] N.Trinajstić, Ž.Jeričević, J.V.Knop, W.R.Müller, and K.Szymanski: "Computer Generation of Isomeric Structures", Pure Appl. Chem. **55**, 379-390 (1983).

[3] J.V.Knop, K.Szymanski, Ž.Jeričević, and N.Trinajstić: "Computer Enumeration and Generation of Benzenoid Hydrocarbons and Identification of Bay Regions", J. Comput. Chem. **4**, 23-32 (1983).

[4] J.V.Knop, K.Szymanski, Ž.Jeričević, and N.Trinajstić: "Computer Generation and Identification of Carcinogenic Bay Regions in Benzenoid Hydrocarbons", Int. J. Quantum Chem. **23**, 713-722 (1983).

[5] J.V.Knop, K.Szymanski, L.Klasinc, and N.Trinajstić: "Computer Enumeration of Substituted Polyhexes", Comput. Chem. in press.

[6] J.V.Knop, K.Szymanski, G.Jashari, and N.Trinajstić: "Computer-assisted Generation of Aza-benzenoids", Glasnik kemičara i tehnologa Kosova, submitted.

[7] A.T.Balaban: "Enumeration of Cyclic Graphs", in Chemical Applications of Graph Theory, Edited by A. T.Balaban, Academic, London, 1976, pp. 63-105.

[8] F.Harary: Graph Theory , Addison-Wesley, Reading, Mass., 1971, second printing, p. 194 .

[9] D.Cvetković, I.Gutman, and N.Trinajstić: "Graph Theory and Molecular Orbitals. VII . The Role of Resonance Structures", J. Chem. Phys. **61** 2700-2706 (1974) .

[10] B.Džonova-Jerman-Blažič and N.Trinajstić: "Computer-Aided Enumeration and Generation of the Kekulé Structures in Conjugated Hydrocarbons", Comput. Chem. **6** , 121-132 (1982) .

[11] J.V.Knop, K.Szymanski, G.Jashari, and N.Trinajstić: "The Connection Between the Adjacency Matrix and the Boundary Code of Benzenoid Hydrocarbons", Croat. Chem. Acta **56** , 443-450 (1983) .

[12] Ž.Jeričević: private communication (July 1983) .

[13] A.T.Balaban: "Chemical Graphs. XII . Configuration of Annulenes" , Tetrahedron **27** , 6115-6131 (1971) .

[14] J.V.Knop, K.Szymanski, Ž.Jeričević, and N.Trinajstić: "Enumeration of Benzenoid Hydrocarbons", Report at The Eight Meeting of Croatian Chemists, Zagreb, February 14-16, 1983 .

[15] Ref. 8 , p. 32

[16] F.Harary and E.M.Palmer, Graphical Enumeration, Academic, New York, 1973 , p. 6 .