Bul. Ştiinţ. Univ. Baia Mare Ser. B,
Matematică-Informatică, Vol. XIV (1998), Nr. 1, 103-110

Dedicated to Professor Iulian Coroian on his 60th anniversary

A NEW METHOD FOR ORDERING THE VERTICES OF MOLECULAR GRAPHS

Zoiţa BERINDE and Vasile BERINDE

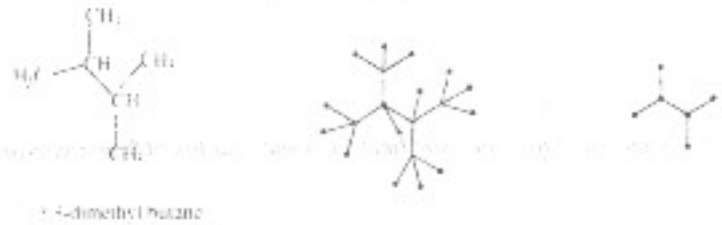
Abstract. In this paper a new algorithm,GAMA, for intramolecular ordering of molecular graphs is described.The algorithm is then applied for ordering acyclic connected graphs (trees) as well as cyclic molecular graphs.
A set of ten cyclic compact graphs has been considered in order to compare the results obtained by means of our algorithm, on one hand, to those obtained by means of others well-known procedures, IVEC [9] and MOLCEN [10], on the other hand. To define this new algorithm we use the concept of weighted electronic distance (d.e.p.) introduced in [6].

A graph, $G = (V,E)$, is an ordered pair of two sets V and E, where V is a non-empty finite set and E denotes a binary relation defined on V. The elements of V are called vertices and the elements of E are called edges.

In this paper we are concerned only with molecular graphs, when the atoms are represented by vertices and the edges are the chemical bonds.

A molecular graph is always connected and non directed graph and the molecular graphs which correspond to acyclic structures are trees.

It is customary in chemistry, on occasion, to write only the carbon skeleton of a molecule for which the hydrogen atoms are assumed to be positioned according to the valence of the carbons. Such a graph of a carbon skeleton is called a hydrogen-suppressed graph. The graph for 2,3-dimetyl-butane along with its hydrogene suppressed graph is shown in Figure 1.



2,3-dimethyl butane

G 1 (a) G 1 (b) G 1 (c)

Figure 1 The molecule 2,3-dimethyl-butane (a), its graph (b) and its hydrogen-suppressed graph (c)

In the present paper, all considerations will be made for hydrogen-suppressed graphs but keeping in mind the presence of hydrogen according to the valence of the carbons. The **topological distance** between two vertices *i* and *j* equals the number of edges on the shortest path from *i* to *j*, two adjacent points (vertices) have the topological distance 1 [3]. We have considered a new distance, introduced in a different manner, by counting the number of electrons of the adjacent vertices *i* and j and we called it "weighted electronic distance", denoted by d.e.p. [4].

The **weighted electronic distance**, for molecular graphs containing only simple covalent bonds, is given by

$$d.e.p.(i,j) = \frac{Z_i}{v_i} + \frac{Z_j}{v_j} \qquad (1)$$

where:

- Z_i denotes the order number of the atom *i* (that is, the number of all electrons in the atom i);
- v_i is the degree of the vertex i (that is, the number of edges incident to *i*).

For the class of saturated hydrocarbons, which are chemical compounds containing only carbon and hydrogen atoms and simple covalent bonds between them, we have $Z_i = Z_j = 6$. The equation (1) will be in this case of the form

$$d_{EP}(v_i, v_j) = \frac{\delta(v_i - v_j)}{r_{ij}^2} \qquad (2)$$

Instead of the usual vertex degree, v_i, we have considered the **weighted degree**, denoted by SEP, and defined as the sum of the weighted electronic distances of all edges incident to i.

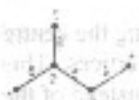
If we denote the set of vertices of a graph G by $V(G) = \{1, 2, \ldots, n\}$ and the set of its edges by E(G), then the adjacency matrix [5], $A = (a_{ij})$, is a quadratic symmetric $n \times n$ matrix given by

$$a_{ij} = \begin{cases} 1, & \text{if } i \neq j \text{ and } (i,j) \in E(G) \\ 0, & \text{if } i = j \text{ or } (i,j) \notin E(G) \end{cases} \qquad (3)$$

By replacing the usual topological distance between two adjacent vertices i and j (always = 1) by the weighted electronic distance, we obtained [4] the so called **electronic weighted matrix of connectivity**, CEP, which is a quadratic and symmetric $n \times n$ matrix. Its entries, $[CEP]_{ij}$, are given by

$$[CEP]_{ij} = \begin{cases} d_{EP}(i,j), & \text{if } i \neq j \text{ and } (i,j) \in E(G) \\ 0, & \text{if } i = j \text{ or } (i,j) \notin E(G) \end{cases} \qquad (4)$$

Figure 2 shows the graph G 2 of 2-methyl butane with the weighted electronic distances represented on each edge, the CEP matrix, the weighted degrees, SEP_i, and the ratio SEP_i / v_i, denoted by I_{v_i}.



G 2

$$CEP(G\,2) = \begin{bmatrix} 0 & 8 & 0 & 0 & 0 \\ 8 & 0 & 5 & 0 & 8 \\ 0 & 5 & 0 & 9 & 0 \\ 0 & 0 & 9 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 \end{bmatrix} \quad \begin{matrix} SEP_i & I_{v_i} \\ 8 & 8 \\ 21 & 7 \\ 14 & 7 \\ 9 & 9 \\ 8 & 8 \end{matrix}$$

Figure 2. The CEP matrix for G 2.

The degree of the vertex i, v_i is just the sum of the entries on row i in the adjacency matrix. The weighted degree of the vertex i, SEP_i, may be defined by analogy as being the sum of the entries on row i in the weighted electronic matrix of connectivity, CEP:

$$SEP_i = \sum_{j=1}^{n} [CEP]_{ij} \qquad (5)$$

We have also considered a local invariant of the vertex i, denoted by I_{v_i} [6] and defined by

$$I_{v_i} = \frac{SEP_i}{v_i} \qquad (6)$$

The determination of the centre of a tree (molecular acyclic graph) is very important for codifying uniquely the structure of acyclic chemical compounds in linear notation systems [7] and may be used for obtaining topological centric indices [8]. It is well-known a fact independently discoved by Jordan and Sylvester that every tree has a unique centre, consisting from a vertex or an edge (bicentre), see [1], [7], [9]. Each vertex of a centre (bicentre) has the smallest excentricity.

The **excentricity** of an edge i denoted by $e(i)$ is defined by

$$e(i) = \max \{ d(i,j) / j \in V(G) \}$$

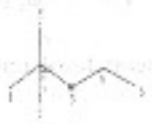
where d(i,j) is the topological distance between i and j.

The **radius** of a graph G, denoted by $\rho(G)$, is defined by

$$\rho(G) = \min \{ e(i) / i \in V(G) \}$$

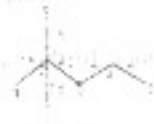
The **centre** of a graph G is the set of vertices i for which we have $\rho(G) = e(i)$.

The main aim of this paper is to introduce a new method for finding the centre of molecular graphs and also for the intramolecular ordering of its vertices. This method is obtained by using the weighted electronic distance (d e p.) instead of the usual topological distance in order to find the excentricity, the radius and the centre of the graph. Figure 3 shows, for the graph of 2,2-dimethyl pentane, the usual topological distances and the centre obtained (G 3(a)) as well as the weighted electronic distances and the centre obtained (G 3(b)). One also can use the algorithm for the computation of the excentricities.



G 3(a)

e(1) = 3; e(2) = 3; e(3) = 2;
e(4) = 3; e(5) = 4; e(6) = 4
i=3



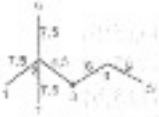
G 3(b)

e(1)=27; e(2) = 18; e(3) =15
e(4) = 18; e(6) = 22; e(7) = 27.
i=3

Figure 3. Finding the centre of a tree

The same result will be obtained if we will use the local invariant on vertex, $IC_i = SEP_i / v_i$. The smallest value of IC_i corresponds to the centre of the graph.

If we find more than two vertices having the same minimum value for IC_i, then we shall apply the principle of maximum excentricity, using d.e.p., and we will choose the one having the smallest excentricity. Figure 4 shows the finding of the centre of the graph G.3(b) by means of the latter procedure.



G 3(b)

$CEP(G.3.b) =$

0	7,5	0	0	0	0	0	7,5	7,5
7,5	0	4,5	0	0	7,5	7,5	27	6,75
0	4,5	0	6	0	0	0	10,5	**5,25** ⇐
0	0	6	0	9	0	0	15	7,5
0	0	0	9	0	0	0	9	9
0	7,5	0	0	0	0	0	7,5	7,5
0	7,5	0	0	0	0	0	7,5	7,5

Figure 4. Finding the centre of the graph by using IC_i

Due to its practical relevance, the finding of the centre of cyclic graphs was considered by many authors. Bonchev [9] proposed his hierarchical criteria 1D-3D, used by the algorithm IVEC [7], and Diudea [10] introduced the algorithm MOLORD for the finding of the centre of a graph and for the intramolecular ordering of its vertices.

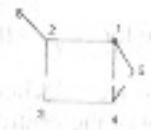
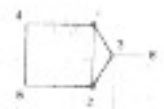
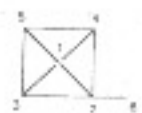
The algorithm we consider in this paper, GAMA, is based on the increasing ordering of the vertices of a graph using the values I^{γ}_{γ}, $\gamma = 1,0$. The smallest value of I^{γ}_{γ} correspond to the centre of the graph. Our hierarchical criteria are the following two:

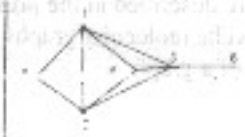
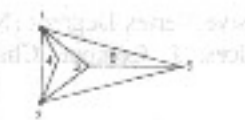
1G. the smallest value of I^{γ}_{γ}.

2G. the smallest excentricity (calculated with weighted electronic distances).

The centric ordering obtained by applying IVEC, MOLCEN (testing topological indices of C(LK) and X(LK) type) and respectively by means of GAMA, for a set of polycyclic graphs taken from [11], is given in Table 1.

Table 1 **IVEC, MOLCEN and GAMA ordering**

GRAPH	ALGORITHM	CENTRIC ORDERING
	IVEC	(1),(2),(3),(4),(5),(6)
	C(LK)	(1),(2),(3),(4),(5),(6)
	X(LK)	(1),(4),(2),(3),(5),(6)
	GAMA	**(1),(4),(3),(5),(2),(6)**
	IVEC	(1,2),(3),(4,5),(6)
	C(LK)	(1:2),(3),(4,5),(6)
	X(LK)	(1,2),(3),(4:5),(6)
	GAMA	**(1:2),(3),(4;5),(6)**
	IVEC	(1),(2),(3,4),(5,6)
	C(LK)	(1),(2),(3;4),(5,6)
	X(LK)	(1),(2),(3:4),(5,6)
	GAMA	**(1),(3;4),(2),(5;6)**
	IVEC	(1),(2),(3,4),(5),(6)
	C(LK)	(1),(2),(3,4),(5),(6)
	X(LK)	(1),(2),(3:4),(5),(6)
	GAMA	**(1),(3;4),(5),(2),(6)**

	IVEC	(1;2),(3),(4),(5),(6)
	C(LK)	(1;2),(3),(4),(5),(6)
	X(LK)	(1;2),(3),(4),(5),(6)
	GAMA	**(1;2),(4),(3),(5),(6)**
	IVEC	(1;2;3),(4),(5),(6)
	C(LK)	(1;2;3),(4),(5),(6)
	X(LK)	(1;2;3),(4),(5),(6)
	GAMA	**(1;2;3),(5),(4),(6)**
	IVEC	(1),(2),(3),(4),(5),(6)
	C(LK)	(1),(2),(3),(4),(5),(6)
	X(LK)	(1),(4),(2),(3),(5),(6)
	GAMA	**(1),(3),(4),(2),(5),(6)**
	IVEC	(1),(2),(3),(4),(5),(6)
	C(LK)	(1),(2),(3),(4),(5),(6)
	X(LK)	(1),(4),(2),(3),(5),(6)
	GAMA	**(1),(3),(2),(4),(5),(6)**
	IVEC	(1),(2,3),(4,5),(6)
	C(LK)	(1),(2,3),(4,5),(6)
	X(LK)	(1),(2,3),(4,5),(6)
	GAMA	**(1),(2;3),(4;5),(6)**
	IVEC	(1;2),(3),(4),(5),(6)
	C(LK)	(1;2),(3),(4),(5),(6)
	X(LK)	(1;2),(3),(4),(5),(6)
	GAMA	**(1;2),(4)(3),(5),(6)**

The following general remarks may be made for the obtained results given in Table 1.

a) the ordering obtained by means of GAMA method is approximately the same as the ordering obtained by others procedures.

b) for all the ten structures considered the centre of the graph is the same;

c) for all the ten structures considered the most peripherical vertex is the same, independently of the method used.

This enable us to conclude that the GAMA procedure described in the present paper is useful for finding the centre of both acyclic and cyclic molecular graphs. It is also useful for intramolecular ordering of the vertices of a graph.

REFERENCES

[1] *Harary F.*, Graph Theory , Addison-Wesley, Reading Mass., 1969

[2] *Basak, S. C., Magnuson, V. R., Niemi, G.J., Regal, R. R., Veith, G. D.*, Determining structural similarity of chemicals using graph-theoretic indices. Discr. Appl. Math., 19(1988), 17-44

[3] *Trinajstic, N.*, Chemical Graph Theory, CRC Press Inc.,Boca Raton, Florida, 1983

[4] *Bernde, Z.*, A new connectivity matrix, Bulletins for Applied Mathematics, vol LXXX VI-a / 98, 1518, 77-86

[5] *Sylvester, J.J.*, Am. J. Math., 1(1874), 64

[6] *Bernde, Z.*, Topological indices obtained from the CEP matrix of connectivity, Proceed. of First Southern Symposium on Computing, Univ. of Southern Mississippi (to appear)

[7] *Bonchev, D., Mekenyan, O., Balaban, A.T.*, Iterative procedure for the generalized graph center in polycyclic graphs, J Chem. Inf. Comput. Sci., 29 (1989) 91-97

[8] *Balaban, A.T., Motoc, I.*, Correlation between Octane Number and Topological Indices of Alkanes, Mathch, 5(1979), 197-218

[9] *Bonchev, D., Balaban, A.T., Randić M.*, The graph center concept for polycyclic graphs, Int. J. Quantum Chem.,19(1981), 61-82

[10] *Dražeta, M., Minailuc, O.M., Balaban, A.T.*, Regressive Vertex Degrees (New Graph Invariants) and Derived Topological Indices, J. Comput. Chem., 12(1991), 527-535

Received 20.07.1998

Zorja BERINDE
North University of Baia Mare
Department of Chemistry
Victoriei 76, 4800 Baia Mare
ROMANIA

Vasile BERINDE
North University of Baia Mare
Department of Mathematics and Computer Science
Victoriei 76, 4800 Baia Mare
ROMANIA
E-mail: vberinde@univer.ubm.ro