On molecular graphs and some topological indices

Zoita Berinde[1], Vasile Berinde[2] and Laszlo Balog[2]
[1] *"V. Lucaciu" High School Baia Mare*
[2] *Department of Mathematics and Computer Science Faculty of Sciences, North University of Baia Mare*

Abstract

The paper is devoted to the study of some physico-chemical properties of chemical substances, by means of mathematical approaches, i.e. by using molecular graphs and topological indices.

Key words: *molecular graph, adjacency matrix, distance matrix, topological index, correlation*

1 Introduction

The classical mathematical model of molecular complexity based on concepts from graph theory is cosidered in order to automatically generateand study various properties of alkanes by means of topological indices. The enumeration of chemical isomers, in particular the constitutional isomers of alkanes C_2H_{2n+2} was a challenging mathematical problem to which, since 1874-1875, Cayley applied the graphical concept of a tree.

If only carbon atoms of alkanes are depicted, there is a one-to-one correspondence between these isomers and trees, that is graphs of at most agree 4. The application of graph theory to the equivalence or non-equivalence of such two different structures becomes the task of finding an invariant (a quantity characteristic of graph that does not change when the graphis embedded differently), which is called a topological index.

A topological index expresses in numerical form the topology of chemical species it represents. It is obtained in the following way: starting the hydrogen-suppressed graph of the species under study, a matrix is set up to reflect the topological structure of the graph. This matrix is then converted into a number by various formulas, that is into a topological index.

At their origins, topological indices were developed for the purpose of obtaining correlations with a wide variety of the physico-chemical substances. In the same manner, the aim of this paper is to describe a procedure for studying the correlation of some topological indices with various physico-chemical properties. This procedure is integrated in a computer programme for automatically generating of all isomers of alkanes for $n = 2, 3, \ldots, 12$.

These investigations are very important in the design of drugs and pharmaceutical industry [4], [5].

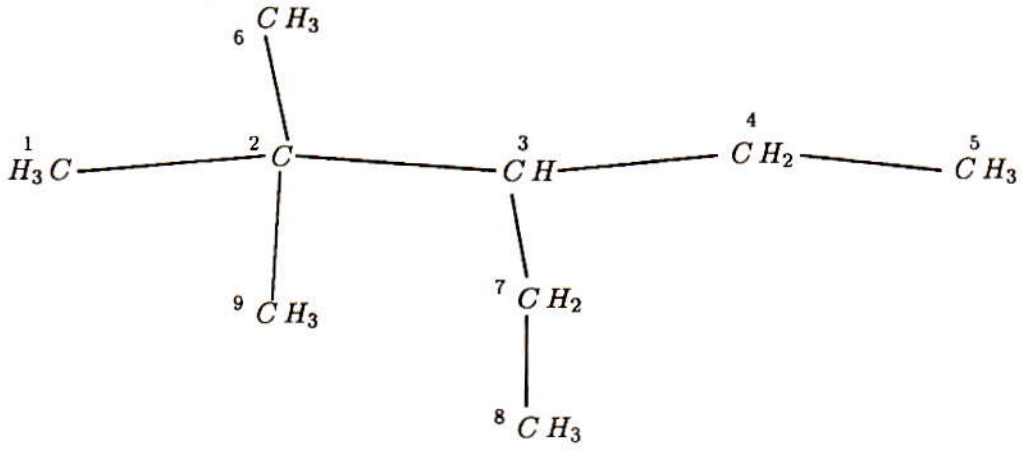
2 Graph representations of chemical structures

Among the various kinds of graphs used in chemistry (chemical graphs), much attention is paid to the molecular (constitutional) graphs since constitutional formulas are depicted by

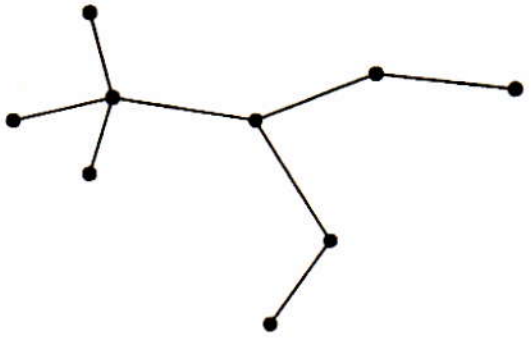
points representing atoms, and by lines representing covalent bonds. The basic mathematical structure which maps a certain molecule (molecular graph G) is the adjacency matrix of the molecular graph $A(G)$. For a molecule having N atoms (of carbon, when the hydrogen atoms are supressed), $A(G)$ is square $N \times N$ matrix. Its entries a_{ij} have only two different values: 1 and 0, due to the fact that two atoms in a molecule are in binary relation, being either connected or notconnected:

$$a_{ij} = 1, \quad \text{for } i, j \text{ - adjacent}$$
$$a_{ij} = 0, \quad \text{otherwise.}$$

For following chemical compound having the structure



and called 2,2-dimethyl, 3-ethilpentane, the corresponding chemical graph G is



and the adjacency matrix $A(G)$ is given by

$$A(G) = \begin{array}{c|ccccccccc|c} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & d_i \\ \hline 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 2 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 4 \\ 3 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 3 \\ 4 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 2 \\ 5 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 6 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 7 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 2 \\ 8 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 9 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

REMARK 1 *$A(G)$ is a symetric matrix, so we can use only the reduced adjacency matrix, as in [1]-[2]. By summing the entries on the row (column) i in the adjacency matrix we obtain the degree of the vertex v_i, denoted by d_i. The number d_i gives the chemical character of the atom i.*

An other matrix associated to G is the so called distance-matrix, denoted by $D(G)$ or simply D, $D = (d_{ij})$. Here d_{ij} stands for the topological distancel between the vertices i and j. For G given above, $d_{15} = 4 = d_{18}$, $d_{16} = 2, \ldots$.

3 Topological indices

The adjacency matrix expresses in a certain way the complexity of the molecule (isomer). We have to convert the information given by this matrix (or other matrix) into a number in such a way that the same value will be obtained for isomorphic structures, that is, for different isomers we have to obtain different values.

There exists a pletora of topological indices [3]. Such topological indices are, for instance

$$A' = \sum_{i,j=1}^{N} a_{ij} \quad \text{(total adjacency index)}$$

and

$$s(i) = \sum_{j=1}^{N} d_{ij}.$$

If v_i is the degree of vertex i (the number of bonds) then we can write some of the so-called Zagreb Group indices

$$M_1 = \sum_{i=1}^{N} v_i^2$$

$$M_2 = \sum_{all\ edges} v_i v_j$$

$$\chi_R = \sum_{all\ edges} (v_i v_j)^{-\frac{1}{2}}, \quad \text{and so on.}$$

It was observed that such a topological index generally correlates, well or not, with various physico-chemical properties. So we can obtain very important information about such properties for various chemical compounds.

4 A method for generating the isomers of alkanes

The number of the isomers of alkanes increaseas geometrically as N increases

N	Numbers of izomers
3	1
4	2
5	3
6	5
7	9
8	18
9	35

In order to simplify the study of various topological indices, a special algorithm (and a corresponding programme in PASCAL) was obtained (see [1], [2]). We used this programme to generate the isomers of alkanes for $N = 10$, $N = 11$ and $N = 12$ and we obtained 75, 155 and 355 isomers, respectively (for $N = 20$, there exist 366 319 isomers).

Then some topological indices has been tested in what concern correlation with certain physico-chemical properties so boiling point, heats of vaporization, etc.

The results of the statistical study indicate good correlations for χ_R and heat of vaporization. For $N = 8$ (18 isomers) these results given here show a good correlation (94%):

Isomers	1	2	3	4	5	6
ΔH_{vap} observed	8.410	8.826	8.402	8.915	9.209	8.897
ΔH_{vap} computed	8.304	8.747	8.747	9.074	9.177	8.747

Isomers	7	8	9	10	11	12
ΔH_{vap} observed	9.014	9.272	9.029	9.051	9.484	9.081
ΔH_{vap} computed	8.853	9.177	9.177	9.177	9.493	9.074

Isomers	13	14	15	16	17	18
ΔH_{vap} observed	9.476	8.973	9.316	9.521	9.483	9.915
ΔH_{vap} computed	9.493	9.074	9.177	9.493	9.493	9.801

The regression line is $y = 15,618000 - 0,298547x$ and the standard deviation is $s = 0,137352$.

References

[1] Berinde Z., Berinde V.: *Matrix representation and posibilities of study and generation of all structures of a chemical substance* (in Romanian), Analele Univ. Oradea, 1(1991), Fasc. Fizica-Chimie, 108-123.

[2] Berinde Z.:*On a method of representation and counting of the structures of alkanes in Romanian*, Bul. Stii. Univ. Baia Mare, Fac. Fizica-Chimie, 8(1991), 231-238.

[3] Diudea M., Ivanciuc O.: *Molecular topology* (in Romanian), Editura COMPREX, Cluj-Napoca, 1995.

[4] Kier L.B., Hall L.H.: *Molecular Connectivity in Structure-Activity Analysis*, John Willey & Sons, New York, 1986.

[5] Kier L.S., Hall L.H: *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976.