Variable Neighborhood Search for Extremal Graphs. 10. Comparison of Irregularity Indices for Chemical Trees

Ivan Gutman,[†] Pierre Hansen,*[,‡] and Hadrien Mélot[§]

GERAD - Ecole des Hautes Etudes Commerciales, 3000, chemin de la Cote-Sainte-Catherine, Montréal, Quebec H3T 2A7, Canada, Faculty of Science, University of Kragujevac, Serbia and Montenegro, and Department of Computer Science, University of Mons − Hainaut, Belgium

Received November 28, 2003

Chemical graphs, as other ones, are *regular* if all their vertices have the same degree. Otherwise, they are *irregular*, and it is of interest to measure their irregularity both for descriptive purposes and for QSAR/QSPR studies. Three indices have been proposed in the literature for that purpose: those of Collatz-Sinogowitz, of Albertson, and of Bell's variance of degrees. We study their properties for the case of chemical trees. Structural conjectures are generated with the system *AutoGraphiX*, and most of them proved later by mathematical means. Analytical expressions for extremal values are obtained, and extremal graphs are characterized for the two last indices.

1. INTRODUCTION

A graph is said to be regular if all its vertices have the same degree (= number of first neighbors). Otherwise, it is irregular. This, of course, applies also to molecular graphs.[1−6] Regular graphs dealt with in chemical graph theory are mainly of degree 2 (molecular graphs of annulenes and cycloalkanes) and 3 (molecular graphs of fullerenes). The interest toward regular molecular graphs has much increased after the experimental discovery of fullerenes and the elaboration of their theory.[7]

Despite this, the vast majority of graphs of interest to chemists is irregular. Clearly, some are more irregular than others. It is our intention to make this notion more precise, through some irregularity measure.

The purpose of the present paper is 3-fold.

First, we wish to inform the chemical community of the irregularity measures that already exist in the mathematical literature.

Second, we establish some properties of these irregularity measures, applicable to molecular graphs, chemical trees in particular. By doing this we reveal the close connection between the concept of molecular-graph-irregularity and branching. We show that some irregularity measures are, in fact, branching indices. Some, on the other hand, are unrelated to branching.

Third, according to Balaban,[8] in order that a graph invariant be acceptable as a usable molecular structure descriptor (topological index) it must monotonically increase (or decrease) with the increase of the extent of branching, especially in the case of chemical trees. Thus, our analysis implies that some irregularity measures are, whereas some are not, expected to be of practical value in QSPR/QSAR studies.

Three main irregularity measures have been proposed in the literature:

(i) The *Collatz-Sinogowitz index*,[9] defined for any graph $G = (V, E)$ with vertex set V and edge set E by

$$\mathrm{CS}(G) = \lambda_1 - \frac{2m}{n} \tag{1}$$

where λ_1 is the *index* or largest eigenvalue of the *adjacency matrix* $A = (a_{ij})$ (with $a_{ij} = 1$ if vertices i and j are joined by an edge and 0 otherwise), m is the number of edges and n the number of vertices of G ($2m/n$ is the average degree);

(ii) the *variance of degrees*[10] defined for any graph $G = (V, E)$ by

$$\mathrm{VAR}(G) = \frac{1}{n}\sum_{i=1}^{n-1} n_i\left(i - \frac{2m}{n}\right)^2 \tag{2}$$

where n_i denotes the number of vertices of degree i for $i = 1, 2,..., n - 1$;

(iii) the *Albertson index*,[11] defined for any graph $G = (V, E)$ by

$$\mathrm{A}(G) = \sum_{(i,j)\in E} |d_i - d_j| \tag{3}$$

where d_i denotes the degree of vertex i for $i = 1, 2,..., n$.

For general graphs, it was conjectured in ref 9 that stars maximize the difference between λ_1 and $2m/n$. This conjecture was refuted.[12] CS(G) and VAR(G) are compared in ref 10. Tight upper bounds on CS(G) and VAR(G) expressed as functions of n and m for both connected and disconnected graphs are also given there. It is shown in ref 11 that

$$\mathrm{A}(G) < \frac{4n^3}{27}$$

and that this bound can be approached arbitrarily closely. A tight upper bound on A(G) expressed as a function of m and

* Corresponding author e-mail: Pierre.Hansen@gerad.ca.
† University of Kragujevac.
‡ HEC Montreal.
§ University of Mons − Hainaut.

10.1021/ci0342775 CCC: $30.25
Published on Web 02/09/2005

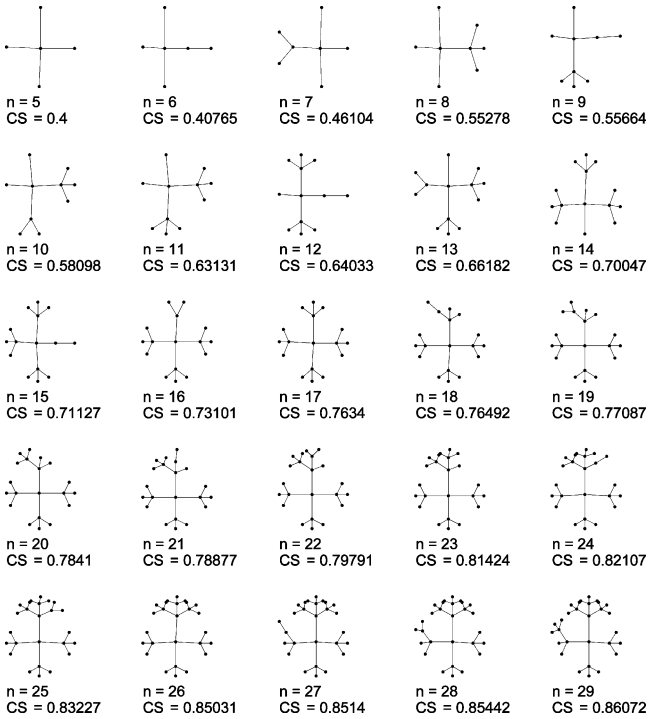

Figure 1. Extremal graphs found by *AGX* for CS(*T*).

n is given in ref 21, and it is shown that extremal graphs are a particular class of *split* graphs (which consist of a clique, an independent set, and some edges joining a vertex in the clique to another one in the independent set).

All three indices are nonnegative and equal to 0 if and only if the graph *G* is regular.

More can be said if one restricts attention to some particular class of graphs, such as e.g. chemical graphs. In this paper, we begin such a study by considering chemical trees, i.e., trees with maximum degree 4.

The system *AutoGraphiX* (*AGX*)[13−21] is first used to obtain extremal or near-extremal chemical trees for each of the three indices, taking *n* as parameter. This leads immediately to conjectures on the structure of such trees, from where analytical expressions for the maximum of VAR(*T*) and A(*T*) are deduced. Then, these conjectures are proved.

The paper is organized as follows: the next three sections are devoted to results on CS(*T*), VAR(*T*), A(*T*), together with their proofs (when available), respectively. Section 5 compares the indices. Conclusions are presented in Section 6.

2. COLLATZ-SINOGOWITZ INDEX

2.1. Extremal Graphs. It was shown that CS(*T*) is minimum for paths.[22] This automatically determines the chemical tree with minimal CS-value. Chemical trees with maximum CS(*T*) have recently been investigated, on the basis of a complete enumeration of all chemical trees with $n \leq 21$.[23] The use of *AGX* to find chemical trees with maximum and near maximum CS(*T*) for $n \leq 29$ (see Figure 1) fully corroborated these results.

2.2. Conjectures. Extremal chemical trees found for $n \leq 29$, belong to the class of *dendrimers*. In what follows we refer to them as the D_n-dendrimers. These are defined as follows: a first vertex, 1, at level 0 is connected to up to 4 vertices 2, 3, 4, 5 at level 1 and exactly that number if $n \geq 5$. Then each vertex at level *k* in order of increasing indices

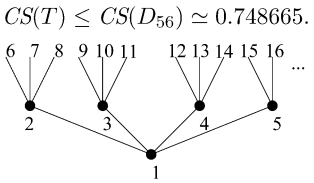

Figure 2. Construction of dendrimers, which are extremal graphs for CS(*T*).

is connected to 3 new vertices at level $k + 1$, for $k = 1, 2,...$ until vertices are exhausted (see Figure 2).

However, no mathematical proof that these trees are extremal with regard to CS(*T*) for all *n* is as yet available. So we reiterate

Conjecture 1 (Ref 23). The chemical trees with maximum Collatz-Sinogowitz index are the *D_n-dendrimers and are unique for each n*.

Several surprising results about D_n-dendrimers have been recently found[24] and are presented in ref 23:

(i) they are chemical trees with minimum *Wiener index* or sum of distances between pairs of vertices (and hence with minimum average distance). For a review on the Wiener index of trees and an exhaustive bibliography see ref 25; (ii) they are chemical trees with minimum *greatest Laplacian eigenvalue*;[26] (iii) it follows from the characterization of chemical trees with minimum *Randic index* Ra(*T*),[27] where

$$\mathrm{Ra}(T) = \sum_{(i,j)\in E} \frac{1}{\sqrt{d_i d_j}} \tag{4}$$

given in ref 16 that they have minimum Randic index if $n = 3k + 2$ for $k = 1, 2,...$

Moreover, it is conjectured that for these latter values of *n*, dendrimers D_n also have minimum *energy* E,[28] where

$$\mathrm{E} = \mathrm{E}(T) = \sum_{i=1}^{n} |\lambda_i| \tag{5}$$

and the λ_i denote the eigenvalues of the adjacency matrix, as well as minimum *Hosoya index* Z, where

$$\mathrm{Z} = \mathrm{Z}(T) = \sum_{k\geq 0} m(T, k) \tag{6}$$

and the $m(T, k)$ are the numbers of *k*-matchings of *T* (see ref 5 pp 127−134).

An analytical expression for CS(D_n) as a function of *n* is also missing at present. However, some insight may be obtained from its graphical representation, given in Figure 3. The curve of maximum CS(*T*) first increases rapidly, with kinks corresponding to successive sets of 3 values of *n*, then goes through a maximum and decrease slightly. A zoom on the values for $n = 45$ to 65, separated according to *n*(mod 3), is presented in Figure 4. This leads to

Conjecture 2. For all chemical trees *T*

$$\mathrm{CS}(T) \leq \mathrm{CS}(D_{56}) \simeq 0.748665$$

3. VARIANCE OF DEGREES

3.1. Extremal Graphs. The extremal chemical trees found by *AGX* for the variance of degrees are presented in Figure 5. These trees belong to 3 families and contain (as for the CS index) vertices of degree 1 and 4 as well as one vertex

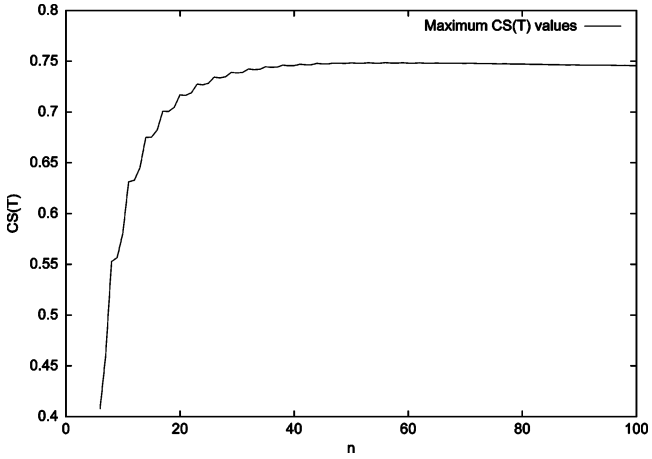

Figure 3. Bounds on *CS*(*T*) for a chemical tree *T* of order *n*.

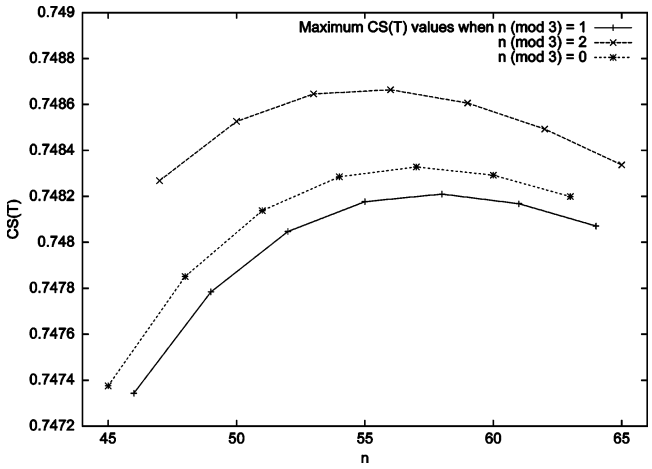

Figure 4. Bounds on *CS*(*T*) for a chemical tree *T* of order $n \in$.[45,65]

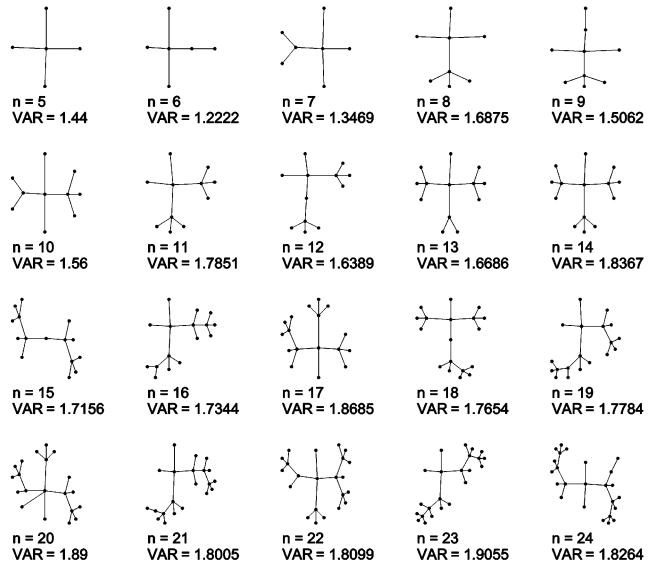

Figure 5. Extremal graphs found by *AGX* for VAR(*T*).

of degree 2 when $n(\mathrm{mod}\ 3) = 0$ and one vertex of degree 3 when $n(\mathrm{mod}\ 3) = 1$.

As $2n_4 + n_3 = n_1 - 2$ for all chemical trees, one can show that increasing n_4 and n_1 increases the variance of degrees more than increasing n_3 or n_2. This shows that trees which are conjectured to be extremal for CS are also extremal for the variance of degrees.

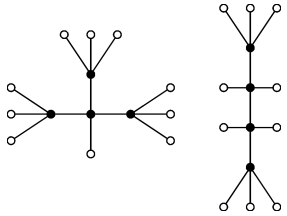

Figure 6. All extremal graphs for VAR when $n = 14$.

However, they are a subset of them, because any tree having the same vertex degree distribution as D_n has the same VAR-value as D_n. Transformations which keep the distribution of degrees constant are easy to make. Recall that a *caterpillar* is a tree composed of a path and pending edges appended to some or all of its vertices. It thus appears that the following are representative elements of the class of *n*-vertex chemical trees with maximum VAR:

(i) caterpillars with all inner vertices of degree 4 when $n(\mathrm{mod}\ 3) = 2$; (ii) as in (i) with one edge added at an end vertex of a longest path when $n(\mathrm{mod}\ 3) = 0$; (iii) as in (i) with two edges added at an end vertex of a longest path when $n(\mathrm{mod}\ 3) = 1$.

In fact, in the trees just described, all vertices of degree 4 are on a path. Other extremal graphs for VAR(*T*) can be obtained easily: for family (i) we have to form a chemical tree with n_4 vertices and add the pending edges to obtain a degree 4 for each one of these vertices. A similar construction can be done for families (ii) and (iii) where we first have to choose a vertex which will be of degree 2 or 3, respectively. This leads to the following characterization of extremal trees with maximum variance of degrees.

Theorem 1. *Extremal Graphs for VAR of Family (i). In this case* $n(\mathrm{mod}\ 3) = 2$. *These are the trees with* $n_1 = (2n + 2)/3$, $n_2 = n_3 = 0$, $n_4 = (n - 2)/3$. *There are k different extremal chemical trees of order n where k is the number of chemical tree of order (n-2)/3 (the chemical trees with* n_4 *vertices for which we have to add the pending edges).*

Extremal Graphs for VAR of Family (ii). In this case $n(\mathrm{mod}\ 3) = 0$. *These are the trees with* $n_1 = 2n/3$, $n_2 = 1$, $n_3 = 0$, $n_4 = (n - 3)/3$. *There are* $\sum_{i=1}^{l} m_i$ *different extremal chemical trees of order n where l is the number of chemical tree of order n/3 and* m_i *the number of possibilities for choosing the vertex of degree 2, up to symmetry, for each of the l chemical trees (i.e., a vertex of degree 1 or 2 in the graph of order n/3).*

Extremal Graphs for VAR of Family (iii). In this case $n(\mathrm{mod}\ 3) = 1$. *These are the trees with* $n_1 = (2n + 1)/3$, $n_2 = 0$, $n_3 = 1$, $n_4 = (n - 4)/3$. *A similar construction as for family (ii) applies, but here we choose a vertex of degree 3, among those of degree 1, 2, or 3 in a chemical tree of order (n-1)/3.*

The validity of this theorem lies in the fact that all extremal graphs of a family have the same variance of degrees, and Theorem 2 will ensure that these values are extremal.

For example if $n = 14$ (family (i)), the $k = 2$ different extremal graphs, based on the star S_4 and on the path P_4, are shown in Figure 6.

If $n = 12$ (family (ii)), $l = 2$ (S_4 and P_4), $m_1 = 1$, for the star and $m_2 = 2$ for the path. Such a construction is shown in Figure 7, and for $n = 13$ (family (iii)) the construction of the extremal graphs is represented in Figure 8.

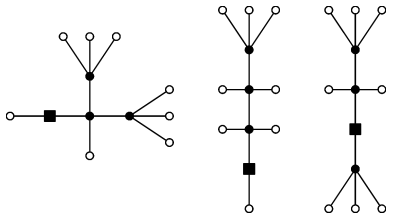

Figure 7. All extremal graphs for VAR when $n = 12$.

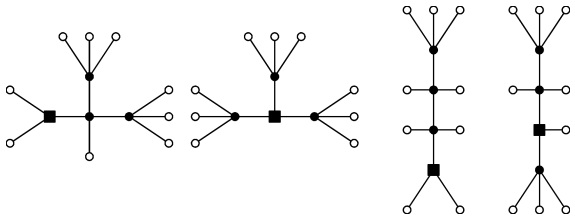

Figure 8. All extremal graphs for VAR when $n = 13$.

3.2. Bounds. Theorem 2. *For all chemical trees T_n with $n \geq 2$ vertices*

$$\mathrm{VAR}(T_n) \leq \frac{1}{n^2}[2n^2 - 2n - 4] \text{ if } n(\mathrm{mod}\ 3) = 2$$

$$\leq \frac{1}{n^2}[2n^2 - 4n - 4] \text{ otherwise}$$

and the bounds are sharp for all such n.

Proof. A simple rearrangement of the definition of $Var(G)$ in (2) yields

$$\mathrm{VAR}(G) = (M_1)/n - (2m/n)^2$$

where $M_1 = \sum_{i=1}^{l} i^2 n_i$ is the so-called *first Zagreb index*.[29−32] In case of a chemical tree T

$$\mathrm{VAR}(T) = (M_1)/n - [2(n-1)/n]^2 \qquad (7)$$

and

$$M_1 = n_1 + 4n_2 + 9n_3 + 16n_4 \qquad (8)$$

For all chemical trees

$$n_1 + n_2 + n_3 + n_4 = n \qquad (9)$$

$$n_1 + 2n_2 + 3n_3 + 4n_4 = 2(n-1) \qquad (10)$$

Solve these equations in n_1 and n_4 and substitute into (7) and (8). This yields

$$\mathrm{VAR}(T) = (6n - 10 - 2n_2 - 2n_3)/n - [2(n-1)/n]^2 \qquad (11)$$

Clearly, for a fixed value of n, VAR(T) will be maximal if $n_2 = n_3 = 0$ (if this is possible), or if n_2 and n_3 are as close to zero as possible.

If n(mod 3) = 2 it is possible to construct chemical trees (that are not unique) with $n_2 = n_3 = 0$. If n(mod 3) = 0 it is possible to construct chemical trees (that are not unique) with $n_2 = 1$, $n_3 = 0$. If n(mod 3) = 1 it is possible to construct chemical trees (that are not unique) with $n_2 = 1$, $n_3 = 0$. Substituting these choices of n_2 and n_3 into (11) results in Theorem 2.

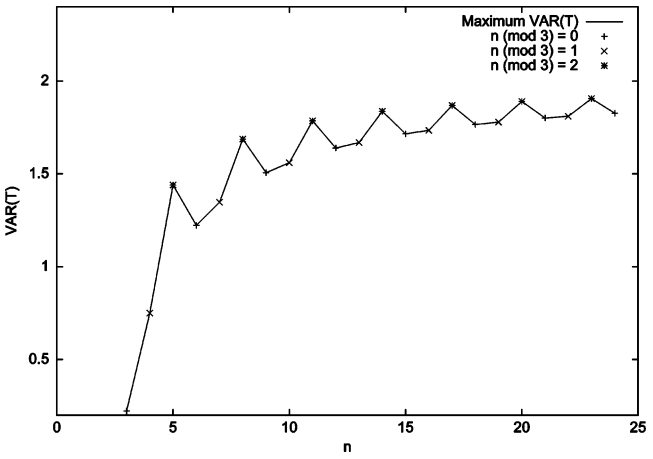

Figure 9. Bounds on *VAR*(*T*) for a chemical tree *T* of order *n*.

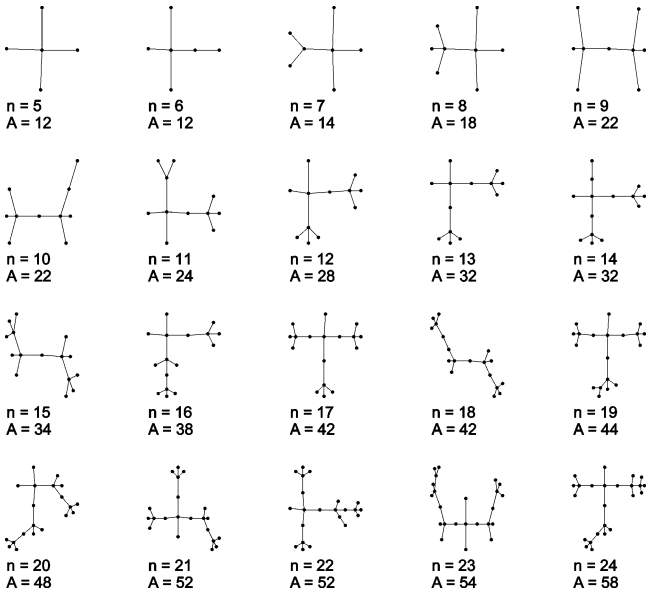

Figure 10. Extremal graphs found by *AGX* for A(*T*).

Figure 9 shows the curve of values defined in Theorem 2.

4. ALBERTSON INDEX

4.1. Extremal Graphs. The extremal chemical trees found by *AGX* for the Albertson index are presented in Figure 10. These trees have vertices of degree 1 and 4, a single vertex of degree 3 if n(mod 4) = 3, and one vertex of degree 2 if n(mod 4) = 2 as well as for all of them $\lceil n/4 \rceil - 2$ additional vertices of degree 2.

These trees are not unique. There are four classes, having the following representative elements which are caterpillars:

(i) caterpillars with vertices of degree 1, 4, 2, 4, 2,..., 4, 1 along the longest path; (ii)−(iv) as in (i) with 1 (resp 2,3) additional pending edges incident with an endvertex of the (or a) longest path.

4.2. Bounds. A sharp lower bound on A(T) for all trees is readily obtained:

Theorem 3. *For any tree T*

$$\mathrm{A}(T) \geq \max_i d_i(d_i - 1) \qquad (12)$$

the bound being sharp if and only if T is homeomorphic to a star.

Proof. Consider the vertex j with maximum degree; then there are d_j disjoint paths from j to a pendant vertex l_k for $k = 1, 2, ..., d_j$. The sum of contributions to A(T) of edges along each of these paths is at least $d_j - 1$ with equality if and only if the degrees of successive vertices along the path from j to l_k decrease monotonically (not necessarily strictly). This yields the lower bound. Moreover, if there is a vertex of degree > 2 different from j along one such path, there is at least one edge contributing positively to A(T) and not in the union of these paths. Hence T is either a star or a tree homeomorphic to a star, i.e., such that removal of vertices i of degree $d_i = 2$ along any path, and merging the corresponding incident edges $\{r, i\}$ and $\{i, s\}$ gives a star.

A sharp upper bound is given by

Theorem 4. *For any chemical tree T with n vertices*

$$A(T) \leq \begin{cases} \dfrac{5n\text{-}1}{2} & \text{if } n(\text{mod } 4) = 1, \text{and} \\ \dfrac{5n - n(\text{mod } 4) + 4}{2} & \text{otherwise} \end{cases}$$

Moreover this bound is sharp for all n ≥ 5.

Proof. We use a proof technique based on linear programming, first introduced in ref 16 and applied several times.[33,34]

Let x_{ij} be the number of edges with end vertices of degree i and j. Then

$$A(T) = x_{12} + 2x_{13} + 3x_{14} + x_{23} + 2x_{24} + x_{34} \quad (13)$$

We first solve the following system of six linear equations which are satisfied by all chemical trees:

$$x_{12} + x_{13} + x_{14} = n_1 \quad (14)$$

$$x_{12} + 2x_{22} + x_{23} + x_{24} = 2n_2 \quad (15)$$

$$x_{13} + x_{23} + 2x_{33} + x_{34} = 3n_3 \quad (16)$$

$$x_{14} + x_{24} + x_{34} + 2x_{44} = 4n_4 \quad (17)$$

$$n_1 + 2n_2 + 3n_3 + 4n_4 = 2n - 2 \quad (18)$$

$$n_1 + n_2 + n_3 + n_4 = \text{n} \quad (19)$$

with unknowns x_{14}, x_{24}, n_1, n_2, n_3, and n_4. That gives

$$x_{14} = \frac{n+3}{2} - \frac{3x_{12}}{2} - \frac{7x_{13}}{6} - \frac{x_{22}}{2} - \frac{x_{23}}{6} + \frac{x_{33}}{6} + \frac{x_{34}}{3} + \frac{x_{44}}{2} \quad (20)$$

$$x_{24} = \frac{n\text{-}5}{2} + \frac{x_{12}}{2} + \frac{x_{13}}{6} - \frac{x_{22}}{2} - \frac{5x_{23}}{6} - \frac{7x_{33}}{6} - \frac{4x_{34}}{3} - \frac{3x_{44}}{2} \quad (21)$$

$$n_1 = \frac{n+3}{2} - \frac{x_{12}}{2} - \frac{x_{13}}{6} - \frac{x_{22}}{2} - \frac{x_{23}}{6} + \frac{x_{33}}{6} + \frac{x_{34}}{3} + \frac{x_{44}}{2} \quad (22)$$

$$n_2 = \frac{n\text{-}5}{4} + \frac{3x_{12}}{4} + \frac{x_{13}}{12} + \frac{3x_{22}}{4} + \frac{x_{23}}{12} - \frac{7x_{33}}{12} - \frac{2x_{34}}{3} - \frac{3x_{44}}{4} \quad (23)$$

$$n_3 = \frac{x_{13}}{3} + \frac{x_{23}}{3} + \frac{2x_{33}}{3} + \frac{x_{34}}{3} \quad (24)$$

$$n_4 = \frac{n\text{-}1}{4} - \frac{x_{12}}{4} - \frac{x_{13}}{4} - \frac{x_{22}}{4} - \frac{x_{23}}{4} - \frac{x_{33}}{4} + \frac{x_{44}}{4} \quad (25)$$

Replacing x_{14} and x_{24} by (20) and (21) respectively gives

$$A(T) = \frac{5n\text{-}1}{2} - \frac{5x_{12}}{2} - \frac{7x_{13}}{6} - \frac{5x_{22}}{2} - \frac{7x_{23}}{6} - \frac{11x_{33}}{6} - \frac{2x_{34}}{3} - \frac{3x_{44}}{2} \quad (26)$$

which is maximal for a fixed number of vertices when the values x_{12}, x_{13}, x_{22}, x_{23}, x_{33}, x_{34}, and x_{44} are equal to zero. However, in the case of n-vertex chemical trees the condition

$$x_{12} = x_{13} = x_{22} = x_{23} = x_{33} = x_{34} = x_{44} = 0 \quad (27)$$

can be satisfied only if $n(\text{mod } 4) = 1$. Any chemical tree satisfying (27) has no vertices of degree 3, all its vertices of degree 2 are adjacent to two vertices of degree 4, and no two vertices of degree 4 are mutually adjacent. Denote the type of such n-vertex chemical trees by (i-1) (the type i when $n(\text{mod } 4) = 1$).

Hence, if $n(\text{mod } 4) = 1$, then for any n-vertex chemical tree, $A(T) \leq (5n - 1)/2$. Equality is attained if and only if T is of type (i-1).

This completes the proof of Theorem 4 in the case $n(\text{mod } 4) = 1$.

If $n(\text{mod } 4) \neq 1$, then the conditions (27) cannot be satisfied by n-vertex chemical trees. Then, to find the chemical tree(s) with maximal A-values we have to find the values of the parameters x_{12}, x_{13}, x_{22}, x_{23}, x_{33}, x_{34}, and x_{44} as close to zero as possible compatible to the existence of a chemical tree, i.e., for which the right-hand sides of eqs 20−25 are integers and for which a graph exists. In the following discussion, we call x_{12}, x_{13}, x_{22}, x_{23}, x_{33}, x_{34}, and x_{44} the *observed parameters*. Looking at (26), we see that increasing by 1 any of the observed parameters subtracts a certain value from the maximal value of A(T), expressed in sixths:

observed parameter	x_{12}	x_{13}	x_{22}	x_{23}	x_{33}	x_{34}	x_{44}
decrease on $A(T)$, in sixths	15	7	15	7	11	4	9

The gap between the maximum value of A(T), which is 5n-1/2 when all observed parameters are equal to zero, and the value of A(T) when some of these parameters are positive will be called the *decrease*. The proof for the different cases is based on the examination of all possible increases of the values of the observed parameters, choosing those with minimal decrease. Fortunately, we can reduce the number of possibilities with some simple observations. *AGX* gives us some (presumably) extremal graphs for each case. The decrease of these feasible solutions is a good upper bound on the minimal decrease. We can limit the enumeration taking only solutions with a decrease less or equal to this bound. Moreover eq 24 can considerably reduce the number of possibilities: if only one observed parameter is not equal to zero, the values of x_{13}, x_{23}, x_{33}, or x_{34} must be multiples of 3; and in general $x_{13} + x_{23} + 2x_{33} + x_{34}$ has to be a multiple of 3.

• Case $n(\text{mod } 4) = 2$. A first lower bound on the maximal A-value is the graph obtained by *AGX* and described above: in this case $x_{12} = 1$ and all other observed parameters are equal to zero. This solution has a decrease of 15 sixths. We will examine all possibilities for the observed parameters with a decrease less than or equal to 15 sixths, and for which eq 24 is not violated.

First we consider the solutions where only one of the observed parameters is different from zero.

nonzero parameter	decrease (1/6)	comment
$x_{12} = 1$	15	starting solution: type (i-2)
$x_{22} = 1$	15	feasible: type (ii-2)
$x_{34} = 3$	12	impossible in view of eq 20
$x_{44} = 1$	9	impossible in view of eq 25

We then examine the possible choices of two nonzero observed parameters with a decrease less than or equal to 15 sixths (and for which eq 24 is not violated).

nonzero parameters	decrease (1/6)	comment
$x_{13} = 1$ and $x_{34} = 2$	15	feasible: type (iii-2)
$x_{23} = 1$ and $x_{34} = 2$	15	feasible: type (iv-2)
$x_{33} = 1$ and $x_{34} = 1$	15	RHS are integers but this solution is not feasible because by eq 24 we have that $n_3 = 1$ which is impossible because of $x_{33} = 1$

There are no possible choices of more than two nonzero observed parameters with decrease less than or equal to 15 sixths. Consequently, the four types (i-2), (ii-2), (iii-2), and (iv-2) of trees have the same A-value, equal to $(5n - 1)/2 - 5/2$, and this is the greatest value that A can assume among n-vertex chemical trees.

This implies the validity of Theorem 4 for n(mod 4) = 2.

• Case n(mod 4) = 3. The analysis is analogous. We start again with a solution of *AGX* where $x_{13} = 2$ and $x_{34} = 1$. This solution has a decrease of 18 sixths. We will examine the solutions with a decrease less than or equal to 18 sixths and for which eq 24 is not violated.

First we consider the solutions where only one of the observed parameters is different from zero.

nonzero parameter	decrease (1/6)	comment
$x_{12} = 1$	15	impossible in view of eq 20
$x_{22} = 1$	15	impossible in view of eq 20
$x_{34} = 3$	12	impossible in view of eq 23
$x_{44} = 1$	9	impossible in view of eq 20
$x_{44} = 2$	18	feasible: type (i-3)

We then examine the possible choices of two nonzero observed parameters.

nonzero parameters	decrease (1/6)	comment
$x_{13} = 1$ and $x_{33} = 1$	18	RHS are integers but this solution is not feasible because by eq 24 we have that $n_3 = 1$ which is impossible because of $x_{33} = 1$
$x_{13} = 1$ and $x_{34} = 2$	15	impossible in view of eq 25
$x_{13} = 2$ and $x_{34} = 1$	18	feasible: type (ii-3) (starting solution)
$x_{23} = 1$ and $x_{33} = 1$	18	RHS are integers but this solution is not feasible because by eq 24 we have that $n_3 = 1$ which is impossible because of $x_{33} = 1$
$x_{23} = 1$ and $x_{34} = 2$	15	impossible in view of eq 25
$x_{23} = 2$ and $x_{34} = 1$	18	feasible: type (iii-3)
$x_{33} = 1$ and $x_{34} = 1$	15	impossible in view of eq 25

There is only one possible choice of exactly three nonzero observed parameters with a decrease less than or equal to 18 sixths: $x_{13} = 1$, $x_{23} = 1$, and $x_{34} = 1$. This solution is feasible and has a decrease of 18/6 (type (iv-3)).

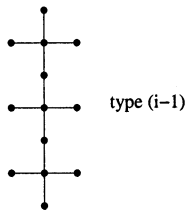

Figure 11. The unique extremal graph for A when $n = 13$.

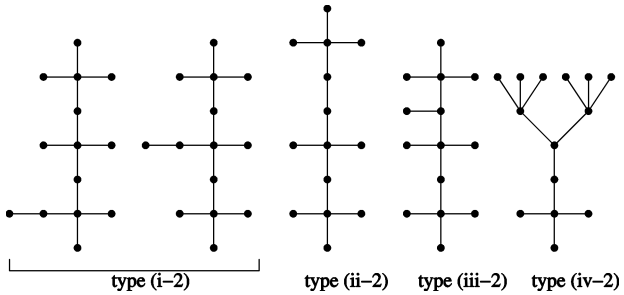

Figure 12. All extremal graphs for A when $n = 14$.

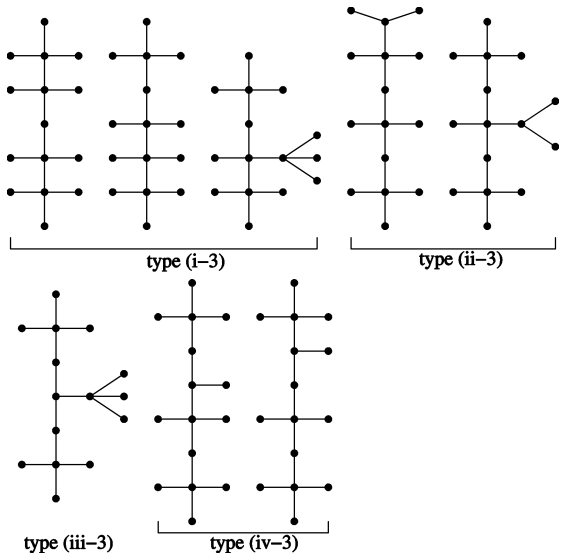

Figure 13. All extremal graphs for A when $n = 15$.

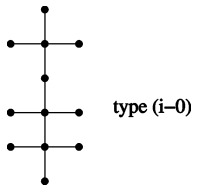

Figure 14. The unique extremal graph for A when $n = 12$.

Consequently, the four types (i-3), (ii-3), (iii-3), and (iv-3) of trees have the same A-value, equal to $(5n - 1)/2 - 3$, and this is the greatest value that A can assume among n-vertex chemical trees.

This implies the validity of Theorem 4 for n(mod 4) = 3.

• Case n(mod 4) = 0. This case is easier because the starting solution (type (i-0)) given by *AGX* has a decrease of 9 sixths ($x_{44} = 1$), and no other solution can be obtained with a decrease less than or equal to 9 sixths without violating eq 24. These trees have $A = (5n - 1)/2 - 3/2$, verifying the respective part of Theorem 4.

Figures 11−14 present all the types of extremal graphs presented in Theorem 4 when $n \in \{12,..,15\}$.

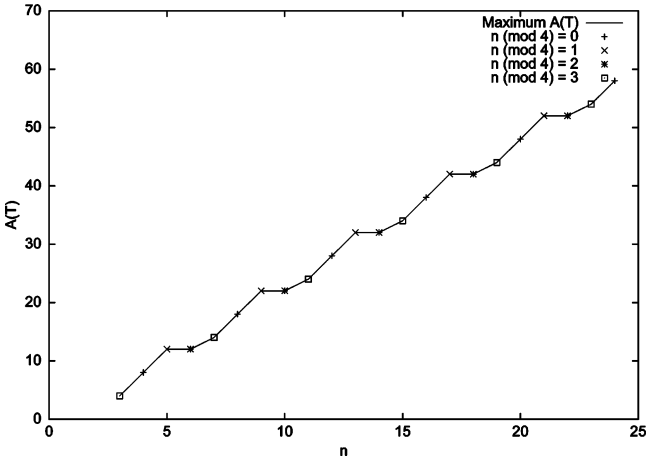

Figure 15. Bounds on $A(T)$ for a chemical tree T of order n.

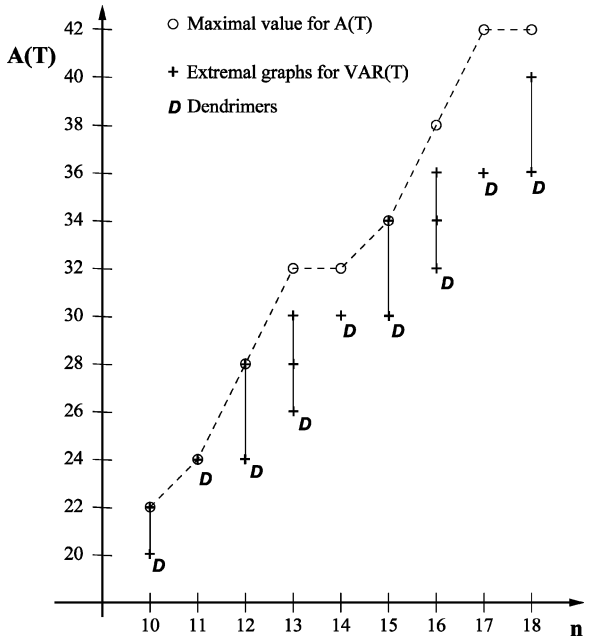

Figure 16. A(T) values for the extremal graphs for VAR(T).

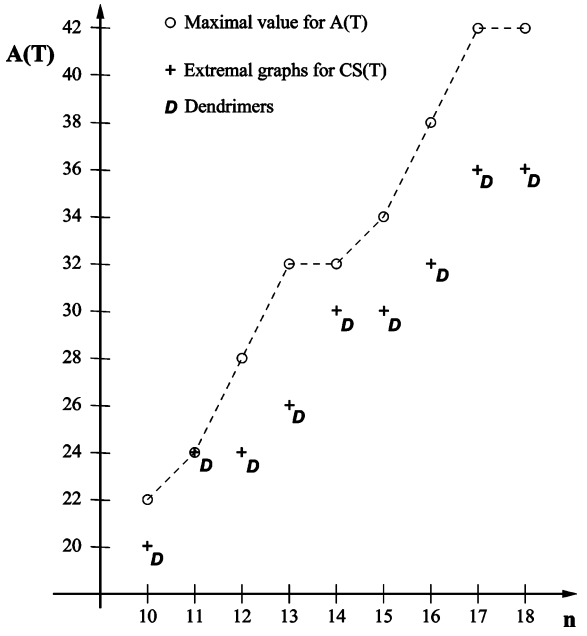

Figure 17. A(T) values for the extremal graphs for CS(T).

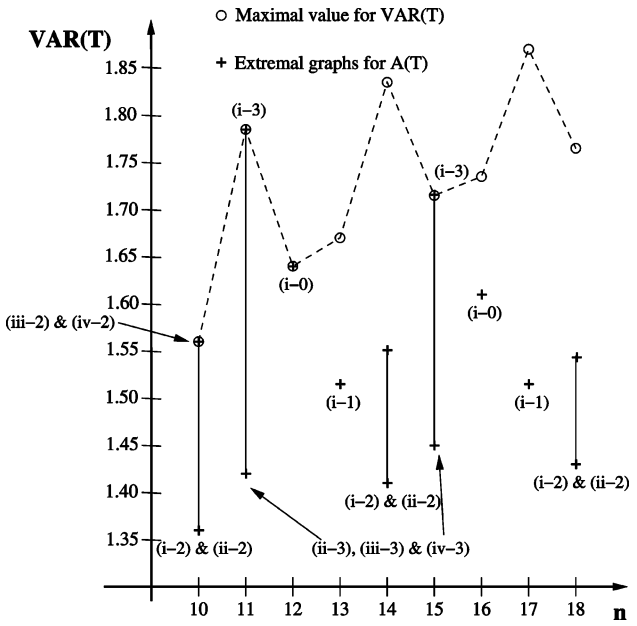

Figure 18. VAR(T) values for the extremal graphs for A(T).

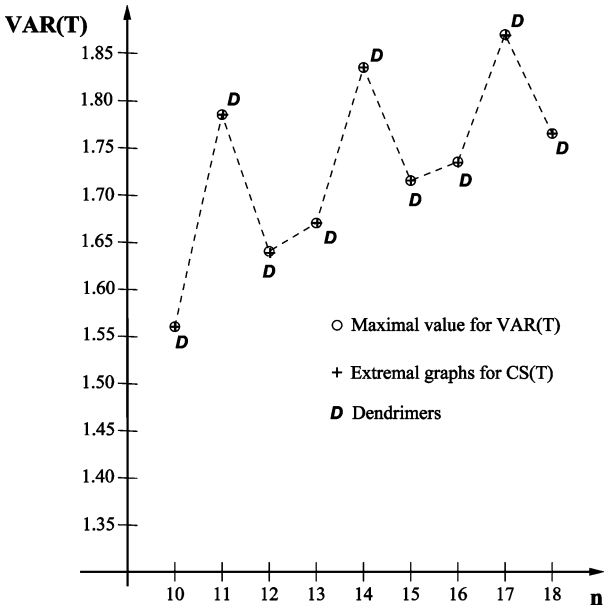

Figure 19. VAR(T) values for the extremal graphs for CS(T).

Figure 15 shows the curve of values defined in Theorem 4.

5. COMPARISON OF THE INDICES

To compare the irregularity indices, we studied how the extremal chemical trees for each index are evaluated according to the two other ones. The results of such observations are shown in Figures 16−21. In each figure we indicate the value of an irregularity index i_1 for chemical trees of order n (where $n \in^{10,18}$), extremal with regard to a different index i_2. The extremal graphs for given n and i_2 are indicated by +, connected by a line, and their i_2 values are represented on the y-axis. The maximum values for i_1 (computed from the previous results) are depicted by circles, linked by a dotted line. Some additional information is added, such as the location of the D_n-dendrimer, indicated by D (this does not imply that the dendrimers are the only extremal graphs with this value in the case of the variance of degrees).

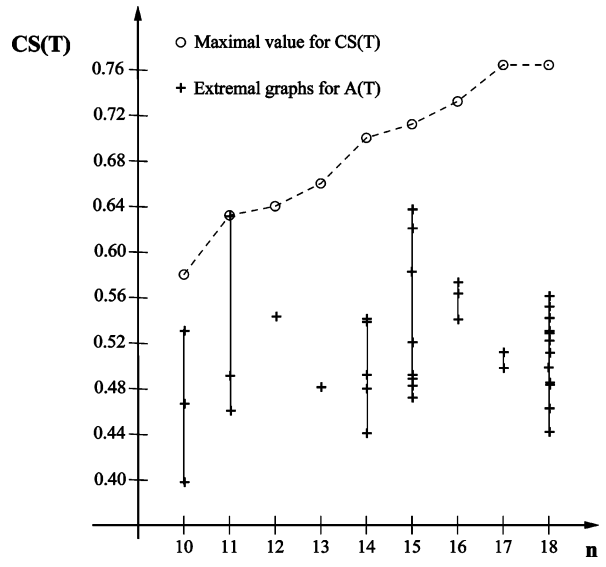

Figure 20. CS(*T*) values for the extremal graphs for A(*T*).

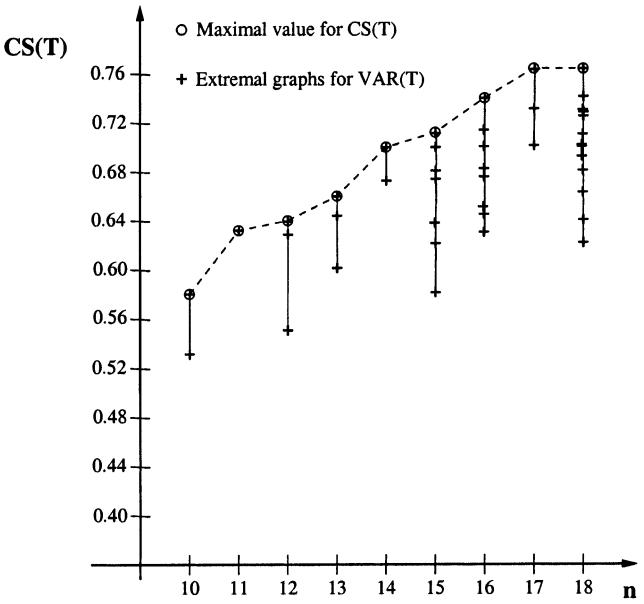

Figure 21. CS(*T*) values for the extremal graphs for VAR(*T*).

Figures 16 and 17 show the A-values of the extremal graphs for VAR(*T*) and CS(*T*). One can see again that the D_n-dendrimers are extremal for VAR(*T*) and CS(*T*). An interesting fact is that the D_n-dendrimers are systematically on the minimum of each interval of values of A(*T*) among the extremal graphs for *VAR*(*T*) (Figure 16). This is due to the fact that degrees are nonincreasing along any path from the root of a D_n-dendrimer to any of its pendant vertices. With some exceptions, maybe due to the fact that *n* is small, the extremal graphs do not attain the maximal A-value.

Figures 18 and 19 present the VAR-values of the chemical trees extremal with regard to A(*T*) and CS(*T*). In Figure 18 the different types of extremal graphs for A(*T*) are represented. Some of these families have the same variance of degrees, which is easy to compute from their characterization. No obvious link can be conjectured between VAR(*T*) and A(*T*). Figure 19 reiterates that the D_n-dendrimers are extremal for VAR(*T*) and CS(*T*).

Figures 20 and 21 show the CS-values of the extremal graphs for A(*T*) and VAR(*T*). With one exception (when $n = 11$), all the extremal graphs for VAR(*T*) have a different CS-value. It is also the case for A(*T*) (for $n = 12$ and $n = 13$, extremal graphs for *A*(*T*) are unique). Comparing differences between optimal CS(*T*) and CS-values for optimal graphs for A(*T*) and for VAR(*T*), it appears that the differences are much larger in the former case (see Figure 20) than in the latter one (see Figure 21).

In summary, A(*T*) (which has a period of 4 in the construction of the families) is apart from the two others (with a period of 3). It seems to express a different notion of irregularity.

6. CONCLUSIONS

The vast majority of chemical graphs is irregular, i.e., not all degrees of their vertices are the same. Clearly, some of these graphs are more irregular than others, hence the interest in irregularity measures. We studied and compared three of them, i.e., the Collatz-Sinogowitz index CS(*T*), the variance of degrees VAR(*T*), and the Albertson index A(*T*), applied to the class of chemical trees. To this effect, extremal graphs were generated with the system *AGX*, conjectures derived and most of them proved by mathematical means.

The families of extremal chemical trees for VAR(*T*) and for A(*T*) are characterized, and the best possible bounds on these indices for all $n \geq 5$ are given.

It appears that CS(*T*) and VAR(*T*) are similar, extremal graphs having a periodicity of 3 in their construction. In contrast extremal trees for A(*T*) have a periodicity of 4 and this index appears to express a different notion of irregularity than the two other ones. It also appears that CS and VAR measure not only irregularity but also the extent of branching. This shows that "branching", a concept which has been extensively studied in the chemical literature, and the much less familiar concept of "irregularity" are not fully independent. This coincidence of the two concepts seems to hold as long as "irregularity" is defined globally, so as to depend on the number of certain structural features of the (molecular) graph but not on their distribution within the graph. The Albertson irregularity index is of a different kind: it depends not so much on the vertex degree distribution but on the differences of the vertex degrees of neighbors—a local structural feature. Thus, the Albertson index A(*T*) reflects properties of (molecular) graphs fully different from "branching" and would therefore deserve due attention in the future.

REFERENCES AND NOTES

(1) *Chemical Applications of Graph Theory*; Balaban, A. T., Ed.; Academic Press: New York, 1976.
(2) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
(3) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; Research Studies Press: John Wiley and Sons: Letchworth, England, 1986.
(4) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, FL, 1983.
(5) Gutman, I.; Polansky, O. E. *Mathematical Concepts in Organic Chemistry*; Springer-Verlag: Berlin, 1988.
(6) *Chemical Graph Theory: Introduction and Fundamentals (Mathematical Chemistry)*; Bonchev, D., Rouvray, D. H., Eds.; Taylor and Francis: 1991.
(7) Fowler, P. W.; Manopoulos, D. E. *An Atlas of Fullerenes*; Oxford University Press: Oxford, 1995.
(8) Balaban, A. T. Topological Indices and Their Uses: A New Approach for Coding of Alkanes. *J. Mol. Struct. (THEOCHEM)* **1988**, *165*, 243−253.

(9) Collatz, L.; Sinogowitz, U. Spektren endlicher Grafen. *Abh. Math. Sem. University Hamburg* **1957**, *21*, 63−77.
(10) Bell, F. K. A Note on the Irregularity of Graphs. *Linear Algebra Appl.* **1992**, *161*, 45−54.
(11) Albertson, M. O. The Irregularity of a Graph. *Ars Combinatoria* **1997**, *46*, 219−225.
(12) Cvetković, D.; Rowlinson, P. On Connected Graphs with Maximal Index. *Publ. Inst. Math. (Beograd)* **1988**, *44*, 29−34.
(13) Caporossi, G.; Hansen, P. Variable Neighborhood Search for Extremal Graphs 1. The AutoGraphiX System. *Discrete Math.* **2000**, *212*, 29−44.
(14) Caporossi, G.; Cvetković, D.; Gutman, I.; Hansen, P. Variable Neighborhood Search for Extremal Graphs 2. Finding Graphs with Extremal Energy. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 984−996.
(15) Cvetković, D.; Simić, S.; Caporossi, G.; Hansen, P. Variable Neighborhood Search for Extremal Graphs 3. On the Largest Eigenvalue of Color-Constrained Trees. *Lin. Multilin. Algebra* **2001**, *2*, 143−160.
(16) Caporossi, G.; Gutman, I.; Hansen, P. Variable Neighborhood Search for Extremal Graphs 4. Chemical Trees with Extremal Connectivity Index. *Comput. Chem.* **1999**, *23*, 469−477.
(17) Caporossi, G.; Hansen, P. Variable Neighborhood Search for Extremal Graphs 5. Three Ways to Automate Finding Conjectures. *Discrete Math.* **2004**, *276*, 81−94.
(18) Hansen, P.; Mélot, H. Variable Neighborhood Search for Extremal Graphs 6. Analysing Bounds for the Connectivity Index. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1−14.
(19) Fowler, P. W.; Hansen, P.; Caporossi, G.; Soncini, A. Polyenes with Maximum HOMO−LUMO Gap. (Variable Neighborhood for Extremal Graphs 7). *Chem. Phys. Lett.* **2001**, *342*, 105−112.
(20) Aouchiche, M.; Caporossi, G.; Hansen, P. Variable Neighborhood Search for Extremal Graphs 8. Variations on Graffiti 105. *Congr. Numer.* **2001**, *148*, 129−144.
(21) Hansen, P.; Mélot, H. Variable Neighborhood Search for Extremal Graphs 9. Bounding the Irregularity of a Graph. To appear in *Graphs and Discovery, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*; Fajtlowicz, S. et al., Eds.; American Mathematical Society: Providence.
(22) Lovász, L.; Pelikan, J. On the eigenvalues of trees. *Periodica Math. Hungar.* **1973**, *3*, 175−182.
(23) Fischermann, M.; Gutman, I.; Hoffmann, A.; Rautenbach, D.; Vidović, D.; Volkmann, L. Extremal Chemical Trees. *Z. Naturforsch.* **2002**, *57a*, 49−52.
(24) Fischermann, M.; Hoffmann, A.; Rautenbach, D.; Székely, L.; Volkmann, L. Wiener Index Versus Maximum Degree in Trees. *Discrete Appl. Math.* **2002**, *122*, 127−137.
(25) Dobrynin, A. A.; Entringer, R.; Gutman, I. Wiener Index of Trees: Theory and Applications. *Acta Appl. Math.* **2001**, *66*, 211−249.
(26) Gutman, I.; Vidović, D. The Largest Eigenvalues of Adjacancy and Laplacian Matrices, and Ionization Potentials of Alkanes. *Ind. J. Chem.* **2002**, *41A*, 893−896.
(27) Randić, M. On characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.
(28) Gutman, I. The energy of a graph: Old and new results. In *Algebraic Combinatorics and Applications*; Betten, A., Kohnert, A., Laue, R., Wassermann, A., Eds.; Springer: Berlin, 2001; pp 196−211.
(29) Gutman, I.; Ruščić, B.; Trinajstić, N.; Wilcox Jr., C. F. Graph Theory and Molecular Orbitals. XII. Acyclic Polyenes. *J. Chem. Phys.* **1975**, *62*, 3399−3405.
(30) Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*, volume 11 of *Methods and Principles in Medicinal Chemistry*; Wiley-VCH: 2000.
(31) Nikolić, S.; Kovacević, G.; Milicević, A.; Trinajstić, N. The Zagreb Indices 30 Years After. *Croat. Chem. Acta* **2003**, *76*, 113−124.
(32) *Topological Indices and Related Descriptors in QSAR and QSPR*; Devillers, J., Balaban, A. T., Eds.; Gordon and Breach: The Netherlands, 1999.
(33) Gutman, I.; Miljković, O.; Pavlović, L. On Graphs with Extremal Connectivity Indices. *Bull. Acad. Serb. Sci. Arts* **2000**, *121*, 1−14.
(34) Fischermann, M.; Hoffmann, A.; Rautenbach, D.; Volkmann, L. A Linear-Programming Approach to the Generalized Randic Index. *Discrete Appl. Math.* **2003,** *128*, 375−385.

CI0342775