MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **59** (2008) 239-252

ISSN 0340 - 6253

Algorithmic generation of molecular graphs with large Merrifield-Simmons index.*

Robert F. Tichy and Stephan G. Wagner

Institute of Mathematics A (Analysis and Computational Number Theory)

Graz University of Technology

Steyrergasse 30, A-8010 Graz, Austria

tichy@tugraz.at, wagner@finanz.math.tugraz.at

(Received April 26, 2007)

Abstract

The Merrifield-Simmons index is related to several physicochemical characteristics and is thus of use in combinatorial chemistry, e.g. in drug design and molecular recognitions. In this paper, we show how one can algorithmically construct databases of acyclic molecular graphs with large Merrifield-Simmons index. Our algorithm can deal with a large number of atoms (several hundreds) in short time by means of a reduction process, and its results lead to a general conjecture on the structure of acyclic molecular graphs with maximal Merrifield-Simmons index. Finally, we show that the algorithm can be modified to deal with another popular index, the Hosoya index.

1 Introduction

Topological indices provide an interesting and powerful tool to study the structure of molecules and their physicochemical properties. Formally, a topological index is merely a map from the set of isomorphism classes of molecular graphs to the real numbers. In

*This work was supported by the Austrian Science Foundation FWF, project S9611.

mathematical chemistry, a vast variety of different indices has been investigated in the past decades, the first prominent example probably being the Wiener index due to the chemist Harold Wiener [26].

Two other very popular instances are the Hosoya index or Z-index (introduced by Haruo Hosoya in 1971, see [9, 10]) and the Merrifield-Simmons index or σ-index (due to R. Merrifield and H. Simmons [15]). The Hosoya index is given as the total number of independent edge subsets (matchings) of a molecular graph, the Merrifield-Simmons index is the total number of independent vertex subsets. Here, a set of edges/vertices is said to be independent if it contains no pair of adjacent edges/vertices (the empty set is counted as an independent set as well). The connections between these indices and various physicochemical characteristics such as boiling points, entropy and heat of vaporization are well studied in several papers—we refer to [6, 18, 20, 23] and the references therein.

In combinatorial chemistry, topological indices are an interesting means of constructing molecular databases which in turn can be used for drug discovery and other purposes (cf. [4, 21, 24]). This is the reason why lots of papers have been written on extremal questions related to these indices (i.e. finding graphs from a prescribed class with large or small index, cf. [3, 5, 11, 13, 14, 27, 28]) and the inverse problem (given a certain index value, construct a graph from a prescribed class with this index value, cf. [4, 12, 25]). Mainly, acyclic systems (in graph-theoretic terminology, trees) and certain classes of molecules involving hexagonal or pentagonal cycles are of interest in this context.

For the Hosoya index and the Merrifield-Simmons index, the trees with maximal or minimal index value are well known (cf. [6, 17] and others); the tree with largest σ-index and smallest Z-index, given the number of vertices, is the star, whereas the tree with smallest σ-index and largest Z-index is the path (s. Figure 1). However, for chemical

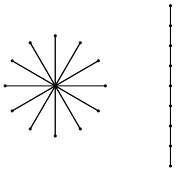

Figure 1: The star and the path.

applications, it is usually necessary to restrict the degree of the vertices—typically, the maximum degree is assumed to be bounded above by 4, the valency of carbon. This poses no particular problems if one is interested in trees with small Merrifield-Simmons index or large Hosoya index, but obviously, the star has larger maximum degree than 4 if the number of vertices becomes larger than 5.

We will call a tree with maximum degree ≤ 4 a *chemical tree*. Not too much is known about chemical trees with large σ- or small Z-index (cf. [3]). In this paper, we are going to present an algorithmic approach which enables us to compute the extremal chemical trees up to a reasonable number of vertices. The computer calculations will also lead us to a general conjecture about how chemical trees with large σ- or small Z-index have to be shaped.

2 Preliminaries

We use the standard graph-theoretic nomenclature—for all notational conventions, we refer to [2]. For a graph G and a vertex $v \in V(G)$, let $\sigma(G)$ denote the Merrifield-Simmons index of G, i.e. the number of independent vertex subsets of G. Furthermore, $\sigma_0(G, v)$ is the number of independent vertex subsets of G not containing v, and $\sigma_1(G, v) = \sigma(G) - \sigma_0(G, v)$ is the number of independent vertex subsets of G containing v. It is very easy to see that we always have $\sigma_0(G, v) \geq \sigma_1(G, v)$. σ, σ_0 and σ_1 satisfy several recursive properties, which we list in the following theorem (cf. [12] for instance):

Theorem 1 *Let T be a tree and v a vertex of T. Denote the components of $T \setminus v$ by $T_1, T_2, \ldots, T_k$ and the neighbors of v by $v_1, v_2, \ldots, v_k$, where v_i belongs to T_i. Then we have*

$$\sigma_0(T, v) = \prod_{i=1}^{k} \sigma(T_i) \quad \text{and} \quad \sigma_1(T, v) = \prod_{i=1}^{k} \sigma_0(T_i, v_i). \tag{1}$$

Let G_1, G_2 be arbitrary graphs with disjoint vertex sets and v_1, v_2 vertices of G_1, G_2 respectively. We construct a new graph $H = (G_1, v_1) \circ (G_2, v_2)$ by taking the union of G_1, G_2 and connecting v_1 and v_2. Then we obtain

$$\sigma(H) = \sigma_0(G_1, v_1)\sigma_0(G_2, v_2) + \sigma_0(G_1, v_1)\sigma_1(G_2, v_2) + \sigma_1(G_1, v_1)\sigma_0(G_2, v_2). \tag{2}$$

The recursive formulas given in the above theorem suggest the use of *rooted trees*. A tree with a distinguished vertex v is called a *rooted tree* and v its root. If $T_1, T_2, \ldots$ and $v_1, v_2, \ldots$ are taken as in Theorem 1, we call the rooted trees T_i the *subtrees* of T and v_i the *children* of v. v_i is regarded as the root of the subtree T_i; in this manner, a recursive structure is imposed on the family of rooted trees. The number of children is called the *outdegree* of a vertex.

Assigning the vector $\mathbf{\Sigma}(T, v) = (\sigma(T), \sigma_0(T, v))$ to every rooted tree, it is possible to compute $\mathbf{\Sigma}(T, v)$ recursively from the values $\mathbf{\Sigma}(T_i, v_i)$ by (1). This also enables us to compute the σ-index of a rooted tree recursively from its subtrees (cf. also [12]).

In particular, *ternary (rooted) trees* will be of interest in our context. A ternary tree may only have up to 3 subtrees, each of which is again a ternary tree. By this means it is

ensured that every vertex has at most degree 4; the degree of the root is at most 3. Every chemical tree can be represented as a ternary tree by choosing a vertex of degree ≤ 3 as the root.

A naïve approach to the problem of finding the chemical trees of given size with maximal σ-index would thus consist of generating all ternary trees together with their σ-indices in a recursive manner and compare. However, this is highly inefficient, since the number of ternary trees of a given size grows exponentially (which is very well known (s. [7])—the investigation of tree counting problems and their chemical applications goes back to Pólya [1, 16]).

Thus, at first, we state a reduction result which greatly reduces the set of candidates for the chemical trees of maximal σ-index:

Theorem 2 *Let T be the tree on n vertices with maximum degree $\leq D$ and maximal Merrifield-Simmons index. Then all vertices of T, except possibly one, have degree 1 or D.*

For the proof of this theorem, we only need a simple lemma:

Lemma 3 Let T be a tree and v,w two different vertices of degree ≥ 2. Furthermore, let T_v be a nonempty component of $T \setminus \{v\}$ not containing w and let T_w be a nonempty component of $T \setminus \{w\}$ not containing v. Then the σ-index increases either if T_v is transferred from v to w or if T_w is transferred from w to v (see Figure 2).

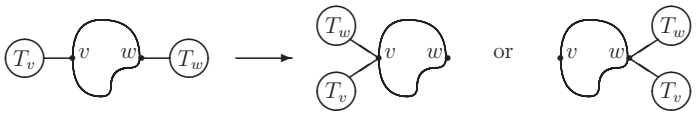

Figure 2: Transferring components of a tree.

Proof: Let

- A be the number of independent subsets of $T \setminus (T_v \cup T_w)$ such that v and w are not contained,
- B_1 be the number of independent subsets of $T \setminus (T_v \cup T_w)$ such that v is contained, but w isn't,

- B_2 be the number of independent subsets of $T \setminus (T_v \cup T_w)$ such that w is contained, but v isn't,
- and C be the number of independent subsets of $T \setminus (T_v \cup T_w)$ such that v and w are both contained.

Clearly, $B_1, B_2 > 0$. Furthermore, let v_1, w_1 be the neighbors of v, w in T_v resp. T_w. We use the notations $x_1 = \sigma(T_v)$, $y_1 = \sigma_0(T_v, v_1)$, $x_2 = \sigma(T_w)$ and $y_2 = \sigma_0(T_w, w_1)$. Then obviously $x_1 > y_1$ and $x_2 > y_2$. Let T_1 be the tree that emerges if T_v is transferred from v to w, and T_2 the tree that emerges if T_w is transferred from w to v. Then we have, by consecutive application of Theorem 1 to v and w,

$$\begin{aligned}\sigma(T) &= x_1x_2A + y_1x_2B_1 + x_1y_2B_2 + y_1y_2C,\\ \sigma(T_1) &= x_1x_2A + x_1x_2B_1 + y_1y_2B_2 + y_1y_2C,\\ \sigma(T_2) &= x_2x_2A + y_1y_2B_1 + x_1x_2B_2 + y_1y_2C.\end{aligned}$$

Now assume, without loss of generality, that $B_1 \geq B_2$ (the other case being symmetric). Then it follows that

$$\sigma(T_1) - \sigma(T) = (x_1 - y_1)(x_2B_1 - y_2B_2) \geq B_2(x_1 - y_1)(x_2 - y_2) > 0$$

and thus $\sigma(T_1) > \sigma(T)$. □

Now, Theorem 2 follows easily: suppose that T has two vertices v, w whose degree is > 1 and $< D$. Then, there exist nonempty components T_v, T_w of $T \setminus \{v\}$ resp. $T \setminus \{w\}$ as in the lemma; both transformations of the lemma leave the maximum degree $\leq D$, since only the degrees of v and w increase resp. decrease by 1, and one of them yields a tree with larger σ-index. This contradiction finishes the proof of the theorem. □

Note that we can now restrict ourselves to the study of ternary trees with the property that all vertices—except the leaves and possibly the root—have outdegree 3. In analogy with the chemical interpretation, we are going to call these ternary trees *saturated*.

Next, we consider an auxiliary problem. Suppose that we have a graph G together with a vertex v. We want to determine a graph G' from a given family of graphs and a vertex v' of G' such that $\sigma(H)$ is as large as possible for the resulting graph $H = (G, v) \circ (G', v')$ (defined as in Theorem 1). By formula (2), this is equivalent to maximizing

$$\begin{aligned}&\sigma_0(G, v)\sigma_0(G', v') + \sigma_0(G, v)\sigma_1(G', v') + \sigma_1(G, v)\sigma_0(G', v')\\ &\qquad = \sigma_0(G, v)\left(\sigma(G', v') + \frac{\sigma_1(G, v)}{\sigma_0(G, v)}\sigma_0(G', v')\right).\end{aligned}$$

Thus, given the fraction $\rho = \frac{\sigma_1(G,v)}{\sigma_0(G,v)}$ (of which we already know that $0 \leq \rho \leq 1$), we only have to maximize $\sigma(G', v') + \rho\sigma_0(G', v')$. Since this paper is mainly concerned with saturated ternary trees, we introduce the notion of α-optimality for this class:

Definition 1 A saturated ternary tree T with n vertices and root v is called α-optimal for $\alpha \in [0,1]$ if

$$\sigma(T) + \alpha\sigma_0(T,v) =$$
$$\max\{\sigma(S) + \alpha\sigma_0(S,w) \,:\, S \text{ is a saturated ternary tree with } n \text{ vertices and } w \in V(S)\}.$$

REMARK: A 0-optimal saturated ternary tree is a tree for which $\sigma(T)$ is maximal, which is exactly what we are looking for.
The described construction principle of joining graphs may be interpreted as some kind of a *fusion* of molecules (Figure 3). Given a certain molecule, we want to fuse it with another one at a certain point in order to maximize the resulting Merrifield-Simmons index.

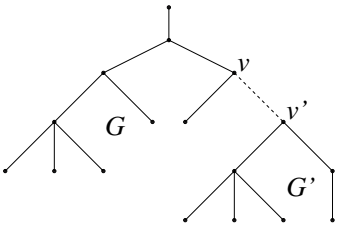

Figure 3: Fusing molecular graphs.

In the following, we will shortly write α-optimal tree instead of "α-optimal saturated ternary tree". One might expect that the subtrees of α-optimal trees are also optimal in some sense. The following observation shows that this expectation is indeed correct:

Theorem 4 *The subtree T_i of an α-optimal tree is always β_i-optimal for some β_i.*

It is not difficult to justify this fact: simply note that

$$\begin{aligned}\sigma(T) + \alpha\sigma_0(T,v) &= (1+\alpha)\prod_{j=1}^{k}\sigma(T_j) + \prod_{i=1}^{k}\sigma_0(T_i,v_i)\\ &= (1+\alpha)\left(\prod_{\substack{j=1\\ j\neq i}}^{k}\sigma(T_j)\right)\sigma(T_i) + \left(\prod_{\substack{j=1\\ j\neq i}}^{k}\sigma_0(T_j,v_j)\right)\sigma_0(T_i,v_i),\end{aligned}$$

which means that T_i has to be β_i-optimal with

$$\beta_i = \frac{1}{1+\alpha}\prod_{\substack{j=1\\ j\neq i}}^{k}\frac{\sigma_0(T_j,v_j)}{\sigma(T_j)}. \tag{3}$$

3 The algorithmic approach

After all the preliminaries in the preceding section, the procedure to determine α-optimal trees (and, in particular, 1-optimal trees) is quite obvious: we start with a tree T consisting of a single vertex for which $\sigma_0(T) = \sigma_1(T) = 1$ and construct larger saturated ternary trees recursively. We only need a tool to filter α-optimal trees. However, this task can be reduced to the well-known calculation of the *upper envelope* of a set of linear functions. Given a set of rooted trees $T_1, T_2, \ldots$ of equal size with roots $v_1, v_2, \ldots$, we consider the linear functions

$$f_i(t) = \sigma(T_i) + t\sigma_0(T_i, v_i).$$

The upper envelope of the functions $f_1(t), f_2(t), \ldots$ is defined as

$$f(t) = \max(f_1(t), f_2(t), \ldots).$$

If the f_i are linear functions, as they are in our case, it can be determined algorithmically in time $O(n \log n)$, where n is the number of functions (cf. [8]). A tree T_i is α-optimal for some α if and only if $f_i(\alpha) = f(\alpha)$. Thus we only have to filter the trees T_i for which $f_i(t) = f(t)$ holds within some interval.
Now, the algorithmic procedure is clear: first, we only consider saturated ternary trees with the property that the root has outdegree 3 (the size of such a ternary tree is always of the form $3m + 1$). Now, we make the following steps to generate lists L_m of all saturated ternary trees on $3m + 1$ vertices:

- Start with the ternary tree consisting only of the root—this is the only element of the first list L_0.
- In the m-th step, consider all possible triples of trees $T_1 \in L_a$, $T_2 \in L_b$, $T_3 \in L_c$ such that $a \leq b \leq c$, $a + b + c = m - 1$. Each of these triples defines a new ternary tree of size $3m + 1$ whose subtrees are T_1, T_2, T_3.
- Each of these new trees is a potential candidate for the list L_m. First, we check whether the optimality conditions of Theorem 4 (β_i-optimality with β_i given by equation (3)) are satisfied for T_1, T_2 and T_3 and some $\alpha \in [0, 1]$ (T_3 might be chosen from a restricted range right away).
- Determine $\mathbf{\Sigma}(T, v) = (\sigma(T), \sigma_0(T, v))$ for the newly generated tree T recursively from $\mathbf{\Sigma}(T_1), \mathbf{\Sigma}(T_2), \mathbf{\Sigma}(T_3)$ according to Equation (1).
- Determine the upper envelope of the functions $\sigma(T) + \alpha\sigma_0(T)$ in the interval $[0, 1]$, where T runs over all generated trees. Select the trees which are optimal within some interval for the list L_m and store the list of interval borders as well.

- Repeat this procedure iteratively.

Now, the trees of maximal Merrifield-Simmons index can be generated easily for any given size n as follows:

- For $n = 3m+1$, Theorem 2 shows that there is exactly one vertex of degree 3. Taking it as the root (thus obtaining a saturated ternary tree), our problem is obviously equivalent to finding a 0-optimal tree.

- For $n = 3m + 2$, all vertices have to have degree 1 or 4, which means that the tree T we are looking for is a saturated ternary tree with one additional leaf attached to the root. If v is this leaf and $T' = T \setminus v$, Lemma 1 shows that $\sigma(T) = \sigma(T') + \sigma_0(T')$, which means that T' is 1-optimal.

- Finally, if $n = 3m + 3$, there is exactly one vertex of degree 2 by Theorem 2. Taking this vertex as the root, the two subtrees have to be α_i-optimal ($i = 1, 2$) for some α_1, α_2. Therefore, it is sufficient to consider all trees that are obtained by attaching two trees from lists L_a, L_b with $a + b = m$ (so that the number of vertices is as desired) to a common root and compare their Merrifield-Simmons indices in order to find the optimal tree.

4 Results and Discussion

4.1 Results

The algorithm given in the preceding section was implemented in C++ as well as in Mathematica (all files are available at `http://finanz.math.tugraz.at/~wagner/molecular`). Additionally, the results of our computations were checked for a number of vertices ≤ 20, making use of G. Royle's database of trees [19]. Here, we would like to show the trees on $25 \leq n \leq 30$ vertices with maximal Merrifield-Simmons index for instance (Figure 4).
The algorithm gives the lists L_m up to $m = 100$ in considerably short time, so one can find the optimal trees up to a size of several hundred vertices easily by means of our algorithm. Note, for comparison, that the number of trees with 100 vertices and maximum degree ≤ 4 is approximately $6 \cdot 10^{39}$. Table 1 lists CPU times for the computations in Mathematica on a 3.4 GHz Windows machine with 1 GB RAM.
The following figure shows a log-plot of the maximal Merrifield-Simmons indices for trees with ≤ 50 vertices—it seems that they grow exponentially with an exponential base of approximately 1.711 (Figure 5).
The structure of the optimal trees leads us to the following conjecture for the general form of the trees with maximal Merrifield-Simmons index.

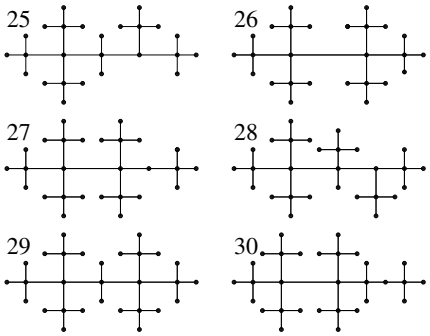

Figure 4: Optimal trees with $25 \leq n \leq 30$ vertices.

n	Number of saturated ternary trees with $\leq 3n+1$ vertices	CPU time for calculating all L_m with $m \leq n$
10	880	< 0.1 sec
25	$1.15 \cdot 10^9$	0.5 sec
50	$6.91 \cdot 10^{19}$	25 sec
75	$6.49 \cdot 10^{30}$	8 min
100	$7.29 \cdot 10^{41}$	2 hrs

Table 1: CPU times required for our computations.

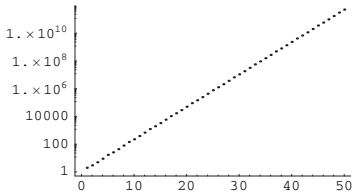

Figure 5: Maximal Merrifield-Simmons indices for trees with ≤ 50 vertices.

Conjecture 1 All optimal trees consist of a path with two subtrees attached to each vertex of the path, except possibly the ends, and these subtrees are complete ternary trees (i.e. all leaves have the same distance to the root), not necessarily of the same size—see Figure 6.

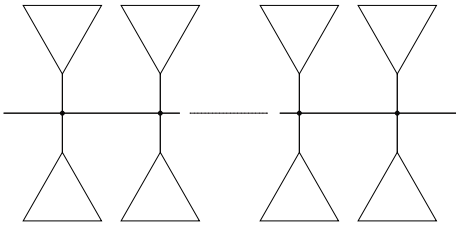

Figure 6: Conjectured form of the optimal trees.

4.2 Complexity

Note that the complexity of the algorithm mainly depends on the length of the lists L_m since all operations on the generated lists can be done within a time bounded by a polynomial of the lengths. Unfortunately, the only estimates we are able to give for these lists are exponential in m. However, the following list of values (Table 2) suggests that the length of L_m increases quite slowly with m.

m	0	1	2	3	4	5	6	7	8	9	10
Length of L_m	1	1	1	2	2	2	4	3	2	3	4
m	15	20	25	30	40	50	60	70	80	90	100
Length of L_m	7	7	5	10	6	12	12	13	9	16	15

Table 2: Lengths of the lists L_m.

4.3 Other Problems

Of course, the idea of our algorithm is not restricted to ternary trees only. It can be adapted to any given maximum degree D (or other classes of trees that can be defined in a similar recursive manner), and even more general problems based on degree restrictions: for instance, suppose we are interested in trees maximizing the Merrifield-Simmons index which have at most one vertex of degree 4 and otherwise maximum degree 3. Then we can take the vertex of degree 4 as the root, and from our considerations it follows that the

subtrees have to be α-optimal for some α among all rooted trees with maximum outdegree 2. Hence all one has to do is to generate a list of all possible trees with this property by means of our algorithm and compare.

Furthermore, the algorithm can also be used for another popular topological index, namely the Z-index or Hosoya index which has already been defined in the introduction. In analogy to the σ-index, one defines $Z_0(G, v)$ as the number of independent edge subsets of G not containing an edge incident with v and $Z_1(G, v)$ as the number of independent edge subsets containing such an edge. Then,

$$Z_0(T, v) = \prod_{i=1}^{k} Z(T_i) \quad \text{and} \quad Z_1(T, v) = \sum_{j=1}^{k} \left(Z_0(T_j, v_j) \prod_{\substack{i=1 \\ i \neq j}}^{k} Z(T_i) \right).$$

Furthermore, if $H = (G_1, v_1) \circ (G_2, v_2)$ as in Theorem 1, we have

$$Z(H) = Z(G_1)Z(G_2) + Z_0(G_1, v_1)Z_0(G_2, v_2).$$

Now, Theorems 2 and 4 hold in an analogous manner when one considers trees of *minimal* Z-index (the (chemical) tree of maximal Z-index is the path):

In the setting of Lemma 3, let this time

- A be the number of independent edge subsets of $T \setminus (T_v \cup T_w)$ such that no edge is incident with v or w,
- B_1 be the number of independent edge subsets of $T \setminus (T_v \cup T_w)$ such that v is incident with one of the edges, but w isn't,
- B_2 be the number of independent edge subsets of $T \setminus (T_v \cup T_w)$ such that w is incident with one of the edges, but v isn't,
- and C be the number of independent edge subsets of $T \setminus (T_v \cup T_w)$ such that v and w are both incident with one of the edges.

With $x_1 = Z(T_v)$, $y_1 = Z_0(T_v, v_1)$, $x_2 = Z(T_w)$, $y_2 = Z_0(T_w, w_1)$ and T_1, T_2 as in Lemma 3, we have

$$\begin{aligned}
z(T) &= x_1x_2(A + B_1 + B_2 + C) + x_1y_2(A + B_1) + y_1x_2(A + B_2) + y_1y_2A, \\
z(T_1) &= x_1x_2(A + B_1 + B_2 + C) + x_1y_2(A + B_1) + y_1x_2(A + B_1), \\
z(T_2) &= x_1x_2(A + B_1 + B_2 + C) + x_1y_2(A + B_2) + y_1x_2(A + B_2).
\end{aligned}$$

If we assume, without loss of generality, that $B_1 \leq B_2$, we obtain

$$z(T) - z(T_1) = (B_2 - B_1)y_1x_2 + y_1y_2A > 0,$$

proving the analogue of Lemma 3 and thus the following theorem:

Theorem 5 *Let T be the tree on n vertices with maximum degree $\leq D$ and minimal Hosoya index. Then all vertices of T, except possibly one, have degree 1 or D.*

We can define α-optimality in the exact same way as for the Merrifield-Simmons index and obtain a statement analogous to Theorem 4 by observing that

$$Z(T) + \alpha Z_0(T, v) = \left(1 + \alpha + \sum_{j \neq i} \frac{Z_0(T_j, v_j)}{Z(T_j)}\right) \left(\prod_{\substack{j=1 \\ j \neq i}}^{k} Z(T_j)\right) Z(T_i) + \left(\prod_{\substack{j=1 \\ j \neq i}}^{k} Z(T_j)\right) Z_0(T_i).$$

The algorithm can now be implemented along the same lines. The results of our implementation (see `http://finanz.math.tugraz.at/~wagner/molecular` again for the files) suggest that chemical trees of maximal Merrifield-Simmons index coincide with those of minimal Hosoya index (indeed, this was checked for a number of vertices ≤ 100), which is a quite typical phenomenon that can be observed in other cases as well (see, for instance, [5, 28]).

Acknowledgement

Special thanks go to Manfred Madritsch for his kind help in programming issues.

References

[1] A. T. Balaban, editor. *Chemical applications of graph theory.* Academic Press [Harcourt Brace Jovanovich Publishers], London, 1976.

[2] R. Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 2000.

[3] M. Fischermann, I. Gutman, A. Hoffmann, D. Rautenbach, D. Vidović, and L. Volkmann. Extremal Chemical Trees. *Z. Naturforsch.*, 57a:49–52, 2002.

[4] D. Goldman, S. Istrail, G. Lancia, A. Piccolboni, and B. Walenz. Algorithmic strategies in combinatorial chemistry. In *Proc. 11th ACM-SIAM Sym. on Discrete Algorithms*, pages 275–284. 2000.

[5] I. Gutman. Extremal hexagonal chains. *J. Math. Chem.*, 12:197–210, 1993.

[6] I. Gutman and O. E. Polansky. *Mathematical concepts in organic chemistry.* Springer-Verlag, Berlin, 1986.

[7] F. Harary and E. M. Palmer. *Graphical enumeration.* Academic Press, New York, 1973.

[8] J. Hershberger. Finding the upper envelope of n line segments in $O(n \log n)$ time. *Inform. Process. Lett.*, 33:169–174, 1989.

[9] H. Hosoya. Topological Index. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bull. Chem. Soc. Jpn.*, 44:2332–2339, 1971.

[10] H. Hosoya. Topological index as a common tool for quantum chemistry, statistical mechanics, and graph theory. In *Mathematical and computational concepts in chemistry (Dubrovnik, 1985)*, Ellis Horwood Ser. Math. Appl., pages 110–123. Horwood, Chichester, 1986.

[11] Y. Hou. On acyclic systems with minimal Hosoya index. *Discrete Appl. Math.*, 119:251–257, 2002.

[12] X. Li, Z. Li, and L. Wang. The inverse problems for some topological indices in combinatorial chemistry. *J. Computational Biology*, 10:47–55, 2003.

[13] X. Li, H. Zhao, and I. Gutman. On the Merrifield-Simmons index of trees. *MATCH Commun. Math. Comput. Chem.*, 54:389–402, 2005.

[14] S. B. Lin and C. Lin. Trees and forests with large and small independent indices. *Chinese J. Math.*, 23:199–210, 1995.

[15] R. E. Merrifield and H. E. Simmons. *Topological Methods in Chemistry*. Wiley, New York, 1989.

[16] G. Pólya. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.*, 68:145–254, 1937.

[17] H. Prodinger and R. F. Tichy. Fibonacci numbers of graphs. *Fibonacci Quart.*, 20:16–21, 1982.

[18] D. H. Rouvray. The search for useful topological indices in chemistry. *American Scientist*, 61:729–735, 1973.

[19] G. Royle's small graphs. Available online at `http://people.csse.uwa.edu.au/gordon/remote/graphs/`.

[20] A. Sabljić and N. Trinajstić. Quantitative structure-activity relationships: The role of topological indices. *Acta Pharm. Jugosl.*, 31:189–214, 1981.

[21] R. P. Sheridan and S. K. Kearsley. Using a genetic algorithm to suggest combinatorial libraries. *J. Chem. Inf. Comput. Sci.*, 35:310–320, 1995.

[22] R. F. Tichy and S. G. Wagner. Extremal Problems for Topological Indices in Combinatorial Chemistry. *J. Computational Biology*, 12:1004–1013, 2005.

[23] N. Trinajstić. *Chemical graph theory*. CRC Press, Boca Raton, FL., 1992.

[24] V. Venkatasubramanian, K. Chan, and J. M. Caruthers. Evolutionary design of molecules with desired properties using the genetic algorithm. *J. Chem. Inf. Comput. Sci.*, 35:188–195, 1995.

[25] H. Wang and G. Yu. All but 49 numbers are Wiener indices of trees. *Acta Applicandae Mathematicae*, to appear.

[26] H. Wiener. Structural determination of paraffin boiling points. *J. Amer. Chem. Soc.*, 69:17–20, 1947.

[27] A. Yu and F. Tian. A Kind of Graphs with Minimal Hosoya Indices and Maximal Merrifield-Simmons indices. *MATCH Commun. Math. Comput. Chem.*, 55:103–118, 2006.

[28] L. Z. Zhang. The proof of Gutman's conjectures concerning extremal hexagonal chains. *J. Systems Sci. Math. Sci.*, 18:460–465, 1998.