

COMPUTER GENERATION OF ISOMERIC STRUCTURES

N. Trinajstić, Ž. Jeričević

The Rugjer Bošković Institute, P.O.B. 1016,
41001 Zagreb, Croatia

J.V. Knop, W.R. Müller, and K. Szymanski

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

Abstract. The algorithms are developed for computer enumeration and generation of isomeric acyclic structures (alkanes, alkenes, alkynes, alkanols, alkanals, alkanones, etc.), benzenoid hydrocarbons, and aza-benzenoids. They are based on the common strategy which consists of representing a given structure by an integers sequence and the use of the induced order notion.

INTRODUCTION

The object of this report is to present a convenient method for computer enumeration and generation of isomeric organic structures. The concept of chemical isomerism, introduced by Berzelius (1), is over 150 years old and, in spite of this, is still being actively pursued by both experimental and theoretical chemists (2). The study of the phenomenon of (organic chemical) isomerism has always been considered as a vital part of theoretical (organic) chemistry (2-3). Therefore, our report should nicely fit into the scope of the Symposium.

The chemical (and mathematical) problem of enumerating (counting) the distinct isomers of a specified composition is of long standing (2,4-34). The problem of generating (and displaying) the isomeric structures is much more recent (35-49) and attempts to solve it are intrinsically connected with the development of the high-speed computer hardware and software (50). The latter problem, counting by generation, is also sometimes referred to as a direct enumeration of isomers (51).

In the present work we will be concerned with the computer enumeration and generation of structural isomers. Molecular structures will be represented by graphs in the standard way (52-54). Thus, we reduce the chemical isomer enumeration problem to counting and displaying the non-isomorphic molecular graphs. Here we will report the isomer count for various acyclic structures (alkanes, C_nH_{2n+2} , alkenes, C_nH_{2n} , alkynes, C_nH_{2n-2} , alkanols, $C_nH_{2n+1}OH$, alkanals, $C_nH_{2n+1}CHO$, alkanones, $C_nH_{2n+2}CO$, etc.) and for some hexagonal planar systems (benzenoid hydrocarbons and aza-benzenoids).

THE OUTLINE OF APPROACH

The general strategy on which our approach rests consists of three steps: (i) The building of a sequence from distinguished elements into an ordered set to represent a given structure; (ii) The representing of a class of sequences assigned to the same structure by different labeling or a class of structures defined by "natural" symmetries by means of finding an

extremal sequence according to the induced order; and (iii) The ordering of the isomeric structures by means of the induced order.

The roots of this approach are in the works by Gordon and Davison (55), by Saunders (56), by Balaban and Harary (57), and by Balaban (58), where these authors attempted to provide the numerical notations for the ring systems.

As an induced order we have chosen the lexicographic order. The criterion used for selecting the well-defined representative of the class is the lexicographic maximum. However, we could have also taken the lexicographic minimum or a different ordering. The isomeric structures are generated in the reverse lexicographic order.

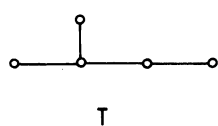
In some cases (e.g. aza-benzenoids) we found that it is more convenient to order the isomers by a different criterion, say by a binary value assigned to each structure.

Acyclic structures

Acyclic hydrocarbons (alkanes, C_nH_{2n+2} , alkenes, C_nH_{2n} , alkynes, C_nH_{2n-2} , etc.) are depicted by (hydrogen-suppressed) trees in which the maximum valency of a vertex is four (52-54). A tree is a connected graph with no cycles (59). We symbolize the tree by T . Acyclic structures containing other atoms besides carbon (alkanols, $C_nH_{2n+1}OH$, alkanals, $C_nH_{2n+1}CHO$, alkanones, $C_nH_{2n+2}CO$, etc.) are depicted by (hydrogen-suppressed) rooted trees in which the maximal vertex is also four. A rooted tree is a tree in which one vertex has been distinguished in some way from others (60). Chemically it corresponds to a substituted carbon. The term hydrogen-suppressed (rooted) trees means that we will consider only bare molecular skeletons.

Trees (and rooted trees) with N vertices may be represented numerically by N -tuples of non-negative integers smaller than N . The N -tuples of (rooted) trees can be produced in the following way. We map the (rooted) trees onto N -tuples of non-negative integers by induction: the trivial (rooted) tree with one vertex is represented by 1-tuple (0). In order to simplify the following discussion we introduce a term the starting vertex for a vertex of a tree at which we start N -tuple. In the case of rooted trees a root-vertex is a point at which N -tuple starts. Thus, a given (rooted) tree with $N > 1$ vertices and M edges incident to the starting vertex (the root-vertex) produces M (rooted) subtrees by removing the starting vertex (the root-vertex) and all of its edges. The (rooted) subtrees (taking as the starting point (the root-vertex) in the subtree the neighbour of the removed vertex) with L_1, L_2, \dots, L_M vertices (where $L_1 + L_2 + \dots + L_M = N - 1$) are by induction provided by L_i subtuples. We concatenate the 1-subtuple (M) and these $N - 1$ subtuples, and get a tuple of $1 + L_1 + L_2 + \dots + L_M = N$ components which we define to be representative for the (rooted) tree.

As an illustrative example we give in Fig. 1 N -tuple representation of a branched tree with 5 vertices which depicts 2-methylbutane. All vertices were used one by one as starting points of the sequences and it is clearly seen that the tree may have several different N -tuples. In order to select a unique N -tuple representation of a given tree we use the concept of lexicographic order. Thus, a K -tuple $(a_1 a_2 \dots a_K)$ of integers is defined lexicographically smaller than L -tuple $(b_1 b_2 \dots b_L)$, if there exists an index j with $1 \leq j \leq L$ so that $a_i = b_i$ for $1 \leq i \leq j$ and $a_j < b_j$ or $K < L$ and $a_i = b_i$ for $1 \leq i \leq N$. The unique N -tuple representation of a tree is the lexicographically highest among all the possible representations



(1 2 1 0 0)
 (3 1 0 0 0)
 (2 2 0 0 0)
 (1 1 2 0 0)

Fig. 1 N -tuples belonging to the tree depicting 2-methylbutane. The lexicographically highest N -tuple is indicated by an arrow.

Among the four different N-tuples representing the carbon skeleton of 2-methylbutane the lexicographically highest N-tuple : (3 1 0 0 0) represents the structure uniquely.

The starting point of N-tuple representing a rooted tree is the root-vertex.



T

(2 3 0 0 1 0)

Two non-isomorphic (rooted) trees cannot have the same N-tuple. Therefore, we state the following theorem.

Theorem: Let T_1 and T_2 be two (rooted) trees with the same associated N-tuple A. Then T_1 and T_2 are isomorphic (rooted) trees (isomorphism being "through" the associated N-tuple A).

Proof: (1) By induction: for $N \leq 4$ the assertion is clear.

(2) Let $\psi_i: T_i \rightarrow A$ be mappings that map each vertex of T_i onto its element in A. These ψ_i 's are 1-1, so we can define $\phi: T_1 \rightarrow T_2$ by $\phi = \psi_2^{-1} \circ \psi_1$, i.e. $\phi(x) = \psi_2^{-1}(\psi_1(x))$ for all $x \in V(T_1)$, ϕ being 1-1 too. Let us prove that ϕ is an isomorphism. Let $\{x, y\} \in E(T_1)$. (Note that $V(T)$ and $E(T)$ stand for the vertex-set and edge-set of T).

(a) Suppose that neither x nor y is the root-vertex of T_1 . Then there exists a subtree L_i of T_1 such that $\{x, y\} \in E(L_i)$. According the definition of N-tuples, the image of L_i is a subtuple of A, $\psi_1(L_i) = A_i$. Let us take a subtree K_i of T_2 and define it by $K_i = \psi_2^{-1}(A_i)$. (It is easy to see that any acyclic graph, associated with a subtuple, is a subtree). Then K_i and L_i are trees with the same associated tuple (namely A_i) and are, according to the inductive assumption, isomorphic with isomorphism ϕ , i.e. $\{\phi(x), \phi(y)\} \in E(K_i) \subseteq E(T_2)$.

(b) Suppose that x is the root-vertex of T_1 . Then y is the root of a subtree L_i . But $\phi(x)$ is the root-vertex of T_2 and according the inductive assumption $\phi(y)$ is the root of a subtree $K_i = \psi_2^{-1}(\psi_1(L_i))$ of T_2 , so that $\{\phi(x), \phi(y)\} \in E(T_2)$, q.e.d. (Because we can replace T_1 by T_2 and vice versa).

The computer programme was developed on the basis of the described method, the details of which are available elsewhere (44). The programme generates the N-tuples representing (rooted) trees in reverse lexicographic order, i.e. starting with the most complicated structure and ending with the chain.

Planar hexagonal structures

Planar hexagonal structures considered are benzenoid hydrocarbons and aza-benzenoids containing one or more nitrogens. The skeletons of these systems are depicted by networks of hexagons which are constructed in the plane by assembling regular hexagons in such a way that (i) two hexagons have exactly one common edge or are disjoint and (ii) the covered area in the plane is simply connected. The number of hexagons h making up a given hexagonal structure is the degree of the structure.

For the numerical representation of a hexagonal structure we use the digits assigned to the vectors covering the edges of the boundary of the structure and construct the sequence following the direction of vectors. The boundary of the hexagonal structure in the plane is a cycle in a graph theoretical sense, because the covered area is simply connected (61). We define two hexagonal structures to be equal if their boundaries are equal in the geometric sense. Since the interior of the hexagonal structure is reconstructable because of the uni-

formity of the interior as far as the boundary is known, we will represent a hexagonal structure only by its boundary and give a numerical description only of the boundary. The interior of a polyhexagonal structure consists of vertices and edges which do not lie on the boundary of the structure (61-62).

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



In Fig. 2 we give the vector description of the boundary of a graph depicting triangulene (phenalene). Graphs depicting benzenoid hydrocarbons are called benzenoid graphs (54,63). We symbolize graphs depicting hexagonal structures by G.

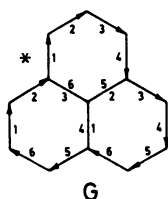


Fig. 2 The vector description of a benzenoid graph G depicting triangulene

We note that the edges of hexagons making up G in Fig. 2, labeled with 1 and 4, 2 and 5, and 3 and 6 coincide in the parent structure. These pairs of edges are said to be inverse to each other.

The boundary of G may be labeled by different sequences of numbers depending on at which edge we start the sequence. Let us start the sequence at the edge of G labeled by the asterisk. The sequence, thus obtained, which represents a numerical notation of G, is given below,

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

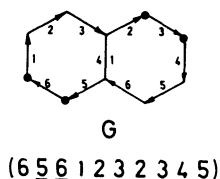
However, it is clear that one may start the sequence at any edge on the boundary. Therefore, the above sequence is not unique. We give below all sequences generated for G by cyclic permutations of the digits on the boundary starting from the edge next to the edge denoted by the asterisk,

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

We choose among all these sequences that one (denoted by an arrow) which represents the lexicographic maximum as the unique representation of G.

The generating algorithm works in an inductive way using the fact that the hexagonal structure of degree 1, i.e. planar six-membered ring, is known. Assuming that all hexagonal structures of degree h are generated, a hexagonal structure of degree $h+1$ can be obtained by joining a new hexagon at the boundary edges of the structure of degree h . In order to generate the whole set of isomeric hexagonal structures of degree $h+1$ obtainable from the given parent structure of degree h , all edges on the boundary of the parent structure must be considered as joining positions with the inverse edges of the added hexagon. If the inverse digits appear in the sequence which are adjacent, they denote inner edges and should be omitted from the sequence.

In the case of aza-benzenoids the edges leading to black dots, representing the positions of nitrogens in the structure, will be underlined in the boundary sequence. Below we give as an example the boundary sequence for 1,2,5,6-tetraazanaphthalene.



If only the vertices of valency 2 in G are considered and if those underlined are labeled by unity and all others by zero, the following binary vector (b_1, b_2, \dots, b_8) representing G is obtained,

$$(1, 1, 0, 0, 1, 1, 0, 0)$$

From this binary vector we can compute the integer characteristic for a given aza-benzenoid, the so-called binary value of G (64),

$$\sum_{i=1}^m b_i 2^{i-1}$$

where m is the number of vertices with valency two in G . Therefore, aza-benzenoids belonging to a given parent polyhex can be ordered according to their binary values instead of using the lexicographic ordering. We could, of course, if we wanted, use the lexicographic order for sorting aza-benzenoids, but it is more convenient to order them by means of their binary values since they are structures whose skeletons consist of only two kinds of atoms, i.e. they represent an example of natural binary structures. Thus aza-benzenoids were generated and ordered (by increased binary value) according to the above approach. Further details about the approach and the description of the computer programme may be found elsewhere (64).

RESULTS AND DISCUSSIONS

Because of the limited space allotted for the article we report here only a few results obtained. However, additional numerical and graphical results may be obtained from the authors on request.

Acyclic structures

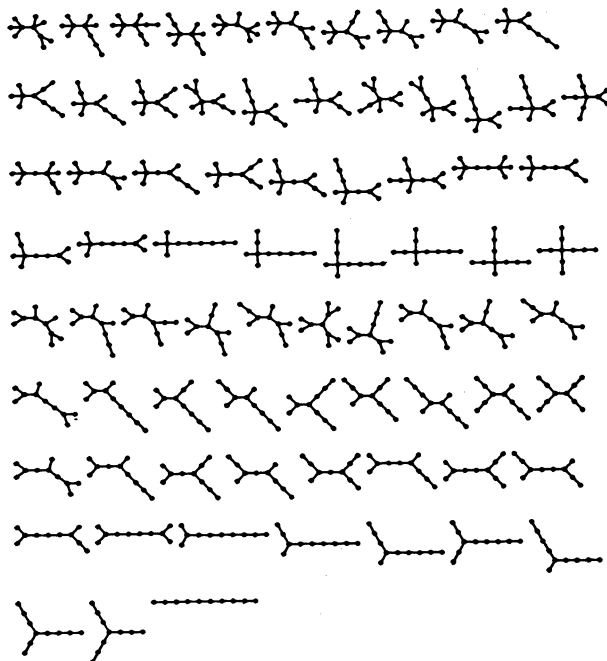
In Table 1 we give the number of isomeric alkanes, alkenes and alkynes with n carbon atoms.

TABLE 1. The number of structural isomers of alkanes, alkenes and alkynes with n carbon atoms

Number of carbon atoms n	Alkanes	Alkenes	Alkynes
	C_nH_{2n+2}	C_nH_{2n}	C_nH_{2n-2}
1	1		
2	1	1	1
3	1	1	1
4	2	3	2
5	3	5	3
6	5	13	7
7	9	27	14
8	18	66	32
9	35	153	72
10	75	377	171
11	159	914	405
12	355	2281	989
13	802	5690	2426
14	1858	14397	6045
15	4347	36564	15167
16	10359	93650	38422
17	24894	240916	97925
18	60523	623338	251275
19	148284	1619346	648061
20	366319	4224993	1679869

In Table 2 we give, as an example, the diagrams for all alkane graphs with 10 vertices.

TABLE 2. Alkane skeletons with 10 atoms (a copy of a computer output)



The total number of isomeric alkyl alcohols (alkanols) for a given number of carbon atoms is reported in Table 3. In this table we also report the isomer count for primary, secondary, and tertiary alcohols, respectively.

TABLE 3. The number of alkanols

Number of carbon atoms n	Alkanols (alkyl alcohols) $C_nH_{2n+1}OH$			
	Primary alcohols	Secondary alcohols	Tertiary alcohols	Total
1	1			1
2	1			1
3	1	1		2
4	2	1	1	4
5	4	3	1	8
6	8	6	3	17
7	17	15	7	39
8	39	33	17	89
9	89	82	40	211
10	211	194	102	507
11	507	482	249	1238
12	1238	1188	631	3057
13	3057	2988	1594	7639
14	7639	7528	4074	19241
15	19241	19181	10443	48865
16	48865	49060	26981	124906
17	124906	126369	69923	321198
18	321198	326863	182158	830219
19	830219	849650	476141	2156010
20	2156010	2216862	1249237	5622109

The number of isomeric alkanals, $C_nH_{2n+1}CHO$, and alkylamines, $C_nH_{2n+1}NH_2$, is identical to the isomeric count for primary alcohols, $C_nH_{2n+1}CH_2OH$. The number of isomeric alkanones, $C_nH_{2n+2}CO$, is identical to the isomer count for secondary alcohols, $C_nH_{2n+2}CHOH$.

The results in Table 1 and Table 3 are checked against the values reported by Read (29) and in all cases of overlap these two sets of values agree.

Though being very efficient for generating all (rooted) trees of a given kind, the method proved to be inadequate when only counts were wanted. For example, in order to obtain only the numbers given by Harary (65) for trees up to 26 vertices, it would have been necessary to spend about 2 000 000 seconds or 24 days of CDC-CYBER 76 CPU-time, as the numbers increase exponentially with a factor of about 2.9 for every next higher number of vertices and the CPU-time augments even faster. Fortunately, it was found after an extensive analysis that the representation by N-tuples provides a possibility for a much cheaper way to get these counts. The method leads to recursive sums of the same form as those obtained from the evaluation of the Pólya's formulae (17), and straightforward optimization by hand leads to algorithms which need for structures up to n vertices storage of the same order n and time of the order $n^2 \log n$ in the worst case. This contrasts pleasantly with the order 1.3^n for the direct evaluation of the recursive functions or the order 2.9^n for the counting by generation. Obviously, trees with very large number of vertices n need multiple precision arithmetics for the counters, and therefore, an additional factor n for storage, and n^2 for time must be then applied. We computed these isomer counts based on our method on TR 445 up to a number n=280 of vertices (we stopped there because higher counts would not fit together with the number n on a single 132 character print line) using only 750 CPU-seconds (equivalent to about 40 CPU-seconds on a CDC-CYBER 76).

Planar hexagonal structures

Benzenoid hydrocarbons were generated up to degree 11. In Table 4 we give the number of benzenoids. We note that here we do not give the number of isomeric structures in the traditional chemical view, but according to number of assembled hexagons (47).

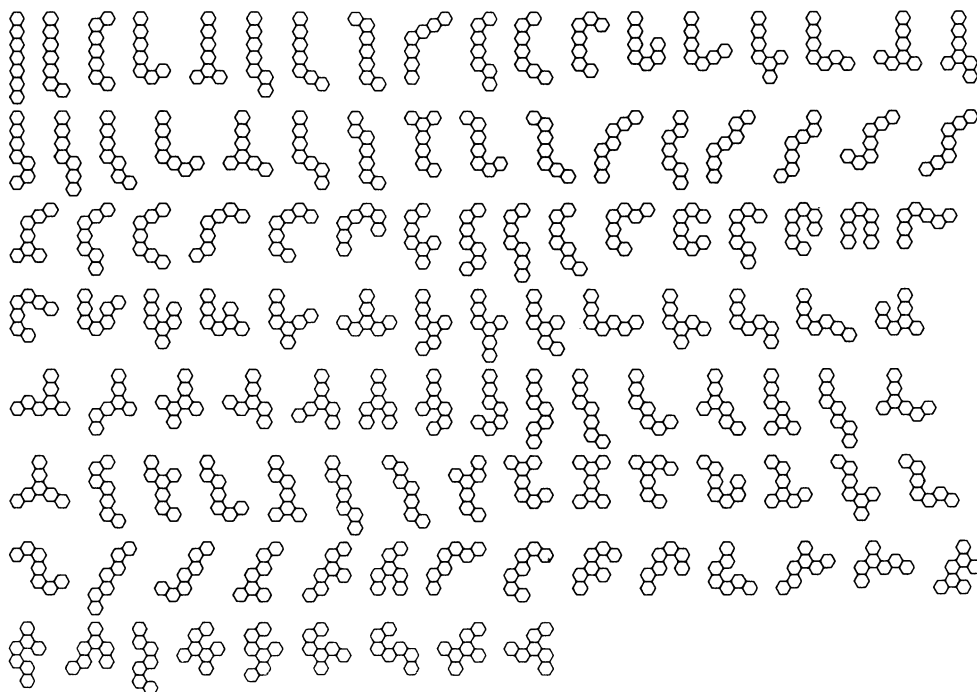
TABLE 4. The number of benzenoid hydrocarbons

Number of hexagons h	Our results		Results from literature	
	Cata-condensed	Total	Cata-condensed ^a	Total ^b
1	1	1	1	1
2	1	1	1	1
3	2	3	2	3
4	5	7	5	7
5	12	22	12	22
6	36	81	37	82
7	118	331	123	333
8	411	1435	446	1448
9	1489	6505	1689	6572
10	5572	30086	6693	30490

^a F. Harary and R.C. Read, Proc. Edinburgh Math. Soc. ser. II 17, 1 (1970).

^b W.F. Lunnon, in Graph Theory and Computing, Edited by R.C. Read, Academic, New York, 1972, p. 87.

TABLE 5. The computer output containing all strictly planar cata-condensed benzenoids of degree 7 sorted in reverse lexicographic order



Since in our work we considered only geometrically planar benzenoids, non-planar structures like helicenes (66) are not included. In addition, we did not include in our study rings of hexagons, i.e. circulenes (47,67-69) either. This makes our total numbers different from those reported by Lunnon for $h \geq 6$, because he considered circulenes. Our numbers for cata-condensed benzenoids differ from those reported by Harary and Read for $h \geq 6$, because they included helicenes in their work.

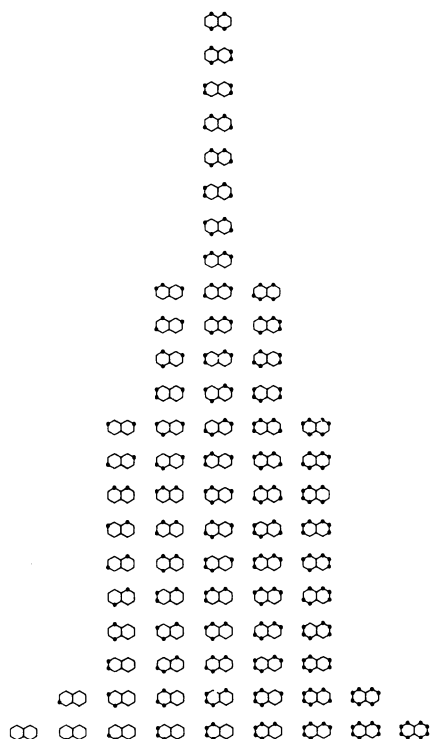
In Table 5 we give as an example a computer output containing all strictly planar cata-condensed benzenoid hydrocarbons with $h=7$.

Here we also report the isomer count for poliaza-benzenoids. In Table 6 we give the number of isomeric aza-benzenoids up to $h=10$ with 1, 2, ..., 8 nitrogen atoms.

TABLE 6. The number of isomeric aza-benzenoids

Number of hexagons h	Number of nitrogen atoms							
	1	2	3	4	5	6	7	8
1	1	3	3	3	1	1		
2	2	10	14	22	14	10	2	1
3	10	48	109	194	216	187	100	42
4	43	243	730	1620	2442	2802	2276	1410
5	210	1326	4918	12982	24611	35384	38500	32326
6	1026	7349	32043	98765	223717	388936	525764	561378
7	5130	41030	204417	720519	1894520	3862098	6229444	8069192
8	25770	228694	1277866	5073805	15169211	35475697	66405417	101090191
9	130350	1270663	7860506	34707100	116213717	306558871	652844870	1141425811
10	661458	7031737	47678819	231739037	8588186640	2521963738	6018548260	11879179287

TABLE 7. The computer output containing all isomeric aza-naphthalenes ordered according their increasing binary values. Black dots denote the positions of nitrogens in the structure.



The example of the computer output containing isomeric aza-naphthalenes is given in Table 7. Other numerical results and the corresponding diagrams may be obtained on request from the Computer Centre of the University of Düsseldorf.

CONCLUSIONS

Here we presented a novel method for computer generation of isomeric structures. It has been applied to some acyclic and cyclic planar structures. The method has several excellent features: (i) It enumerates isomeric structures with high accuracy as any method in literature; (ii) It produces directly graphs of studied structures; and (iii) It may serve as a suitable basis for developing a general algorithm for chemical synthesis by computer. However, we have still to find an algorithm to generate efficiently the stereo-isomers in order to make our approach chemically realistic and competitive to other proposals (29,37-43). Work in this direction is in progress.

Acknowledgements. We are thankful to Professor W.J. Wiswesser (Frederick, MD) for the correspondence on the isomer enumeration problem. We are also grateful to Professor A.T. Balaban (Bucuresti) and Professor M. Randić (Ames) for the correspondence and discussions about the chemical isomerism and its theoretical descriptions. We wish also to thank Professor J.N. Murell (Falmer, Sussex) and Professor R.G. Parr (Chapel Hill) for their interest in this work, encouragement and useful discussions. The mathematical aspects of the approach have been discussed with young and gifted mathematician Dr. Pavel Křivka (Pardubice).

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

REFERENCES

1. J.J. Berzelius, Poggendorffs Ann. Phys. Chem. **19**, 305 (1830).
2. Z. Slanina, Adv. Quantum Chem. **13**, 89 (1981).
3. W. Hückel, Theoretische Grundlagen der Organischen Chemie, Akademischer Verlagsgesellschaft, Leipzig, 1961 (ninth edition); L.N. Ferguson, The Modern Structural Theory of Organic Chemistry, Prentice Hall, Eaglewoods Cliffs, New Jersey 1963; R.L. DeKock and H.B. Gray, Chemical Structure and Bonding, Benjamin/Cummings, Menlo Park, California 1980.
4. A. Cayley, Phil. Mag. **13**, 172 (1875); ibid. **67**, 444 (1874); Ber. **8**, 1056 (1875); Am. J. Math. **4**, 248 (1881).
5. O.J. Lodge, Phil. Mag. **50**, 367 (1875).
6. H. Schiff, Ber. **8**, 1542 (1875).
7. F. Herrmann, Ber. **13**, 792 (1880).
8. F. Tiemann, Ber. **26**, 1595 (1893).
9. M. Delannoy, Bull. Soc. Chim. France **11**, 239 (1894).
10. S.M. Lozanitsch, Ber. **30**, 1917 (1897).
11. A. Goldberg, Chemiker Zeitung **22**, 395 (1898).
12. H. Kauffmann, Ber. **33**, 2131 (1900).
13. M. Trautz, Lehrbuch der Chimie, De Gruyter, Berlin, 1924, Vol. 3, p. 23.
14. J.H. Redfield, Am. J. Math. **49**, 433 (1927).
15. A.C. Lunn and J.K. Senior, J. Phys. Chem. **33**, 1027 (1929).
16. H.R. Henze and C.M. Blair, J. Am. Chem. Soc. **53**, 3042 (1931); ibid. **53**, 3077 (1931).
17. G. Pólya, Acta Math. **68**, 145 (1937).
18. T.L. Hill, J. Phys. Chem. **47**, 253 (1943); ibid. **47**, 413 (1943); J. Chem. Phys. **11**, 294 (1943).

19. R. Otter, Ann. Math. **49**, 583 (1948).
20. E. Rancke-Madsen, Acta Chem. Scand. **4**, 1450 (1950).
21. I.V. Krivoshei, Russ. J. Struct. Chem. **8**, 274 (1967).
22. M.Y. Kornilov, Zhur. Strukt. Khim. **8**, 373 (1967).
23. M.L. Shivar and W.S. DeLoach, J. Elisha Mitchell Sci. Soc. **84**, 367 (1968).
24. W.E. Bennett, Inorg. Chem. **8**, 1325 (1969).
25. A.T. Balaban, Rev. Roumaine Chim. **11**, 1097 (1966); ibid. **15**, 463 (1970); ibid. **17**, 865 (1972); ibid. **17**, 883 (1972); ibid. **18**, 635 (1973).
27. D.H. Rouvray, J. South African Chem. Inst. **26**, 141 (1973).
27. F. Harary and E.M. Palmer, Graphical Enumeration, Academic, New York, 1973.
28. D.H. Rouvray, Chem. Soc. Revs. **3**, 355 (1974).
29. R.C. Read, in Chemical Applications of Graph Theory, Edited by A.T. Balaban, Academic, London, 1976, p. 25.
30. A.T. Balaban, in Chemical Applications of Graph Theory, Edited by A.T. Balaban, Academic, London, 1976, p. 63.
31. K. Balasubramanian, Indian J. Chem. **16B**, 1094 (1978); Ann. New York Acad. Sci. **319**, 33 (1979).
32. R.A. Davidson, J. Am. Chem. Soc. **103**, 312 (1981).
33. H. Dolhaine, Computers & Chemistry **5**, 41 (1981).
34. K. Balasubramanian, Computers & Chemistry **6**, 57 (1982).
35. W.J. Wiswesser, Chem. Eng. News **30**, 3523 (1952); A Line-Formula Chemical Notation, Crowell, New York, 1954.
36. J. Lederberg, Proc. Natl. Acad. Sci. U.S.A. **53**, 134 (1965).
37. J. Lederberg, G.L. Sutherland, D.G. Buchanan, E.A. Feigenbaum, A.V. Robertson, A.M. Duffield, and C. Djerassi, J. Am. Chem. Soc. **91**, 2973 (1969).
38. L.M. Masinter, N.S. Sridharan, J. Lederberg, and D.H. Smith, J. Am. Chem. Soc. **96**, 7702 (1974).
39. R.E. Carhart, D.H. Smith, H. Brown, and N.S. Sridharan, J. Chem. Inf. Comput. Sci. **15**, 124 (1975).
40. Y. Kudo and S.-i. Sasaki, J. Chem. Inf. Comput. Sci. **16**, 43 (1976).
41. Y. Kudo, Y. Hirota, S. Aoki, Y. Takada, T. Tajiri, I. Fujioka, K. Higashino, H. Fujishima, and S.-i. Sasaki, J. Chem. Inf. Comput. Sci. **16**, 50 (1976).
42. J.G. Nourse, J. Am. Chem. Soc. **101**, 1210 (1979); see also J.G. Nourse, J. Chem. Inf. Comput. Sci. **21**, 169 (1981).
43. J.G. Nourse, R.E. Carhart, D.H. Smith, and C. Djerassi, J. Am. Chem. Soc. **101**, 1216 (1979).
44. K. Balasubramanian, J.J. Kaufman, W. Koski, and A.T. Balaban, J. Comput. Chem. **1**, 149 (1980)
45. F. Göbel, J. Comb. Theory B **29**, 141 (1980).
46. J.V. Knop, W.R. Müller, Ž. Jeričević, and N. Trinajstić, J. Chem. Inf. Comput. Sci. **21**, 91 (1981); see also J.V. Knop, W.R. Müller, and N. Trinajstić, QCPE Bull. **1**, 15 (1981).
47. J.R. Dias, J. Chem. Inf. Comput. Sci. **22**, 15 (1982).
48. S.-y. Zhu and J.-p. Zhang, J. Chem. Inf. Comput. Sci. **22**, 34 (1982).
49. J.V. Knop, K. Szymanski, Ž. Jeričević, and N. Trinajstić, J. Comput. Chem., submitted.
50. Network Systems and Software, Infotech State of the Art Report 24, Infotech Information Ltd., Nicholson House, Maidenhead, Berkshire 1975; Information Processing 80, Edited by S.H. Lavington, North-Holland, Amsterdam, 1980.

51. W.J. Wiswesser, private communication (June 2, 1981).
52. I. Gutman and N. Trinajstić, Topics Curr. Chem. **42**, 49 (1973).
53. A. Graovac, I. Gutman, and N. Trinajstić, Lecture Notes in Chemistry, Vol. 4, Springer, Berlin, 1977.
54. N. Trinajstić, Chemical Graph Theory, CRC, Boca Raton, in press.
55. M. Gordon and W.H.T. Davison, J. Chem. Phys. **20**, 428 (1952).
56. M. Saunders, Tetrahedron **23**, 2105 (1967).
57. A.T. Balaban and F. Harary, Tetrahedron **24**, 2505 (1968).
58. A.T. Balaban, Tetrahedron **27**, 6115 (1971).
59. F. Harary, Graph Theory, Addison-Wesley, Reading, Massachusetts, 1972, p. 32.
60. Ref. 59, p. 187.
61. D. Cvetković, I. Gutman, and N. Trinajstić, J. Chem. Phys. **61**, 2700 (1974).
62. B. Džonova-Jerman-Blažič and N. Trinajstić, Computers & Chemistry, in press.
63. I. Gutman, Croat. Chem. Acta **46**, 209 (1974).
64. K. Szymanski, J.V. Knop, and N. Trinajstić, in preparation.
65. Ref. 59, Appendix 3.
66. M.S. Newman and D. Lednicer, J. Am. Chem. Soc. **48**, 4765 (1956).
67. V.D. Hellwinkel, Chemiker Zeitung **94**, 715 (1970).
68. O.E. Polansky and D.H. Rouvray, Math. Chem. (Müllheim/Ruhr) **2**, 63 (1976); ibid. **3**, 97 (1977).
69. I. Agranat, B.A. Hess, Jr., and L.J. Schaad, Pure Appl. Chem. **52**, 1399 (1980).