

Applications of Graph Theory in Chemistry

ALEXANDRU T. BALABAN

Department of Organic Chemistry, Polytechnic Institute, 76206 Bucharest, Roumania

Received March 12, 1985

Graph theoretical (GT) applications in chemistry underwent a dramatic revival lately. Constitutional (molecular) graphs have points (vertices) representing atoms and lines (edges) symbolizing covalent bonds. This review deals with definition, enumeration, and systematic coding or nomenclature of constitutional or steric isomers, valence isomers (especially of annulenes), and condensed polycyclic aromatic hydrocarbons. A few key applications of graph theory in theoretical chemistry are pointed out. The complete set of all possible monocyclic aromatic and heteroaromatic compounds may be explored by a combination of Pauli's principle, Pólya's theorem, and electronegativities. Topological indices and some of their applications are reviewed. Reaction graphs and synthon graphs differ from constitutional graphs in their meaning of vertices and edges and find other kinds of chemical applications. This paper ends with a review of the use of GT applications for chemical nomenclature (nodal nomenclature and related areas), coding, and information processing/storage/retrieval.

INTRODUCTION

All structural formulas of covalently bonded compounds are graphs: they are therefore called molecular graphs or, better, constitutional graphs. From the chemical compounds described and indexed so far, more than 90% are organic or contain organic ligands in whose constitutional formulas the lines (edges of the graph) symbolize covalent two-electron bonds and the points (vertices of the graph) symbolize atoms or, more exactly, atomic cores excluding the valence electrons. Constitutional graphs represent only one type of graphs that are of interest to chemists. Other kinds of graphs (synthon graphs, reaction graphs, etc.) will be mentioned later. This review will try to highlight applications of such graphs in chemistry: graph theory provides the basis for definition, enumeration, systematization, codification, nomenclature, correlation, and computer programming.¹⁻⁷

Chemistry is privileged to be, both potentially and actually, the best documented branch of science. This is mainly due to the facts that most of the chemical information is associated with structural formulas and that structural formulas may be systematically and uniquely indexed and retrieved. In other branches of science, one needs to use words, but in most branches of chemistry, chemists may "talk" with one another by means of formulas only, irrespective of their mother tongue. Of course, one does translate chemical structures into words by means of nomenclature rules, but the time has come when computers can manipulate formulas directly; moreover, chemistry does possess interdisciplinary areas bordering with physics or biology where words are indispensable; one may think about any property or application associated with a chemical structure, and then it is clear that after locating a compound in a database one needs words. One should note, however, that in chemical information words usually come afterward, (the search field being considerably narrowed), whereas in all other fields they come first.

The importance of graph theory (GT) for chemistry stems mainly from the existence of the phenomenon of isomerism, which is rationalized by chemical structure theory. This theory accounts for all constitutional isomers by using purely graph-theoretical methods, which earlier chemists viewed as "tricks with points and lines (valencies)". It is difficult nowadays to realize the awe and wonder this theory caused 125 years ago; the miracle was repeated twice: first when the stereochemistry of carbon compounds was elucidated by Le Bel and Van't Hoff⁸ and then when Werner⁹ extended these ideas to metal complexes.

Defining and finding the constitutional isomers (which correspond to the same given molecular formula) are purely



Alexandru T. Balaban, born in 1931, is Professor of Organic Chemistry at the Polytechnic Institute in Bucharest. He graduated in 1953 and obtained the Ph.D. degree in 1959 with a thesis on $AlCl_3$ -catalyzed reactions (supervisor, C. D. Nenitzescu). He is a corresponding member of the Academy of S. R. Roumania since 1963 and a fellow of the New York Academy of Sciences since 1969. He has published more than 350 papers or book chapters and has edited books on labeled compounds (1969, 1970) and the book "Chemical Applications of Graph Theory" (Academic Press: London, 1976). He is author of six other books, primarily on organic chemistry topics.

graph-theoretical problems. Chemists, however, have used unconsciously graph-theoretical concepts for more than 100 years, just as Molière's "bourgeois gentilhomme" was speaking prose without realizing it.

The problem of isomerism is the real crux of the documentation/retrieval problem. Although molecular formulas can be ordered for indexing purposes according to simple rules (alphanumerically, except for carbon and hydrogen atoms when present), there is no simple and evident means of ordering constitutional graphs with the same numbers and types of vertices, which represent isomers. It is here that more sophisticated techniques of GT may help.

The essence of chemistry is the combinatorics of atoms according to definite rules; hence, the adequate mathematical tools are graph theory and combinatorics, which are closely related branches of mathematics. Until recently, most theoretical chemists viewed mathematics as a tool for processing

(crunching) numerical data, but the present trend toward nonnumerical methods is noticeable, mainly due to the impact of GT.

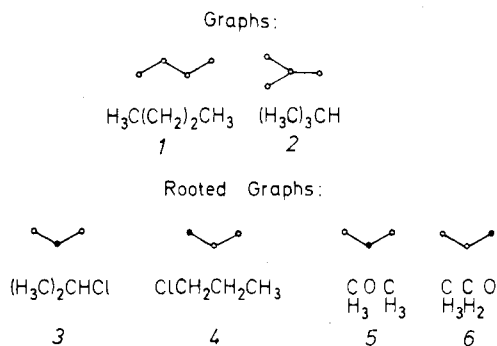
CONSTITUTIONAL AND STERIC ISOMERISM

In graph theory the number of lines meeting at a vertex, i.e., incident to that vertex, is called the *vertex degree*; graphs whose vertex degrees are all equal are called *regular graphs*.

The mathematician Cayley proposed 128 years ago¹⁰ an algorithm for calculating the numbers of constitutional isomers for alkanes, C_nH_{2n+2} , and alkyl derivatives, e.g., $C_nH_{2n+1}Cl$. In graph-theoretical terminology, the latter are called "rooted trees" with the same vertex degree restrictions. Of course, the n vertices of degree 4 correspond to carbon and the $2n + 2$ vertices of degree 1 to hydrogen atoms. Alternatively, one may ignore all hydrogens and look for the equivalent set of trees with n vertices of degrees 1–4; edges now correspond exclusively to C–C bonds, and together with the vertices they form the *hydrogen-depleted graph*, or *skeleton graph*, of the molecule.

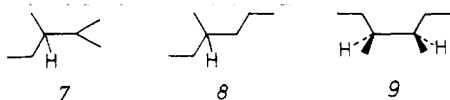
In 1931–1933, the chemists Henze and Blair corrected a few mistakes in Cayley's enumerations and thus perfected an algorithm,¹¹ which nowadays is easily implemented by computer programs for various types of constitutional isomers of alkyl derivatives.¹² A mathematically elegant approach for the same problem was presented by the mathematician Pólya in 1936–1937 when he formulated the important theorem, which since then bears his name.^{13,14}

It may sound a trivial difference, but there is an important distinction between today's formulation (i) that "there exist two isomers each of C_4H_{10} (1 and 2), C_3H_7Cl (3 and 4), and



C_2H_6O (5 and 6) because GT demonstrates¹⁴ that there exist exactly two trees on four vertices and two rooted trees on three vertices," and (ii) the old textbook approach saying "if one tries to write constitutional formulas, one finds only two isomers." The difference is the same as between the correct formulation (i) "the 1st and 2nd principles of thermodynamics assert the impossibility of a perpetuum mobile of 1st and 2nd kind" and (ii) "until now, it has not yet been possible to obtain a perpetuum mobile of 1st or 2nd kind."

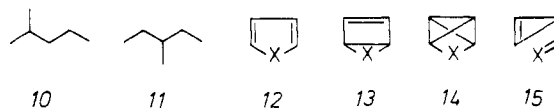
Problems of *stereoisomerism* are also elegantly solved by group-theoretical and graph-theoretical techniques (although graph theory, like topology, ignores geometrical distances and angles) by applying Pólya's¹³ or Otter's theorems¹⁵ to objects with the appropriate (geometrical) symmetry groups. One should, therefore, similarly say "GT demonstrated that (i) there exist two smallest chiral heptanes 7 and 8 where the



chirotopic/stereogenic¹⁶ carbon atoms are shown with a hydrogen atom, and that (ii) the smallest *meso*-alkane is an isomer 9 of octane."¹⁷

VALENCE ISOMERISM

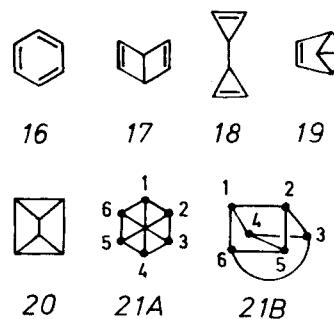
A subclass of constitutional isomers are the valence isomers, defined classically¹⁸ as isomers differing by shifting the positions of simple or double bonds. A clearer definition^{19–21} uses GT concepts: if two hydrogen-depleted graphs have the same partition of vertex degrees, they correspond to valence-isomeric molecules; for vertex-labeled graphs (i.e., molecules containing heteroatoms such as O, N, etc.), vertex degrees associated with each label (atom type) must be conserved. Examples for the smallest alkane valence isomers exist among the five isomeric hexanes: one $Me_2(CH_2)_4$, one $Me_4(CH)_2$, one $Me_4(CH_2)C$, and two valence-isomeric $Me_3(CH_2)_2(CH)$ (10 and 11).



Counterexamples for heteroatom-containing molecules are dimethyl ether (5) and ethanol (6), which are isomeric but not valence isomeric because vertex degrees are not conserved for each atom type separately. There exist four valence isomers of furan, thiophene, or pyrrole, $(CH)_4X$, (12–15), $X = O, S$, or NH ;^{21–23} in these cases when multiple bonding may exist, one should replace the term graphs by multigraphs in order to be more precise.

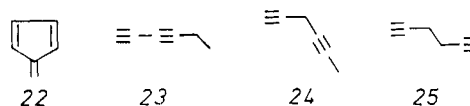
It should be noted from examples 1–15 that the sum of all vertex degrees is identical for isomers; valence isomers are those isomers where the summands are correspondingly equal for each type of atom.

The most interesting case of valence isomerism is provided by annulenes, $(CH)_n$, whose hydrogen-depleted graphs are regular graphs of degree 3 (cubic graphs). A simple graph-theoretical theorem states that the number of vertices in regular graphs of odd degree must be even; therefore, cubic graphs, $(CH)_n$, must have an even n . From the six possible cubic graphs with six vertices, 16–21, the last one, 21, is



nonplanar, i.e., cannot be written on a plane without crossing lines. No such molecule can exist.

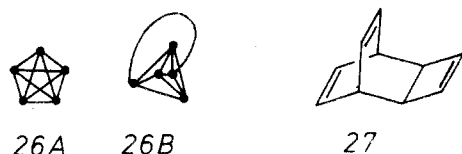
In addition to those six $(CH)_6$ graphs, there are 211 other C_6H_6 graphs^{23,24} that are isomeric (not valence isomeric!) to benzene. They were all depicted in reference 23 and later were included in advertisements of the Jeol Co.²⁵ A few examples of such isomers are fulvene (22) and the three hexadiynes (23–25).



Two graphs are *isomorphic* if there exists a vertex labeling that preserves adjacency; they can be viewed as different geometrical representations of the same abstract graph defined as a set of elements (vertices) $\{v_i\}$, $i \in 1, 2, \dots, n$, and a set of elements (edges) that are unordered duplets from the former set $\{v_i v_j\}$, $i \neq j \in 1, 2, \dots, n$. Thus, 21A and 21B are isomorphic graphs. If the restriction $i \neq j$ is removed, i.e., if one allows a vertex to be connected to itself, one obtains loop graphs; if both loops and multiple edges are allowed, one obtains *general*

graphs having edges $v_i v_i$ (loops, which are counted twice in the degree of vertex i) and multiple edges $v_i v_j, v_i v_j, \dots$ (twice or thrice).

There are two fundamental nonplanar graphs, **21** and **26**, depicted each in two isomorphic representations (A and B).



Kuratowsky demonstrated²⁶ that a necessary and sufficient condition for a graph to be nonplanar is to contain **21** or **26** as a subgraph. If one removes the restriction that a graph edge symbolizes one covalent bond and allows an edge(s) to symbolize a string(s) of several covalent bonds, then *homeomorphic graphs* are obtained. During the last years, several molecules homeomorphic with nonplanar graphs have been obtained.²⁷

Coming back to valence isomers of annulenes, it was shown by a constructive algorithm¹⁹⁻²¹ that there exist two planar cubic multigraphs for $(CH)_4$, five for $(CH)_6$, 19 for $(CH)_8$, 71 for $(CH)_{10}$, and 357 for $(CH)_{12}$. A computer program using an elaborate formula for general cubic multigraphs was developed for checking the above numbers;²⁸ one must calculate separately by recurrence algorithms and subtract the disconnected graphs, the loop graphs, and the nonplanar graphs for arriving at the above numbers. In the formulas that were published for the valence isomers of annulenes,¹⁹⁻²³ benzoannulenes,^{21,29} and homo-^{21,30} or heteroannulenes,³¹ no detailed consideration was given to stereoisomerism so that actually the numbers of possibilities are higher. Exciting experimental discoveries in this field occurred since 1961 when Nenitzescu, Avram, and co-workers³² synthesized the first valence isomer of an annulene, a hydrocarbon $(CH)_{10}$ named by Doering³³ as Nenitzescu's hydrocarbon (**27**). Tetra-*tert*-butyltetrahydrene and its cyclobutadienic valence isomer were prepared by Maier et al.³⁴ and shown to be stable, unlike the less substituted derivatives. All five benzene valence isomers, **16-20**, are known (bicyclopropenyl only as substituted derivatives); Van Tamelen, Kirk, and Pappas prepared Dewar benzene derivatives (a misnomer that has gained large circulation);³⁵ Wilzbach and Kaplan,³⁶ then Katz,³⁷ prepared benzvalene; benzprismane and benzvalene isomerize explosively to benzene.

Bullvalene, $(CH)_{10}$, predicted by Doering and Roth³⁸ and synthesized by Schröder,³⁹ and semibullvalene, $(CH)_8$,⁴⁰ undergo rapid automerizations; cubane, $(CH)_6$, pentaprismane, $(CH)_{10}$, both synthesized by Eaton,⁴¹ and dodecahedrane, $(CH)_{20}$, are spectacular cage-type molecules. The latter was synthesized stepwise by Paquette et al.⁴² and via isomerization by Schleyer et al.⁴³ Many of these valence isomers present considerable interest because they illustrate the brilliant predictions based on the Woodward-Hoffmann rules.⁴⁴

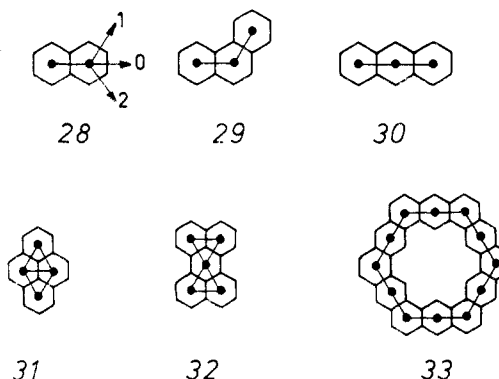
An open problem is the naming of these compounds. As illustrated by the rapidly increasing numbers of planar cubic multigraphs (chemically possible valence isomers of $[n]$ -annulenes) with increasing n , it is not reasonable to coin trivial names (cuneane, hypostrophene, snoutene, etc.); on the other hand, the systematic IUPAC names using the Baeyer system are too cumbersome. It was proposed¹⁹⁻²² to use a string of figures for characterizing such $[n]$ -annulene valence isomers; these figures indicate, in order, the numbers of vertices, double bonds, and three- and four-membered rings and a serial number according to certain conventions. The only drawback of this system, which was adopted by Farnum,⁴⁵ Jones and Scott,⁴⁶ and other chemists interested in valence isomers of annulenes, is that one needs the table of possible valence isomers for retrieving the structure from the notation.

The valence isomers of furan **12-15** can be derived from general cubic graphs by obtaining all such graphs with one loop²² and by assimilating the loop and the vertex to which it is bonded with the heteroatom X. On the other hand, structures of valence isomers of pyridine and other azines may be obtained³¹ by applying Pólya's theorem to valence isomers of benzene.

The most ambitious GT-based computer program for isomer generation developed so far was initiated by Lederberg, Djerassi, Sutherland, Feigenbaum, Duffield, and associates.²⁴ DENDRAL and its heuristic successor programs are able to determine the chemically reasonable structures for a given molecular formula. In conjunction with mass spectrometric and NMR data, this program could enable a completely automatic structure determination for any substance with molecular mass lower than about 1000. The continuing search for extraterrestrial life forms needs such determinations.

CONDENSED POLYCYCLIC AROMATIC HYDROCARBONS (PAH'S)

The classical definitions of cata- and peri-condensed PAH's involve the absence and presence, respectively, of carbon atoms common to three rings.⁴⁷ Though this definition clearly indicates naphthalene (**28**), phenanthrene (**29**), and anthracene

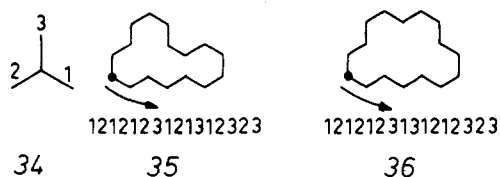


(**30**) to be cata-condensed and pyrene (**31**) or perylene (**32**) to be peri-condensed, it is ambiguous for the recently prepared⁴⁸ kekulene (**33**). A graph-theoretical definition⁴⁹ is based on the dualist graph (the graph having as vertices the centers of benzenoid rings, with two vertices being connected if the corresponding benzenoid rings are condensed): cata-condensed systems have acyclic dualist graphs, peri-condensed systems have dualist graphs possessing only three-membered rings, and corona-condensed systems have dualist graphs possessing larger rings (in this series the last criterion prevails, i.e., peri-condensed systems may have cata-condensed portions, while corona-condensed systems may have both cata- and peri-condensed portions of the molecule).

With this definition, all cata-condensed systems with the same number p of benzenoid rings are isomeric. GT methods provided enumerations and structures of all cata-condensed systems with $p \leq 7$ benzenoid rings,⁴⁹⁻⁵¹ but peri-condensed systems are too complicated and were enumerated only by means of computer programs.⁵²

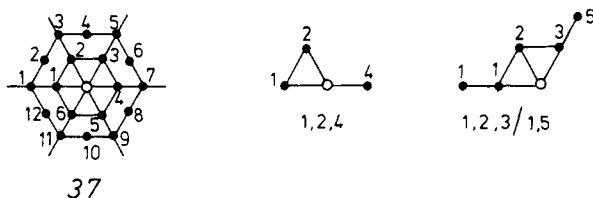
From this dualist-graph approach, one can derive a notation for cata-condensed aromatic systems. Benzenoid polycyclic hydrocarbons (PAH or polyhexes) may undergo cata-annulation according to three directions to afford homologous *catafusenes* as seen for **28**; therefore, a three-digit code (3DC) was proposed with 0 symbolizing straight and 1 or 2 symbolizing kinked annulation (with the proviso that the smallest resulting number is selected from all possibilities).⁴⁹ This notation was extended to branched catafusenes,⁵⁰ and may be generalized to cata-condensed non-benzenoid systems.⁵³ Thus, phenanthrene (**29**) is coded in 3DC by 1 and anthracene by 0.

On the basis of specifying in digital form the three possible orientations (34) of carbon-carbon bonds in the graphite



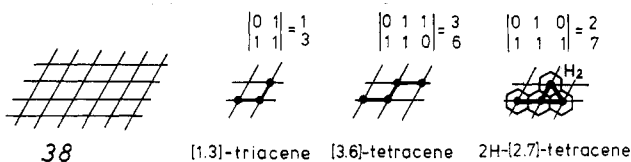
lattice,⁵⁴ one may obtain a coding⁵⁵ of planar annulene geometries (e.g., 35 and 36) that is free from ambiguities existing in other codes.⁵⁶ One selects a vertex and a sense of rotation so that the resulting string of digits forms the minimal number for the given geometry. On the same basis, Trinajstić, Knop, and co-workers⁵⁷ developed algorithms and computer programs for the construction and enumeration of all polyhexes. Dias⁵⁸ has studied the homologous series of polyhexes as functions of their molecular formulas and numbers of benzenoid rings.

Two different approaches have been proposed for the coding of perifusenes: one may specify⁵⁹ the presence of benzenoid rings (vertices of dualist graph) in concentric shells (37) around



the topological center of the polyhex; the center vertex is not specified in the notation; the first shell has numbers 1-6 wherein 1 lies left from the center on a horizontal line; the second shell has numbers 1-12; etc. Shells are delineated by slashes, and the polyhex is oriented so as to minimize the numbers in the first shell according to certain rules, as seen in the two examples beside 37.

Alternatively,⁶⁰ one may superimpose the polyhex on the grid numbered 38 so as to minimize the area of the circum-



scribed parallelogram and to minimize the number of rows and, at equal number of rows, the number of columns.

The presence of a vertex of a dualist graph is read as 1 and the absence as 0, yielding rows of binary numbers that, when translated into decimal numbers, afford a coding of the polyhex.⁶⁰

GRAPH THEORY AND THEORETICAL CHEMISTRY

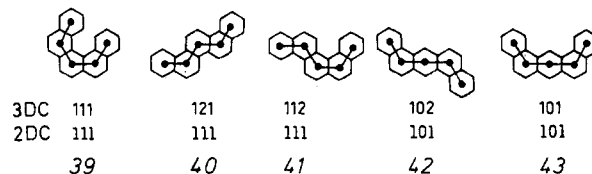
It is a known fact⁶¹ that the simple Hückel MO theory and the VB theory are both topologically based. In particular, the HMO eigenvalues (in β units) are identical with the eigenvalues of the adjacency matrix of the hydrogen-depleted graph.⁶² It is therefore not surprising that much chemical information on cata-condensed systems is encoded in the topology and geometry of their dualist graphs (actually, dualist graphs are exceptional graphs because their angles do matter).

Thus, it was found that the resonance energies and redox potentials are linear functions of two variables: the number p of condensed benzenoid rings and the number s of 180° angles in the dualist graph or of zeroes in the 3DC.^{63,64} The latter number indicates how many anthracenic subunits (which may occur separately or overlap in longer linear portions) exist in the cata-condensed PAH; i.e., phenanthrene has $s = 0$, anthracene and benzo[*a*]anthracene have $s = 1$, tetracene has $s = 2$, etc.

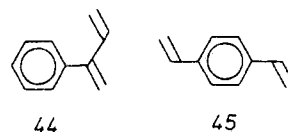
Topological correlations exist also with reactivity data, e.g., with reaction rates for the Diels-Alder cycloaddition of maleic anhydride to cata- and perifusenes; the correlations consider two parameters: the number of benzenoid rings in the longest cata-condensed linear portion and a shape parameter (an integer number expressing the nature of the annelation at the ends of this largest linear portion). The agreement with experiment for more than 100 cata- and perifusenes is better⁶⁵ than in any other correlation using more sophisticated theoretical calculations.⁶⁶

The topology is also important for the carcinogenic activity of PAH's, where the presence of bay regions is responsible for the conversion of an adjacent benzenoid ring into a diol epoxide, probably the ultimate carcinogen. A computer program generates all PAH's, recognizes the presence and number of bay regions, and prints the result and structure of the PAH together with its three-digit code.⁶⁷

By ignoring the difference between digits 1 and 2 in 3DC, i.e., between annelation kinked toward right or left, and by replacing both 1 and 2 by 1 (i.e., by kinked annelation of any kind), one obtains a two-digit notation that no longer codes the complete structure but only its topology: this is called the L transform of the 3DC into a 2DC since the resulting code consisting of 0 and 1 can be still interpreted as a binary number.⁶⁸ Alternatively, Gutman⁶⁹ introduced an analogous description of the topology in terms of the LA sequence (L = linear annelation, A = angular annelation); the sequence of L/A letters is completely equivalent to the sequence of 0/1 letters/digits. Catafusenes having different 3DC but identical 2DC were termed⁷⁰ *isoarithmic* because they give rise to the same number of Kekulé structures. Moreover, many of their spectral or chemical properties are quasi-identical, e.g., electronic and photoelectronic spectra, rates of Diels-Alder cycloaddition with maleic anhydride, etc. Examples of isarithmic catafusenes are tetrahelicene and chrysene, pentahelicene (39), picene (40), and hydrocarbon 41, and the pair of dibenzanthracenes 42 and 43.



The similarity of chemical and physical properties for isarithmic catafusenes demonstrates that the topology encased in the 2DC or the LA sequence is sufficient for determining the frontier MO's. However, none of the isarithmic catafusenes appear to be similar on using the simple MO methods; on the other hand, the known *cospectral* (or *isospectral*) graphs corresponding to molecules that possess the same eigenvalues of the Hückel or adjacency matrices^{71,72} have different physicochemical behavior, as pointed out by Heilbronner.⁷³ Examples of cospectral graphs (the term originates from the equality of the eigenvalues) are 2-phenyl-1,3-butadiene (44) and 1,4-divinylbenzene (45); the corresponding



molecules have different electronic absorption spectra and chemical behavior. These data prove the deficiencies of the simple MO methods and point to the need for a theoretical rationalization.

Dewar⁷⁴ provided a major remedy for one of the deficiencies of the Hückel MO approach by redefining the standard against which resonance energies must be calculated; instead of disjoint

Table I. Total Numbers $K_{1,k}$ of Kekulé Structures and $R_{1,k}$ of Conjugated 6 Circuits (Perfect Benzenoid Rings) and Ratio between Conjugated 6 Circuits and Total Number of Benzenoid Rings for Helicenes and Isoarithmetic Polyhexes

	k				
	1 (28)	2 (29)	3	4 (37-39)	∞
$K_{1,k}$	3	5	8	13	
$R_{1,k}$	4	10	20	40	
$R_{j,k}/[(jk+1)K_{j,k}]$	0.667	0.667	0.625	0.615	0.553

Table II. Numbers $R_{1,k}^{(t)}$ of Conjugated t Circuits in $(1,k)$ Hexes

t	k				
	1	2	3	4	5
6	4	10	20	40	76
10	2	4	10	20	40
14		2	4	10	20
18			2	4	10

ethylene units, one takes connected acyclic graphs with the same number of π -centers and π -electrons. The Dewar resonance energy, originally computed according to Pople's SCF parametrization, was subsequently simplified by making HMO calculations with new standards and parametrizations. The Trinajstić-Gutman topological resonance energy,⁷⁵ the Hess-Schaad parametrization,⁷⁶ Herndon's structure-resonance theory,⁷⁷ and the conjugated circuits model developed by Randić⁷⁸ and independently by Gomes and Mallion⁷⁹ all lead to consistent results. All these newer theoretical approaches have a GT basis and make explicit use of GT: by application of the Sachs theorem, eigenvalues may be readily calculated by recursive formulas; Herndon's revival of valence-bond methods based on Kekulé structure counts benefits from interesting algorithmic⁵⁴ or algebraic⁷⁰ procedures for computing numbers K of Kekulé structures. The conjugated circuits model was applied⁸⁰ to zigzag (j, k) catafusenes having equal numbers j of benzenoid rings in each of their k linear portions (or to all their isoarithmetic isomers): if $j = 1$, one obtains, as known, the Fibonacci sequence, $K_{1,k} = F_{k+2}$ ($F_0 = F_1 = 1$). Interestingly, the ratio between the numbers $R_{j,k}$ of conjugated six circuits ("perfect benzenoid rings", or fully benzenoid rings) having three conjugated double bonds and the total number of six-membered rings, $K_{j,k}(jk+1)$, gives a simple expression when j increases indefinitely:

$$L_j = \lim_{k \rightarrow \infty} (R_{j,k}/[(jk+1)K_{j,k}]) = \frac{1}{j} \left[1 + \frac{j-2}{(j^2+4)^{1/2}} \right]$$

Thus, phenanthrene (29) has $K_{1,2} = 5$ Kekulé structures with a total of $R_{1,2} = 10$ perfect (fully) benzenoid rings. Other values are seen in Table I. Values $K_{j,k}$ for $k > 1$ may be considered to be generalized Fibonacci numbers.

It was established⁸⁰ that the sequence of conjugated t circuits $R_{1,k}^{(t)}$ in (j,k) hexes is the same for $t = 10$ as for $t = 6$ but shifted as seen in Table II; e.g., for picene and its isoarithmetic (j,k) hexes with $k = 4$ and five six-membered rings, the number of fully aromatic rings in all 13 Kekulé structures (conjugated 6 circuits) is 40, the number of conjugated 10 circuits is 20, etc. These data allow the easy calculation of resonance energies of such (j,k) catafusenes according to Randić's parametrization modified so as to consider all conjugated circuits, not only the linearly independent ones.⁷⁸

HEURISTIC COMBINATORICS FOR MONOCYCLIC AROMATIC SYSTEMS

In accordance with Pauli's principle, sp^2 -hybridized atoms that form a planar aromatic ring may have 2, 1, or 0 electrons in the nonhybridized p_z orbital; therefore, they are of exactly three types: X, Y, or Z, respectively. If the ring is M sized

Table III. First Row Atoms That May Form Aromatic Rings Together with "Aromaticity (Electronegativity) Constant" (k_H) (in Parentheses) for the Case $R = H$ (Otherwise, $k_H = k_R + 20\sigma_{p,R}$)

type	group ^a					
	13	4	14	5	15	6
X			$\begin{array}{c} \cdot\cdot \\ \text{C} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{N} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{N} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{O} \\ \\ \text{R} \end{array}$
			(-100)	(-77)	(-26)	(-3)
Y	$\begin{array}{c} \cdot\cdot \\ \text{B} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{C} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{C} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{N} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{N} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{O} \\ \\ \text{R} \end{array}$
	(-72)	(-50)	(0)	(23)	(74)	(97)
Z	$\begin{array}{c} \cdot\cdot \\ \text{B} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{C} \\ \\ \text{R} \end{array}$	$\begin{array}{c} \cdot\cdot \\ \text{C} \\ \\ \text{R} \end{array}$			
	(28)	(50)	(100)			

^aIn this paper the periodic group notation is in accord with recent actions by IUPAC and ACS nomenclature committees. A and B notation is eliminated because of wide confusion. Groups IA and IIA become groups 1 and 2. The d-transition elements comprise groups 3 through 12, and the p-block elements comprise groups 13 through 18. (Note that the former Roman number designation is preserved in the last digit of the new numbering: e.g., III $\rightarrow$ 3 and 13.)

and if only first row atoms are involved (Table III), Hückel's $4N + 2$ π -electrons rule holds; therefore, for an aromatic $X_xY_yZ_z$ ring we have

$$x + y + z = M$$

$$2x + y = 4N + 2$$

For given M and N values, one can construct all possible monocyclic aromatic systems by solving algebraically and then combinatorially either directly⁸¹ or by means of Pólya's theorem⁸² the above system of two equations with three unknowns x , y , and z (which admits several solutions). For given x , y , and z values, we arrive at the typical combinatorial problems known as "the necklace problem" in GT, which consists of finding how many nonisomorphic necklaces one can form with given numbers x , y , and z of differently colored beads. Table IV presents all solutions for a π -electron sextet with no adjacent Z-type atoms. The combinatorial problem with adjacency restrictions was solved by Lloyd.⁸³

The complete set of solutions evidences many unknown systems, especially with boron atoms. However, some of them are too unstable because of unsuitable electronegativities or because they can dissociate into stable molecules (e.g., hexazine $\rightarrow 3N_2$).⁸⁴ A semiempirical measure of electronegativities for this case was proposed,⁸⁵ assigning to each atom type in Table III a constant; summation of these constants over the whole ring should give a sum within certain limits; otherwise, the ring is unlikely to be aromatic and stable.

Consideration of Table IV leads to a generalized definition of mesoaromatic systems indicated by an asterisk (mesoionic compounds are a particular case): aromatic systems having odd numbers of Y-type atoms separated by X- and/or Z-type atoms or chains of such atoms are called mesoaromatic systems. The carbonyl carbon atom behaves as a Z-type atom. So far, many five-membered mesoionic systems have been prepared⁸⁶ and a few 1,3-diboretes⁸⁷ (which are more stable than 1,2-diboretes), but six-membered systems should also be possible; several such systems can be seen in Table IV, and by combination of this table with Table III, many possibilities are generated.

GRAPH THEORY, TOPOLOGICAL INDICES, AND QSAR

For quantitative structure-activity relationships, which are important especially for drug design, one looks for correlations

between physical, chemical, or biochemical properties (possessing a continuous variation and being expressed numerically) and chemical structures. Only a few chemical features may be expressed numerically: electronic effects of substituents (Hammett's σ constants), steric parameters (Taft's or related constants), and hydrophobicity (Hansch's partition coefficient). Usually, one has to compare closely related molecules in order to obtain meaningful correlations.

A different approach was initiated by H. Wiener in 1947⁸⁸ with topological indices (TI's). These TI's are numbers associated with chemical structures via their hydrogen-depleted graphs G . For hydrocarbons, the Wiener index W is the sum of the number of bonds between all pairs of vertices in G . If one defines the (topological) distance between two vertices of a graph as the number of bonds along the shortest path between these two vertices, then W is the sum of all distances in graph G . One can associate with any graph on n vertices several matrices: the adjacency matrix $A(G)$ is a square, symmetrical, $n \times n$ matrix with entries $a_{ij} = 1$ for adjacent (directly bonded) vertices i and j and zero otherwise; the distance matrix $D(G)$ is also a square $n \times n$ matrix with entries $d_{ij} = 0$ on the main diagonal and $d_{ij} = 1$ for adjacent vertices i and j as in A , but all other entries are integers bigger than 1 and represent the topological distance between vertices i and j . The sums over rows i or columns i for $A(G)$ indicate the vertex degrees v_i ; the sums over rows i or columns i for $D(G)$ indicate another graph invariant for each vertex (invariant from the arbitrary vertex labeling $i \in \{0, 1, \dots, n\}$), called distance sum s_i . It is easy to see that $W = \sum s_i/2$.

More than 40 other TI's were devised till now.⁸⁹⁻⁹² From them, Randić's molecular connectivity χ ,^{93,94} Hosoya's⁹⁵ index Z , and the recently introduced⁹⁶ index J (distance sum connectivity) appear to be the most useful:

$$\chi = \sum_{\text{edges}} (v_i v_j)^{-1/2}$$

$$Z = \sum_{k=0}^s p(G, k)$$

$$J = q \sum_{\text{edges}} (d_i d_j)^{-1/2} / (q - n + 2)$$

In the above formulas, $\sum_{\text{edges}}$ stands for summation over all q edges. By multiplying the vertex degrees or the distance sum of end points for each edge of graph G on n vertices, this may be chosen so that no two of them have a common vertex: by definition, $p(G, 0) = 1$ and $p(G, 1) = q$, the number of edges; s is the maximum number of edges disconnected to each other.

A general feature of all topological indices is their degeneracy, i.e., the fact that two or more different graphs have the same value of a TI. The degeneracy is high for W and Z , lower for χ , and very low for J . For reducing degeneracies, Bonchev, Mekenyan, and Trinajstić devised information theoretical indices^{97,98} and proposed a superindex⁹⁹ consisting of a sum of 10 different TI's.

There exist correlations among various TI's.⁹⁹ The systematic exploration of molecular properties that are more closely associated with TI's revealed that the molecular refractivity and the molecular volume are such properties; therefore, TI's may be useful parameters for expressing steric features of molecules or substituents.¹⁰⁰

Various TI's have different correlating abilities with physical, chemical, or biological properties of molecules. Thus, W and Z account well for thermodynamic properties of alkanes. Centric indices,⁹¹ which were not discussed above, correlate well with octane numbers of alkanes but have high degeneracies.¹⁰¹ However, it was with Randić's index that most correlations were made so far, especially for biochemical properties and drug design.⁹⁴ The presence of heteroatoms

along with carbon atoms requires special treatment, which is available till now only for indices χ ⁹⁴ and J .¹⁰²

SEARCH FOR GRAPH INVARIANTS

The problem of cospectral graphs is connected to the search for simple graph invariants, i.e., means of characterizing graphs up to isomorphism. It had been conjectured that the set of eigenvalues or the characteristic polynomial of a graph (obtained from the adjacency matrix by inserting x on the main diagonal and by solving the resulting determinant) is such an invariant;¹⁰³ however, Balaban and Harary⁷¹ provided counterexamples of cospectral hydrogen-depleted graphs. Then, it was again asserted that if hydrogens are included, a graph invariant may result,¹⁰⁴ but Herndon¹⁰⁵ proved the fallacy of this idea. Thus, so far, short of the adjacency matrix or equivalent data, a graph cannot be characterized by simple graph invariants. For establishing whether two graphs are or are not isomorphic, no satisfactory algorithm exists for which the time required should increase less fast than polynomially vs. the number of vertices.

Since the adjacency matrix $A(G)$ characterizes completely a graph G , Randić¹⁰⁶ searched a standard form for presenting $A(G)$. By permuting rows and columns of $A(G)$ and by reading the rows sequentially, one obtains various binary numbers; if one selects the smallest number, then one really has a graph invariant and a binary notation of the graph structure. However, there exist two problems: namely, (i) the long search for the minimal binary number that is equivalent to the search for a vertex labeling (however, Randić gave several rules for shortening this search) and (ii) the possibility of local minima¹⁰⁷ that prevent one from reaching the real minimum (it appears, however, that for constitutional graphs such cases are unlikely and that triple permutations of rows or columns may provide a way out from local minima).¹⁰⁸ Each vertex of a graph may be characterized¹⁰⁹ by a set of integers representing the numbers of paths with various lengths to the other vertices in the graph (a path is a continuous sequence of graph edges so that all vertices are distinct). An efficient algorithm and computer program exist for obtaining the path numbers for graph vertices. Another idea due to Randić is to combine the topological index χ with the path numbers in order to obtain a molecular identification number, ID.¹¹⁰ However, for multigraphs, the ID does present degeneracies. It is conjectured¹¹¹ that if one adopts a similar approach based on distance sums and path numbers instead of vertex degrees and path numbers, the degeneracy would disappear. Further attempts to devise practically nondegenerate topological indices are based on hierarchically ordered extended connectivities (see below). Finally,¹¹¹ one may combine just two TI's, namely, J summed with an index based on the numbers of second (once removed) neighboring vertices.

REACTION GRAPHS

Constitutional graphs are not the only types of graphs that are of interest for chemistry. If, in order to overview the successive intermediates in multistep reactions one depicts by a point (vertex) each intermediate and by a line (edge) an elementary reaction step (the conversion of one intermediate into the next), one obtains a reaction graph. The first reaction graph to be published described¹¹² the intramolecular isomerization of a pentasubstituted ethyl cation, 46, with five

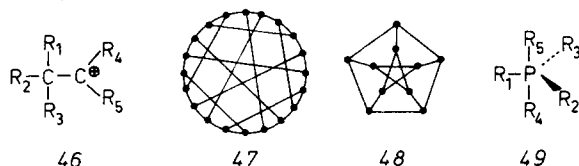


Table IV. Arrangements of Atom Types X, Y, or Z for Monocyclic Aromatic Compounds and π -Electron Sextet with No Adjacent Z-Type Atoms^a

m=4	$\begin{array}{c} X-X \\ X-Z \\ (3,0,1) \end{array}$	$\begin{array}{c} X-X \\ (2,2,0) \end{array}$
m=5	$\begin{array}{c} X-X \\ Z-X-Z \\ (3,0,2) \end{array}$	$\begin{array}{c} (1,4,0) \end{array}$
m=6	$\begin{array}{c} X-X \\ X-Z \\ (2,2,1) \end{array}$	$\begin{array}{c} (Type A) \\ (Type B) \end{array}$
m=6	$\begin{array}{c} Z-X \\ Z-X-Z \\ (3,0,3) \end{array}$	$\begin{array}{c} Z-X \\ (4,1,1) \end{array}$
m=6	$\begin{array}{c} Z-X \\ Z-X-Z \\ (2,2,2) \end{array}$	$\begin{array}{c} (0,6,0) \end{array}$
m=7	$\begin{array}{c} Z-X \\ Z-X-Z \\ (2,2,3) \end{array}$	$\begin{array}{c} (0,6,1) \end{array}$
m=7	$\begin{array}{c} Z-X \\ Z-X-Z \\ (1,4,2) \end{array}$	
m=8	$\begin{array}{c} Z-X \\ Z-X-Z \\ (2,2,4) \end{array}$	$\begin{array}{c} (0,6,2) \end{array}$
m=8	$\begin{array}{c} Z-X \\ Z-X-Z \\ (1,4,3) \end{array}$	

^a Mesoaromatic compounds are indicated by an asterisk, and Y-type atoms are not shown explicitly in order to avoid crowding the formulas.

different graphs and with the two carbon atoms distinguishable (e.g., by isotopic labeling). If, however, there is no means of discriminating the two carbon atoms, the graph reduces from a 20-vertex graph (47) to a 10-vertex graph 48 (the Petersen graph or the five cage), which was also discussed by Dunitz and Prelog¹¹³ in the context of intramolecular rearrangements for five-coordinated trigonal bipyramidal (TBP) species, ignoring their stereoisomerism. If stereoisomerism is considered, the rearrangement of TBP complexes is described by graph 48, which was used in the context of phosphoranes, 49, by Ramirez and co-workers.¹¹⁴ The isomorphism between the pentasubstituted phosphorane graph and the pentasubstituted ethyl cation graph (all substituents being different) was pointed out by Mislow.¹¹⁵ In describing intramolecular rotations, Mislow used reaction graphs. A spectacular success of reaction graphs was in providing a plausible structure for an intermediate product in the isomerization leading to diamantane, which was discovered and investigated by Schleyer and co-workers.¹¹⁶ The corresponding reaction graph has >40 000 vertices, whereas that of the related isomerization affording adamantane involves a reaction graph¹¹⁷ consisting of 16 vertices and 2897 possible reaction paths. For coping with the former graph, a computer program was devised in order to select at each step the energetically most favored path.¹¹⁶ One should note in this context that the systematic enumeration of "diamond hydrocarbons" is the analogue in three dimensions of the two-dimensional catafusene problem and can be solved by applying the dualist-graph approach.¹¹⁸

Further uses for reaction graphs were for intramolecular rearrangements of octahedral complexes or for degenerate

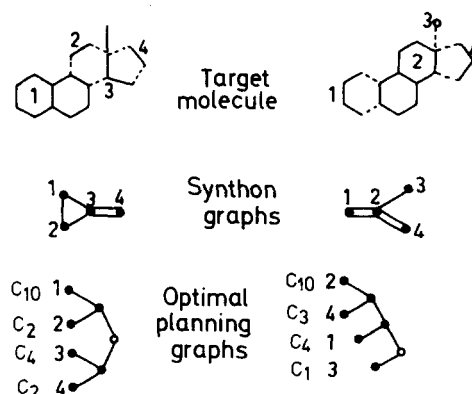


Figure 1. Synthon graphs (second row) for the same target molecule (first row) dissected differently. The order in which the four synthons 1-4 are assembled together is indicated in the bottom row; the number of carbon atoms in each synthon is written to the left of the synthon notation (1-4) on the bottom row, where the target molecule is denoted by the white point.

rearrangements of carbenium ions. In the field of organometallic chemistry, Gielen, Brocas, and Willem used GT and other mathematical tools for rationalizing experimental data.¹¹⁹ New permutational instruments (double cosets) were developed by Ruch,¹²⁰ Klemperer,¹²¹ Ugi, and co-workers.¹²²

In some instances, the number of isomers is so large that only local properties of the reaction graph can be investigated. This is the case of the bullvalene automerization graph,¹²³ which has $10!/3 = 1\,209\,600$ vertices if enantiomerism is taken into account and 604 800 vertices if not.

SYNTHON GRAPHS

Deinton's saying "Chemistry is an art with a bonus"¹²⁴ is nowhere more valid than in organic synthesis. Woodward mastered the intricacies of this art in a hitherto unequalled manner, but Corey and Wipke¹²⁵ inaugurated a new era, which adds the computer's memory and speed to the chemist's insight. Computer-aided design of organic syntheses looks for ways to dissect a molecular graph (the target molecule) into simple, commercially available starting materials, able to react via known reactions to form the target molecule.¹²⁶ The term *synthon graphs* was first used by Hendrickson.¹²⁷ Synthon graphs have vertices symbolizing synthons and edges symbolizing bond sets, i.e., assemblies between synthons. An example is presented in Figure 1, where the steroid skeleton is assembled from four synthons; the bond set is denoted by broken lines and the synthon by solid lines in the upper row.

Nowadays a rich bibliography exists in this field, and sophisticated computer programs have been developed and implemented.¹²⁸⁻¹³⁰ The field is continuously expanding and may eventually reach the heuristic stage when new potential reactions will be uncovered by these computer programs; this trend was adumbrated more than 20 years ago¹³¹ when lists of electrophiles and nucleophiles were matched matrixwise. It was noted on that occasion that whenever protolytic reactions were possible, these occurred preferentially converting the stronger acid-base pair into the weaker one before the coupling of the electrophile with the nucleophile. Thus, whereas the Gattermann-Koch reaction occurs via formyl cation and aromatic compound (and not via carbon monoxide and arenonium cation), the related reaction in the alkene series occurs via carbon monoxide and alkylcarbenium ion (and not via formyl cation and alkene).

GRAPH THEORY AND CHEMICAL INFORMATION

The preceding sections contained occasional information on coding and naming chemical structures. The present section

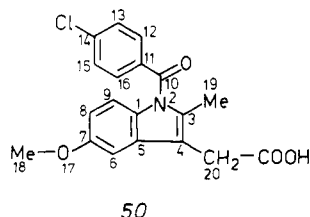
will deal exclusively with problems concerning chemical information such as nomenclature, coding, storage, and retrieval of chemical structures.

Chemical nomenclature has grown in a haphazard way, sanctifying trivial names and collecting several systems that were widely used, rather than being based upon a logical foundation and on first principles. The policy of the IUPAC Commission on the Nomenclature of Organic Chemistry and of similar nomenclature commissions for inorganic or macromolecular chemistry as well as of the International Union of Biochemistry is to assemble and simplify the existing practice of chemical nomenclature, so as to obtain fewer exceptions and a uniform system. However, as stated by Fletcher,¹³² "it appears unlikely that the present rules will ever become systematic." Historical introductions to chemical nomenclature problems were published by Verkade¹³³ and by Goodson,¹³⁴ and a bibliography is available.¹³⁵ One should mention that the 1892 International Congress in Geneva created a Commission for the Reform of Chemical Nomenclature and that nowadays inorganic chemistry (red book) and organic chemistry (blue book) have a set of IUPAC rules¹³⁶ that are consistent for the most part with the practice of Chemical Abstracts Service (CAS).¹³⁷

A proposal for using graph theory for the purpose of chemical nomenclature ("nodal nomenclature") was made by Lozac'h, Goodson, and Powell.^{138,139} Read and Milner¹⁴⁰ had listed the attributes of an ideal nomenclature system, but they developed actually a coding procedure (see below). The nodal names¹³⁸ have a unique numbering of all non-hydrogen and nonsubstituent atoms (nodes) of the graph, unlike the classical nomenclature wherein each moiety (module) has its own numbering. Each node is assigned a unique number used in the descriptor, i.e., the part of the name enclosed in brackets. Modules may be cyclic or acyclic sets of nodes, to be treated as separate entities; a descriptor starting with zero belongs to a cyclic module; otherwise, the module is acyclic. A few examples will be analyzed. *n*-Hexane is called [6]hexanodane, while 3-methylpentane is called [5.1³]hexanodane. In the latter case, the sum of numbers inside the brackets is the number of carbon atoms, i.e., 6, and the upper index 3 indicates which of the five atoms of the main chain is connected to the one-atom side chain. Bicyclo[2.2.2]hexane is called bicyclo-[06.2^{1,4}]octanodane, and indane is called bicyclo[09.0^{1,5}]nonanodane.^{138,139}

By means of the set of seniority rules, and of several examples,¹³⁸ it is quite easy to become familiar with the nodal nomenclature. The advantages of this system are its generality, the consistent treatment of rings, chains, or combinations thereof, systematic numbering of atoms, possibility of application to coordination compounds and to cyclophanes, and its computer compatibility.

The full nodal name of **50** is 14-chloro-10-oxo-2-aza-17-oxatricyclo[(09.0^{1,5})2:10(1)10:11(06)7:17(2)3:19(1)4:20(1)]icosan(1,3-9,11-16)axene-2-carboxylic acid. The modules



are separated by colons, and numbers outside parentheses are locants. The double bonds are indicated individually (e.g., 1) or collectively in conjugated systems (e.g., 3-9, or 11-16) as a prefix to the word axene; the —COOH, —Cl, or =O group as well as heteroatoms is indicated as replacement terms. The descriptor from the nodal name is comparable with the con-

nection table of the CAS system (see below) and can be used for substructure search.

Coding and linear notation systems are older than the nodal nomenclature. In Read's approach,¹⁴⁰ the coding of acyclic chemical graphs is achieved by finding the graph center and by writing chemical formulas either as clusters, delineated by parentheses, attached to the center, or as a linear chain of clusters. No special notation is needed because the code uses normal chemical symbols and, in addition, periods for simple bonds or colons for double bonds in clusters such as carbonyl. The coding is implemented by a computer program both for cyclic and acyclic graphs having a connectivity table as input data (atoms and bonds). No sub- or superscripts are used. Thus, 2-methyl-1-propanol becomes CH(CH₃)₂(CH₂.OH), and malonic acid is CH₂(C=O.OH)₂.

The oldest linear notation system was imagined by Gordon and co-workers,¹⁴¹ but it did not enjoy a widespread use; Dyson's system¹⁴² was recommended by IUPAC, but the most popular is Wiswesser Line Notation (WLN).¹⁴³ All these systems rely heavily on conventions; e.g., the WLN for hydroquinone is QR DR. At present, there exist computer programs that can convert WLN into connectivity tables, and vice versa.¹⁴⁴ Normally, chemical line notations such as WLN are convenient tools for manual or machine registration and substructure searching of files with moderate size (10⁴–10⁵ structures).

The CAS Registry File uses connection tables developed by Morgan.¹⁴⁵ Morgan's algorithm is based on the connectivity of each atom in the hydrogen-depleted graph and on the extended connectivity of its adjacent non-hydrogen atoms. The extended connectivity allows a better discrimination between atoms than the simple connectivity (which for carbon atoms can have only the four values 1–4).

At present, more than 7 million structures are available for CAS online search. Some important features are that one can make substructure search (manually, such a search may be extremely complicated and time consuming with the CA *Chemical Substance* or *Formula Indexes*) and that questions can be compounded with logical symbols (AND, OR, NOT). An impressive effort was invested in computer programs for generating structure diagrams from connection tables either on a computer screen or on paper (photocomposer or electrostatic printer output).^{146,147}

So far, only constitutional and no stereochemical aspects were taken into account. Wipke and Dyott¹⁴⁸ improved Morgan's system so as to include stereochemical information; for implementing Wipke's et al. SECS (Simulation and Evaluation of Chemical Synthesis) computer program, one needs from the outset all stereochemical information about the target molecule and the starting materials. It was for this purpose that the SEMA program (Stereochemical Extension of Morgan's Algorithm) was developed. A simple approach codes configurations at stereocenters (enantiomerism), diastereomerism of double bonds or ring structures, and even conformations around single bonds.¹⁴⁸

Among other systems for structure storage and retrieval that are in widespread use, we shall mention Dubois' DARC system based upon the idea of focal coding;¹⁴⁹ the Institute for Scientific Information¹⁵⁰ makes available this database and also Index Chemicus Online. Less widely used systems were developed by Fugmann,¹⁵¹ by Ugi and co-workers,¹⁵² by the All-Union Institute for Scientific and Technical Information in Moscow and the Mathematics Institute in Novosibirsk,¹⁵³ and by Moreau.¹⁵⁴ Balaban, Mekenyan, and Bonchev, in a series of papers,¹⁵⁵ elaborated hierarchically ordered extended connectivities (the HOC algorithm) for coding the constitution and stereochemistry of molecules, for canonical atom numbering, and for developing on the basis of this numbering new

topological indices a unique topological representation, correlations with ^1H NMR spectra in PAH's, and computer programs for the implementation of this algorithm; it improves the Morgan algorithm and enables more compact storage of structure information in the computer memory. A somewhat similar but less extensive approach for inorganic structures such as polyhedral clusters, including a unique linear notation, was proposed by Herndon and co-workers.¹⁵⁶

REFERENCES AND NOTES

- Balaban, A. T., Ed. "Chemical Applications of Graph Theory"; Academic Press: London, 1967.
- Balaban, A. T.; Rouvray, D. H. In "Applications of Graph Theory"; Wilson, R. J.; Beinecke, L. W., Eds.; Academic Press: London, 1976.
- Rouvray, D. H. *Chem. Soc. Rev.* **1974**, *3*, 355; *Endeavour* **1975**, *34*, 28; *RIC Rev.* **1971**, *4*, 173; *J. Chem. Educ.* **1975**, *52*, 768; *Chem. Br.* **1974**, *10*, 11.
- Trinajstić, N. "Chemical Graph Theory"; CRC Press: Boca Raton, FL, 1983.
- Balaban, A. T. *Math. Chem.* **1975**, *1*, 33.
- King, R. B., Ed. "Chemical Applications of Topology and Graph Theory". *Stud. Phys. Theor. Chem.* **1983**, *23*.
- Biggs, N. L.; Lloyd, E. K.; Wilson, R. J. "Graph Theory 1736-1936"; Clarendon Press: Oxford, 1977; Chapter 4, p 55.
- Le Bel, J. A. *Bull. Soc. Chim. Fr.* **1874**, *22*, 337. Van't Hoff, J. H. *Arch. Neerl. Sci. Exactes Nat.* **1874**, *9*, 445.
- Werner, A. Z. *Anorg. Chem.* **1893**, *3*, 267.
- Cayley, A. *Philos. Mag.* **1857**, *13* (1), 172; **1859**, *18*, 374; **1874**, *47*, 444; **1877**, *3*, 34; *Rep. Br. Assoc. Adv. Sci.* **1875**, *45*, 257; *Ber. Dtsch. Chem. Ges.* **1875**, *18*, 1056.
- Henze, H. R.; Blair, C. M. *J. Am. Chem. Soc.* **1931**, *53*, 3042, 3077; **1933**, *55*, 680; **1934**, *56*, 157. Blair, C. M.; Henze, H. R. *J. Am. Chem. Soc.* **1932**, *54*, 1098, 1538. Coffman, D. D.; Blair, C. M. *J. Am. Chem. Soc.* **1933**, *55*, 252.
- Knop, J. V.; Müller, W. R.; Jericević, Z.; Trinajstić, N. *J. Chem. Inf. Comput. Sci.* **1981**, *21*, 91. Kornilov, M. Yu. *Zh. Strukt. Khim.* **1967**, *8*, 373. Davis, C. C.; Cross, K.; Ebel, M. *J. Chem. Educ.* **1971**, *48*, 675.
- Pölya, G. Z. *Kristallogr.* **1936**, *93*, 415; *Acta Math.* **1937**, *68*, 145.
- Harary, F. "Graph Theory"; Addison-Wesley: Reading, MA, 1969. Harary, F.; Robinson, R. W. "Graphical Enumeration"; Academic Press: New York, 1971.
- Otter, R. *Ann. Math.* **1948**, *49*, 583.
- Mislow, K.; Siegel, J. *J. Am. Chem. Soc.* **1984**, *106*, 3319.
- Robinson, R. W.; Harary, F.; Balaban, A. T. *Tetrahedron* **1976**, *32*, 353.
- Maier, G. "Valenzisomerisierungen"; Verlag Chemie: Weinheim, FRG, 1972.
- Balaban, A. T. *Rev. Roum. Chim.* **1966**, *11*, 1097.
- Balaban, A. T.; Banciu, M. *J. Chem. Educ.* **1984**, *61*, 766. Banciu, M.; Popa, C.; Balaban, A. T. *Chem. Scr.* **1984**, *24*, 28.
- Balaban, A. T.; Banciu, M.; Ciorba, V. "Annulenes, Benzo-, Homo-, Hetero-Annulenes and Their Valence Isomers"; CRC Press: Boca Raton, FL, 1985.
- Balaban, A. T. *Rev. Roum. Chim.* **1970**, *15*, 463.
- Balaban, A. T. *Rev. Roum. Chim.* **1973**, *18*, 635.
- Lederberg, J. In "The Mathematical Sciences"; Boechm, G. A. W., Ed.; MIT Press: Cambridge, MA, 1969; p 37. Lederberg, J. *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *53*, 134. Lederberg, J.; Sutherland, G. L.; Buchanan, B. G.; Feigenbaum, E. A.; Robertson, A. V.; Duffield, A. M.; Djerassi, C. *J. Am. Chem. Soc.* **1969**, *91*, 2973. Lindsay, R. K.; et al. "Applications of Artificial Intelligence for Organic Chemistry: the DENDRAL Project"; McGraw-Hill: New York, 1980. Carhart, R. E.; Smith, D. H.; Brown, H.; Sridharan, N. S. *J. Chem. Inf. Comput. Sci.* **1975**, *15*, 124.
- Jeol advertisement, reverse front cover of *J. Am. Chem. Soc.* **1973**, *95*, issues 15, 16, and 18.
- Kuratovsky, K. *Fundam. Math.* **1930**, *15*, 271.
- Simmons, H. E., III; Maggio, J. E. *Tetrahedron Lett.* **1981**, 287. Benner, S. A.; Maggio, J. E.; Simmons, H. E. *J. Am. Chem. Soc.* **1981**, *103*, 1581. Paquette, L. A.; Vazeux, M. *Tetrahedron Lett.* **1981**, 291. Walba, D. M.; Richards, R. M.; Haliwanger, R. C. *J. Am. Chem. Soc.* **1982**, *104*, 3215. Walba, D. M. In reference 6, p 6. Hisatome, M.; Kawajiri, Y.; Yamakawa, K.; Iitaka, Y. *Tetrahedron Lett.* **1979**, 1777. Hisatome, M.; Kawajiri, Y.; Yamakawa, K.; Harada, Y.; Iitaka, Y. *Tetrahedron Lett.* **1982**, *23*, 1713.
- Balaban, A. T.; Vancea, R.; Holban, S.; Moțoc, I., submitted for publication.
- Balaban, A. T. *Rev. Roum. Chim.* **1974**, *19*, 1185. Banciu, M.; Balaban, A. T. *Chem. Scr.* **1983**, *22*, 188.
- Balaban, A. T. *Rev. Roum. Chim.* **1974**, *19*, 1611.
- Balaban, A. T. *Rev. Roum. Chim.* **1974**, *19*, 1323.
- Avram, M.; Sliam, E.; Nenitzescu, C. D. *Liebigs Ann. Chem.* **1960**, *636*, 184.
- Doering, W. v. E.; Rosenthal, J. W. *J. Am. Chem. Soc.* **1966**, *88*, 2078.
- Maier, G.; Pfriem, F.; Schäfer, U.; Malsch, K. D.; Matusch, R. *Chem. Ber.* **1981**, *114*, 3965. Maier, G.; Pfriem, F.; Malsch, K. D.; Kalinovsky, H. O.; Dehnicke, K. *Chem. Ber.* **1981**, *114*, 3988.
- Van Tamelen, E. E.; Pappas, S. P. *J. Am. Chem. Soc.* **1964**, *84*, 3789. Van Tamelen, E. E. *Angew. Chem.* **1965**, *77*, 759.
- Wilzbach, K. E.; Kaplan, L. *J. Am. Chem. Soc.* **1965**, *87*, 4004.
- Katz, T. J.; Wang, E. J.; Acton, N. *J. Am. Chem. Soc.* **1971**, *93*, 3782.
- Doering, W. v. E.; Roth, W. R. *Angew. Chem.* **1963**, *75*, 27; *Tetrahedron* **1962**, *18*, 67.
- Schröder, G. *Chem. Ber.* **1964**, *97*, 3140.
- Zimmerman, H. E.; Binkley, R. W.; Givens, R. S.; Grunwald, G. L.; Sherwin, M. A. *J. Am. Chem. Soc.* **1969**, *91*, 3316. Maier, G.; Mende, U. *Angew. Chem., Int. Ed. Engl.* **1968**, *7*, 537.
- Eaton, P. E.; Cole, T. W. *J. Am. Chem. Soc.* **1964**, *86*, 3157. Eaton, P. E.; Yat Sun Or; Branca, S. J. *J. Am. Chem. Soc.* **1981**, *103*, 2134.
- Ternansky, R. J.; Balogh, D. W.; Paquette, L. A. *J. Am. Chem. Soc.* **1982**, *64*, 4503. Paquette, L. A.; Ternansky, R. J.; Balogh, D. W.; Kentgen, G. *J. Am. Chem. Soc.* **1983**, *105*, 5446.
- Schleyer, P. v. R.; Prinzbach, H.; Roth, W., unpublished results.
- Woodward, R. B.; Hoffmann, R. "The Conservation of Orbital Symmetry"; Verlag Chemie: Weinheim, FRG, 1970.
- Farnum, D. G.; Ghandi, M.; Raghu, S.; Reitz, T. *J. Org. Chem.* **1982**, *47*, 2598.
- Jones, M.; Scott, L. T. *J. Am. Chem. Soc.* **1967**, *89*, 150.
- Clar, E. "Polycyclic Hydrocarbons"; Academic Press: London, 1964; Springer: Berlin, 1964; "The Aromatic Sextet"; Wiley: London, 1972.
- Diederich, F.; Staab, H. A. *Angew. Chem.* **1978**, *90*, 383. Krieger, C.; Diederich, F.; Schneitzler, D.; Staab, H. A. *Angew. Chem.* **1979**, *91*, 733.
- Balaban, A. T.; Harary, F. *Tetrahedron* **1968**, *24*, 2505.
- Balaban, A. T. *Tetrahedron* **1969**, *25*, 2949.
- Harary, F.; Read, R. C. *Proc. Edinburgh Math. Soc.* **1970**, *17*, 1.
- Lunnon, W. F. In "Graph Theory and Computing"; Read, R. C., Ed.; Academic Press: New York, 1972; p 87. Klarner, D. *Fibonacci Q.* **1965**, *3*, 9; *Can. J. Math.* **1967**, *19*, 851. Dzonova-Jerman-Blazić, B.; Trinajstić, N. *Comput. Chem.* **1982**, *6*, 121; *Croat. Chem. Acta* **1982**, *55*, 347.
- Balaban, A. T. *Rev. Roum. Chim.* **1970**, *15*, 1251.
- Gordon, M.; Davison, W. H. T. *J. Chem. Phys.* **1952**, *20*, 428.
- Balaban, A. T. *Tetrahedron* **1971**, *27*, 6115.
- Oth, J. F. M.; Gilles, J. M. *Tetrahedron Lett.* **1968**, 6259. Oth, J. F. M.; Gilles, J. M.; Antoine, A. *Tetrahedron Lett.* **1968**, 6365.
- Knop, J. V.; Szymanski, K.; Jericević, Z.; Trinajstić, N. *Math. Chem.* **1984**, *16*, 119.
- Dias, J. R. *Math. Chem.* **1983**, *14*, 83; *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 15, 139.
- Bonchev, D.; Balaban, A. T. *J. Chem. Inf. Comput. Sci.* **1981**, *21*, 2.
- Cioslowski, J.; Turek, A. M. *Tetrahedron* **1984**, *40*, 2161.
- Coulson, C. A.; O'Leary, B.; Mallion, R. B. "Hückel Theory for Organic Chemists"; Academic Press, London, 1978.
- Graovac, A.; Gutman, I.; Trinajstić, N. "Topological Approach to the Chemistry of Conjugated Molecules". *Lect. Notes Chem.* **1977**, *4*.
- Sahini, V. E. *J. Chim. Phys.* **1962**, *59*, 177; *Rev. Chim., Acad. Repub. Pop. Roum.* **1962**, *7*, 1265. Sahini, V. E.; Savin, A. *Rev. Roum. Chim.* **1977**, *22*, 39; **1979**, *24*, 165.
- Balaban, A. T. *Rev. Roum. Chim.* **1970**, *15*, 1243.
- Balaban, A. T.; Biermann, D.; Schmidt, W., submitted for publication.
- Biermann, D.; Schmidt, W. *J. Am. Chem. Soc.* **1980**, *102*, 3163, 3173; *Isr. J. Chem.* **1980**, *20*, 312.
- Balasubramanian, K.; Kaufman, J. J.; Koski, W. S.; Balaban, A. T. *J. Comput. Chem.* **1980**, *20*, 312.
- Balaban, A. T. *Rev. Roum. Chim.* **1977**, *22*, 45.
- Gutman, I. *Theor. Chim. Acta* **1977**, *45*, 309; *Z. Naturforsch., A: Phys., Phys. Chem., Kosmophys.* **1982**, *37A*, 69. El-Basil, S.; Gutman, I. *Chem. Phys. Lett.* **1982**, *89*, 145.
- Balaban, A. T.; Tomescu, I. *Math. Chem.* **1983**, *14*, 155.
- Balaban, A. T.; Harary, F. *J. Chem. Doc.* **1971**, *11*, 258.
- Gutman, I.; Trinajstić, N. *Top. Curr. Chem.* **1973**, *42*, 49. Chapter 7, Vol 1, in reference 4. Herndon, W. C.; Ellzey, M. L., Jr. *Tetrahedron* **1975**, *31*, 99. Randić, M.; Trinajstić, N.; Živković, T. *J. Chem. Soc., Faraday Trans. 2*, **1976**, *244*. Živković, T.; Trinajstić, N.; Randić, M. *Mol. Phys.* **1975**, *30*, 517.
- Heilbronner, E.; Jones, T. B. *J. Am. Chem. Soc.* **1978**, *100*, 6506. Heilbronner, E. *Math. Chem.* **1979**, *5*, 105.
- Dewar, M. J. S.; deLiano, C. *J. Am. Chem. Soc.* **1969**, *91*, 789. Dewar, M. J. S.; Harget, A. T.; Trinajstić, N. *J. Am. Chem. Soc.* **1969**, *91*, 6321. Dewar, M. J. S.; Trinajstić, N. *J. Am. Chem. Soc.* **1970**, *92*, 1453; *J. Chem. Soc. A* **1969**, 1754. Dewar, M. J. S. "The Molecular Orbital Theory of Organic Chemistry"; McGraw-Hill: New York, 1969.
- Milun, M.; Sobotka, Z.; Trinajstić, N. *J. Org. Chem.* **1972**, *37*, 139. Aihara, J. *J. Am. Chem. Soc.* **1976**, *98*, 2750.
- Hess, B. A., Jr.; Schaad, L. J. *Pure Appl. Chem.* **1980**, *52*, 1471. Hess, B. A., Jr.; Schaad, L. J.; Agranat, I. *J. Am. Chem. Soc.* **1978**, *100*, 5268.
- Herndon, W. C.; Ellzey, M. L., Jr. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 260.

- (78) Randić, M. *Chem. Phys. Lett.* **1976**, *38*, 68; *J. Am. Chem. Soc.* **1977**, *99*, 444; *Tetrahedron* **1977**, *33*, 1905; *Mol. Phys.* **1977**, *34*, 849; *Pure Appl. Chem.* **1980**, *52*, 1587.
- (79) Gomes, J. A. N. F.; Mallion, R. B. *Rev. Port. Quim.* **1979**, *21*, 82. Gomes, J. A. N. F. *Croat. Chem. Acta* **1980**, *53*, 561; *Theor. Chim. Acta* **1981**, *59*, 333.
- (80) Balaban, A. T.; Tomescu, I. *Croat. Chem. Acta* **1984**, *57*, 391.
- (81) Balaban, A. T. *Acad. Repub. Pop. Rom., Fil. Cluj, Stud. Cercet. Chim.* **1959**, *7*, 257. Balaban, A. T.; Simon, Z. *Rev. Roum. Chim.* **1965**, *10*, 1059.
- (82) Balaban, A. T.; Harary, F. *Rev. Roum. Chim.* **1967**, *12*, 1511.
- (83) Lloyd, E. K. "Proceedings of the 1973 British Combinatorial Conference at Aberystwyth". *Math. Soc. Lect. Notes Ser.* **1973**.
- (84) Ha, T. K.; Cimiriaglia, R.; Nguyen, M. T. *Chem. Phys. Lett.* **1981**, *83*, 317. Vogler, A.; Wright, R. E.; Kunkley, X.; *Angew. Chem.* **1980**, *92*, 745. Dewar, M. J. S. *Pure Appl. Chem.* **1975**, *44*, 767. Huber, H. *Angew. Chem., Int. Ed. Engl.* **1982**, *21*, 64.
- (85) Balaban, A. T.; Simon, Z. *Rev. Roum. Chim.* **1964**, *9*, 99.
- (86) Ollis, W. D.; Ramsden, C. A. *Adv. Heterocycl. Chem.* **1976**, *19*, 3. Ramsden, C. A. In "Comprehensive Organic Chemistry"; Barton, D.; Ollis, W. D.; Eds.; Pergamon Press: Oxford, 1979; Vol. 4, p 1171. Baker, W.; Ollis, W. D. *Q. Rev. Chem. Soc.* **1957**, *11*, 15.
- (87) Van der Kerk, S. M.; Budzadaar, P. H. M.; Van der Kerk-van Hoof, A.; van der Kerk, G. J. M.; Schleyer, P. v. R. *Angew. Chem., Int. Ed. Engl.* **1983**, *22*, 48. Hildenbrand, M.; Pritzka, H.; Zenneck, U.; Siebert, W. *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 371. Wehrmann, R.; Pires, C.; Klusik, H.; Berndt, A. *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 372. Schleyer, P. v. R.; Budzadaar, P. H. M.; Cremer, D.; Kraka, E. *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 374. Krogh-Jespersen, K.; Cremer, D.; Dill, J. D.; Pople, J. A.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1981**, *103*, 2589.
- (88) Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17, 2636; *J. Chem. Phys.* **1947**, *15*, 766; *J. Chem. Phys.* **1948**, *52*, 425, 1082.
- (89) Balaban, A. T.; Chiriac, A.; Motoc, I.; Simon, Z. "Steric Fit in QSAR". *Lect. Notes Chem.* **1980**, *15*, 22.
- (90) Sablic, A.; Trinajstić, N. *Acta Pharm. Jugosl.* **1981**, *31*, 189.
- (91) Balaban, A. T. *Theor. Chim. Acta* **1979**, *53*, 355.
- (92) Balaban, A. T. *Pure Appl. Chem.* **1983**, *55*, 199.
- (93) Randić, M. *J. Am. Chem. Soc.* **1965**, *97*, 6609; *Int. J. Quantum Chem., Symp.* **1978**, *5*, 245.
- (94) Kier, L. B.; Hall, L. H. "Molecular Connectivity in Chemistry and Drug Research"; Academic Press: New York, 1976.
- (95) Hosoya, H. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332; *Theor. Chim. Acta* **1972**, *25*, 215; *J. Chem. Doc.* **1972**, *12*, 181.
- (96) Balaban, A. T. *Chem. Phys. Lett.* **1982**, *89*, 399; *Math. Chem.* **1984**, *16*, 163. Balaban, A. T.; Quintas, L. V. *Math. Chem.* **1983**, *14*, 213.
- (97) Bonchev, D.; Trinajstić, N. *J. Chem. Phys.* **1977**, *67*, 4517.
- (98) Bonchev, D. "Theoretic Information Indices for Characterization of Chemical Structures"; Wiley-Research Studies: Chichester, England, 1983; *Int. J. Quantum Chem., Symp.* **1978**, *12*, 293.
- (99) Motoc, I.; Balaban, A. T. *Rev. Roum. Chim.* **1981**, *26*, 539.
- (100) Motoc, I.; Balaban, A. T.; Mekenyan, O.; Bonchev, D. *Math. Chem.* **1982**, *13*, 369. Bonchev, D.; Mekenyan, O.; Trinajstić, N. *J. Comput. Chem.* **1981**, *2*, 127.
- (101) Balaban, A. T.; Motoc, I. *Math. Chem.* **1979**, *5*, 197.
- (102) Barysz, M.; Jashari, G.; Lall, R. S.; Srivatsva, V. K.; Trinajstić, N. In reference 6, p 222.
- (103) Spialter, L. *J. Chem. Doc.* **1964**, *4*, 261; *J. Am. Chem. Soc.* **1963**, *85*, 2012.
- (104) Kudo, Y.; Yamasaki, T.; Sasaki, S. *J. Chem. Doc.* **1973**, *13*, 225.
- (105) Herndon, W. C. *J. Chem. Doc.* **1974**, *14*, 150.
- (106) Randić, M. *J. Chem. Phys.* **1974**, *60*, 3920; *J. Chem. Inf. Comput. Sci.* **1977**, *17*, 171.
- (107) Mackay, A. L. *J. Chem. Phys.* **1975**, *62*, 308.
- (108) Randić, M. *J. Chem. Phys.* **1975**, *62*, 309.
- (109) Randić, M. *Math. Chem.* **1979**, *7*, 5.
- (110) Randić, M. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 164.
- (111) Balaban, A. T., in press.
- (112) Balaban, A. T.; Farcasiu, D.; Banica, R.; *Rev. Roum. Chim.* **1966**, *11*, 1205.
- (113) Dunitz, J. D.; Prelog, V. *Angew. Chem., Int. Ed. Engl.* **1968**, *7*, 726.
- (114) Leuterbur, P. C.; Ramirez, F. *J. Am. Chem. Soc.* **1968**, *90*, 6722.
- (115) Zon, C.; Mislou, K. *Top. Curr. Chem.* **1971**, *19*, 61.
- (116) Gund, T. M.; Schleyer, P. v. R.; Gund, P. H.; Wipke, W. T. *J. Am. Chem. Soc.* **1975**, *97*, 743. Osawa, E.; Aigami, K.; Takaishi, N.; Inamoto, Y.; Yoshiaki, F.; Majerski, Z.; Schleyer, P. v. R.; Engler, E. M.; Farcasiu, M. *J. Am. Chem. Soc.* **1977**, *99*, 5361.
- (117) Whitlock, H. W., Jr.; Siefken, M. W. *J. Am. Chem. Soc.* **1968**, *90*, 4929.
- (118) Balaban, A. T.; Schleyer, P. v. R. *Tetrahedron* **1978**, *34*, 3599.
- (119) Brocas, J.; Gielen, M.; Willem, R. "The Permutational Approach to Dynamic Stereochemistry"; McGraw-Hill: New York, 1983.
- (120) Ruch, E.; Ugi, I. *Top. Stereochem.* **1969**, *4*, 99; *Theor. Chim. Acta* **1966**, *4*, 287. Ruch, E.; Hässelbarth, W.; Richter, B. *Theor. Chim. Acta* **1973**, *29*, 259.
- (121) Klemperer, W. G. *J. Chem. Phys.* **1972**, *56*, 5478; *Inorg. Chem.* **1972**, *11*, 2668; *J. Am. Chem. Soc.* **1972**, *94*, 6940, 8360; **1972**, *95*, 380, 2105. Nourse, J. G. *J. Am. Chem. Soc.* **1977**, *99*, 2063; *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 2385; *J. Chem. Inf. Comput. Sci.* **1981**, *21*, 168.
- (122) Ugi, I.; Dugundji, J.; Kopp, R.; Marquarding, D. "Perspectives in Theoretical Stereochemistry". *Lect. Notes Chem.* **1984**, *36*.
- (123) Randić, M., submitted for publication.
- (124) Dainton, F. Cited from Guthrie, R. D. *Chem. Int.* **1984**, *No. 5*, 35.
- (125) Corey, E. J.; Wipke, W. T. *Science (Washington, D.C.)* **1969**, *166*, 178. Corey, E. J. *Q. Rev. Chem. Soc.* **1971**, *25*, 455.
- (126) Corey, E. J.; Johnson, A. P.; Long, A. K. *J. Org. Chem.* **1980**, *45*, 2051, and further references therein on the program LHASA.
- (127) Hendrickson, J. B. *J. Am. Chem. Soc.* **1977**, *99*, 5439. And previous parts in the series; Hendrickson, J. B.; Braun-Keller, E. *J. Comput. Chem.* **1980**, *1*, 323, for details on the program SYNGEN.
- (128) Bersohn, M.; Esack, A. *Chem. Rev.* **1976**, *76*, 269.
- (129) Balaban, A. T. *Math. Chem.* **1980**, *8*, 159, and references cited therein.
- (130) Gelernter, H.; et al. *Science (Washington, D.C.)* **1977**, *197*, 1041. Gasteiger, J.; Jochum, C. *Top. Curr. Chem.* **1978**, *74*, 93. Johnson, A. P. *Chem. Br.* **1985**, *21*, 59.
- (131) Balaban, A. T. *Rev. Chim., Acad. Repub. Pop. Roum.* **1964**, *7*, 675.
- (132) Fletcher, J. H. *J. Chem. Doc.* **1967**, *7*, 64. See also, Fletcher, J. H.; Dermer, O. C.; Fox, R. W. "Nomenclature of Organic Chemistry, Principles and Practice". *Adv. Chem. Ser.* **1974**, *No. 126*.
- (133) Verkade, P. E. *Bull. Soc. Chim. Fr.* **1978**, *13*, and previous parts in the series.
- (134) Goodson, A. L. *J. Chem. Inf. Comput. Sci.* **1980**, *20*, 167-72.
- (135) "Chemical Abstracts 1985 Index Guide"; American Chemical Society: Washington, DC, 1985; Appendix IV, Section J.
- (136) IUPAC "Nomenclature of Organic Chemistry, Sections A-F and H"; Pergamon Press: Oxford, 1979; "Definitive Rules for Nomenclature of Inorganic Chemistry 1957"; Butterworths: London, 1959.
- (137) Chemical Abstracts Service "Naming and Indexing of Chemical Substances for Chemical Abstracts". "Index Guide", 1982; American Chemical Society: Washington, DC, 1982; Appendix IV.
- (138) Lozac'h, N.; Goodson, A. L.; Powell, W. H. *Angew. Chem., Int. Ed. Engl.* **1979**, *18*, 887.
- (139) Goodson, A. L. *J. Chem. Inf. Comput. Sci.* **1980**, *20*, 172; *Croat. Chem. Acta* **1983**, *56*, 315.
- (140) Read, R. C.; Milner, R. S. "A New System for the Designation of Chemical Compounds for Purpose of Data Retrieval. I. Acyclic Compounds". Report to the University of West Indies; University of West Indies: 1968.
- (141) Gordon, M.; Kendall, C. E.; Davison, W. H. T. "Chemical Ciphering". *R. Inst. Chem., Monogr., Rep.* **1948**.
- (142) Dyson, M. "A New Notation and Enumeration System for Organic Compounds", 2nd ed.; Longmans: London, 1949; "Rules for IUPAC Notation for Organic Compounds"; Longmans: London, 1961.
- (143) Wiswesser, W. J. "A Line-Formula Chemical Notation"; Crowell: New York, 1954.
- (144) Hyde, E.; Mathews, F. W.; Thompson, L. N.; Wiswesser, W. J. *J. Chem. Doc.* **1967**, *7*, 200. Ebe, T.; Zamora, A. *J. Chem. Inf. Comput. Sci.* **1976**, *16*, 36. Farrell, C. D.; Chauvenet, A. R.; Koniver, D. A. *J. Chem. Inf. Comput. Sci.* **1971**, *11*, 52.
- (145) Morgan, H. L. *J. Chem. Doc.* **1965**, *5*, 107.
- (146) O'Korn, L. J. In "Algorithms for Chemical Computations". *ACS Symp. Ser.* **1977**, *No. 46*, 122.
- (147) Dittmar, P. G.; Mockus, J.; Couvreur, K. M. *J. Chem. Inf. Comput. Sci.* **1977**, *17*, 186. Blake, J. E.; Farmer, N. A.; Haines, R. C. *J. Chem. Inf. Comput. Sci.* **1977**, *17*, 223.
- (148) Wipke, W. T.; Dyott, T. M. *J. Am. Chem. Soc.* **1974**, *96*, 4825, 4834.
- (149) Dubois, J. E. In "Computer Representation and Manipulation of Chemical Information"; Wipke, W. T.; Heller, S. R.; Feldman, R. J.; Hyde, E., Eds.; Wiley: New York, 1974. Reference 1, p 333.
- (150) Institute for Scientific Information, Philadelphia (databases Questel/DARC and Index Chemicus Online).
- (151) Fugmann, R. In "Chemical Information Systems". Ash, J. E.; Hyde, E., Eds.; Horwood: Chichester, U.K., 1975; p 195.
- (152) Schubert, W.; Ugi, I. *J. Am. Chem. Soc.* **1978**, *100*, 37; *Chimia* **1979**, *33*, 183. Dugundji, J.; Ugi, I. *Top. Curr. Chem.* **1973**, *39*, 19.
- (153) Kochetova, A. A.; Skorobogatov, V. A.; Khvorostov, P. V. In "Mashinnye Metody Obnaruzheniya Zakonomernosti, Analiza Struktur i Proektiraniya" Skorobogatov, V. A., Ed.; Novosibirsk: 1982; p 70, and references cited therein.
- (154) Moreau, G. *Nouv. J. Chim.* **1980**, *4*, 17.
- (155) Balaban, A. T.; Mekenyan, O.; Bonchev, D. submitted for publication.
- (156) Herndon, W. C.; Leonard, J. E. *Inorg. Chem.* **1983**, *22*, 554.