

SYMBEQ Program and Its Application in Computer-Assisted Reaction Design

Nikolai S. Zefirov,* Igor I. Baskin, and Vladimir A. Palyulin

Department of Chemistry, Moscow State University, Moscow, 119899, Russia

Received July 17, 1992*

The SYMBEQ computer program created for the search of novel types of organic reactions and based on the formal-logical approach is considered. The application of symbolic equations generated by this program for the design of organic reactions is discussed.

INTRODUCTION

During the last decade, numerous efforts of different groups have resulted in the creation of a new field of computer application in organic chemistry: *computer-assisted reaction design*.¹⁻¹² This area of computer chemistry is based on the use of formal algebraic models (mainly the graph theory) and combinatorial algorithms. These abstract mathematical methods provide a universal tool to deduce the unprecedented types of chemical conversions which can be later verified theoretically and synthetically.

In our work we have used the formal-logical approach to the organic reactions, which was created in the 1970s and 1980s by Zefirov, Tratch, Baskin, and Gamziani and since that time has been intensively exploited in various investigations.^{8,9,13} Fundamentals of this approach were thoroughly discussed in our subsequent publications.¹⁴⁻²¹ Thus, in this paper we shall briefly consider only the basic notions of the formal-logical approach, used by us in designing chemical reactions, and the focus of this paper will be on the computer realization of these principles in the SYMBEQ program. We shall also present a short discussion of some of our results obtained using the SYMBEQ program.

To ensure the proper historical perspective, we have to mention some related approaches. First, the computer-assisted reaction design was elaborated by Ugi and co-workers,¹⁻⁷ who have created the IGOR program. The work of this program is based on the use of the UGI-DUGUNDJI model of constitutional chemistry.²² The structures of chemical compounds as well as their interconversions are represented in this approach by matrices. IGOR can generate chemical structures of educts and products, provided the reaction matrices representing reaction categories are entered manually by a chemist. Some new reactions have already been found using IGOR.²

An alternative approach based on the generation of the complete set of nonisomorphic spanning subgraphs of a given graph or pseudograph was proposed recently by R. Herges.⁷ One should mention also the works of Fujita concerning classification and enumeration of organic reactions.²³

BASIC NOTIONS OF FORMAL-LOGICAL APPROACH TO ORGANIC REACTIONS

In the framework of the formal-logical approach, all types of chemical transformations are described by the corresponding *bond redistributions* (BRs). The BR is called *canonical* if charged atoms and radical centers do not take part in it. With

the help of some formal procedures, each noncanonical BR can be reduced to the canonical one.

The next basic notion of the formal-logical approach is a *symbolic equation* (SEQ), a formal scheme that represents all bonds changing their orders in the BR and all atoms which are connected by those bonds (such atoms are called *reaction centers* (RCs)). Each atom in a SEQ is labeled by an abstract symbol indicating the valence change of the atom: RCs with unchanged valences are called *ordinary* ones and are denoted in SEQs by heavy dots; centers that change their valence by 2, 4, 6, ... units are called *specific* ones and are denoted in SEQs by X, X', X'', etc. (it was shown²⁰ that only even values of valence changes are possible in canonical BR). These equations containing symbols of chemical elements instead of abstract symbols are called *reaction equations*.

The other important notion of the formal-logical approach is the *topology* of BR, which is defined as a contour formed by all bonds changing their orders during the reaction. Each contour can be identified by a number of graphs (*topology identifiers*).

All these notions are illustrated in Figure 1, where the reaction of the opening of epoxides with tertiary phosphines²⁴ is described by chemical equation 1a, with all atoms and bonds, reaction equation 1b, containing only RCs and bonds connecting them, and SEQ 1c, in which the atom of phosphorus increasing its valence by two units is denoted by the symbol X, while the atoms of carbon and oxygen preserving their valences unchanged in the course of the reaction are denoted by heavy dots. Linear-cyclic topology of BR 1e in this reaction is identified by graph 1d, in which one edge (drawn vertically) corresponds to the formation of double bond P=O, another edge (drawn horizontally) corresponds to the formation of the double bond C=C from the single bond, and the two remaining edges correspond to the breaking of bonds.

Figure 1 not only illustrates the classification scheme of organic reactions accepted in the formal-logical approach but also shows how this scheme can be used for the generation of all possible reactions.

GENERATION OF ORGANIC REACTIONS BY THE SYMBEQ PROGRAM

The computer program SYMBEQ^{10,11} is intended for the systematic search for all formally possible symbolic equations and reaction equations describing the reactivity of organic compounds. At the first stage of its work, SYMBEQ generates the complete lists of canonical SEQs for the given topology identifier (the mathematical statement of the generation problem, the algorithm, and its implementation are thoroughly considered in ref 12 and reviewed in the Appendix). In the course of generating SEQs, some restrictions are imposed on

* Abstract published in *Advance ACS Abstracts*, April 1, 1994.

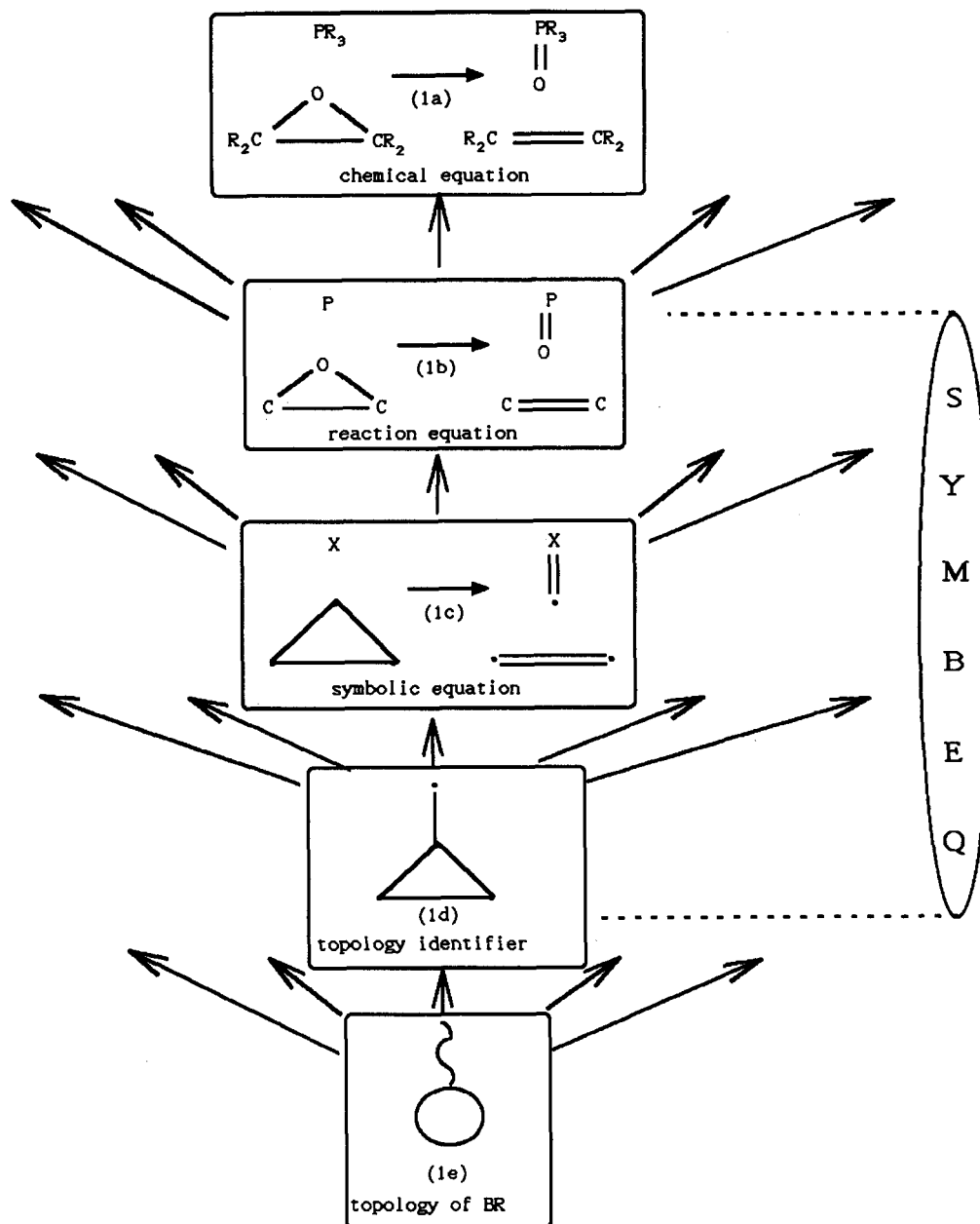


Figure 1. Hierarchical classification tree of organic reactions in the formal-logical approach: chemical equation 1a, reaction equation 1b, SEQ 1c, topology identifier 1d, and the topology of BR 1e.

the allowed numbers of specific RCs and on the maximal value of their *reaction numbers* (e.g., the sums of orders of all their bonds, that each of them must take part in the BR). The possibility for setting allowed types of changes of bond orders in the BRs is also provided. Thus for the graph 1d in Figure 1, a series of SEQs can be generated. One of them will be 1c if the presence of at least one specific RC is allowed and the maximal reaction number is set to be no less than 2. The additional possibility of eliminating SEQs with structures containing allene or triple bonds in cycles is also incorporated into the program. An example of complete lists of SEQs generated by the SYMBEQ program for graph 1d is presented in Figure 2. It should be mentioned that the same list has been obtained by R. Herges,⁷ who has applied his new algorithm for the generation of the complete set of nonisomorphic subgraphs of a given graph or pseudograph. In fact, SEQ 2f, in which all bonds change their orders from zero (i.e., formation of bonds) or to zero (i.e., breaking of bonds), can be associated with the notion of reaction category in the UGI-DUGUNDJI²² approach and corresponds to the R-matrix.

Thus, SYMBEQ can also be used for generating R-matrices representing reaction categories for the IGOR program or for Herges's approach.

For systematic investigations, graphs identifying topologies of BR can be extracted from Harary's book.²⁵ Using the SYMBEQ program, we have systematically studied SEQs corresponding to cyclic,¹⁴⁻¹⁶ linear,⁸ linear-cyclic,⁹ and some more complex types²⁶ of topologies of BRs. The user can choose already prepared graphs or draw any other graph using the graphical editor GIRL²⁷ incorporated in the SYMBEQ program. For convenience, the complete lists of SEQs can be printed on paper in compact form like in Figure 2.

The next part of the SYMBEQ program provides the sequential construction of all reaction equations related to each of the previously generated SEQs; this program includes a special labeling procedure which assigns the symbols C, H, N, O, etc. to all vertices of the SEQ in accordance with the allowed values of valences of the corresponding elements and with the allowed orders of adjacent bonds. The user is also provided with the option to edit a default set of permitted

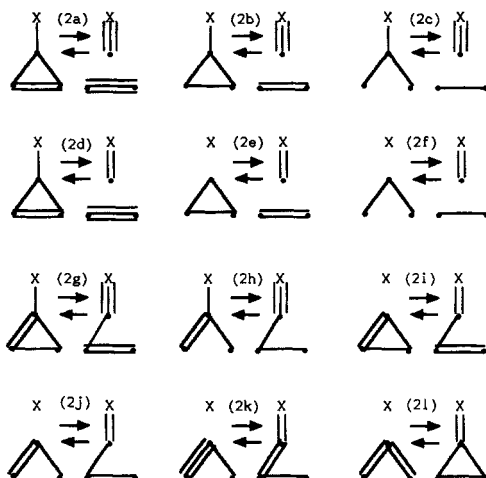


Figure 2. Complete list of SEQs for the topology identifier 1d. These SEQs have been obtained using SYMBEQ for the following restrictions: (1) the number of specific RCs does not exceed 1; (2) the maximal reaction number is set to be 4; and (3) the appearance of triple and allene bonds in cycles is forbidden.

labels for each equivalence class of atoms in SEQs. Thus for SEQ 1c in Figure 1, a series of reaction equations can be generated, and one of them will be 1b. The general flow chart of the work of the SYMBEQ program is represented in Figure 3.

The SYMBEQ program is written in Microsoft QuickBasic language and is intended for IBM PC AT compatibles with EGA or VGA card, mouse, and EPSON printer. Together with the instruction, it is accessible as discussed in ref 35.

APPLICATION OF THE COMPLETE LISTS OF SYMBOLIC EQUATIONS FOR THE DESIGN OF ORGANIC REACTIONS

The SEQs generated by the SYMBEQ program correspond not only to the already known organic reactions but also to the unknown ones. Therefore, they describe the way the bonds are redistributed in these unknown processes, providing the reaction systems for the search for novel types of organic reactions. Hence, complete lists of SEQs can be used as an instrument in searching for novel chemical processes.

The search for novel organic reactions includes (1) the analysis of different ways of replacing abstract symbols in SEQs by the symbols of chemical elements in accordance with the permitted values of valence, (2) the selection of chemical reagents, and (3) experimental investigations of predicted reactions. As an example, let us consider SEQ 3a in Figure 4, by which a practically unknown method for synthesis of 5-membered heterocycles can be described; in this case oxygen, nitrogen, or sulfur atoms must correspond to the ordinary RC, connected in the left part of this SEQ with the specific RC, while any atom which can change its valence by two units (N, S, Se, I, P, etc.) may be related to the specific RC in this SEQ. Some interesting possibilities may be considered. Equations 3b–d show possible methods of furan formation. As the acetylene component, the esters of acetylenecarboxylic acid and acetylenedicarboxylic acid could be used. Equation 3e describes the method which could be proposed for the pyrrole synthesis. For the formation of thiophenes, reaction 3f could be suggested.

The alternative way of reaction design consists of the replacement of reaction fragments in SEQ by their presentations. In many organic reactions, the presentation (4b, Figure 5) corresponds to the reaction fragment (4a). By placing 4b in SEQ 5a, one can obtain a scheme of an interesting chemical

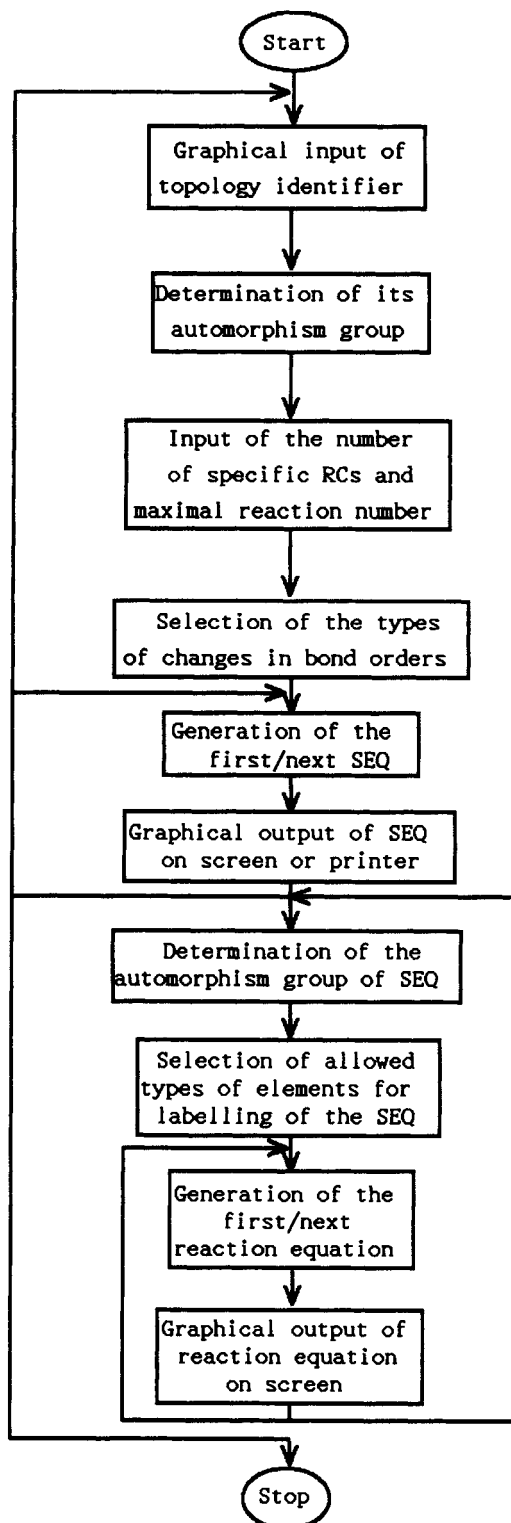


Figure 3. General flow chart of the work of the SYMBEQ program. The flow direction from two branching points is controlled interactively by the user.

process (5b), in which ylides and metallocarbenes can be proposed as the reagent $\text{CH}_2=\text{Y}$.

One of the main principles of the application of SEQs for the design of novel organic reactions is the selection of such reaction systems which can provide the "driving force" for the chemical process, e.g., the formation of strong chemical bonds, the extrusion of energetically stable molecules, and the formation of aromatic systems. Equation 6b in Figure 6 (which can be obtained from SEQ 6a) shows one of several possible directions of the decay of the vinyl ester of α -diazacarboxylic

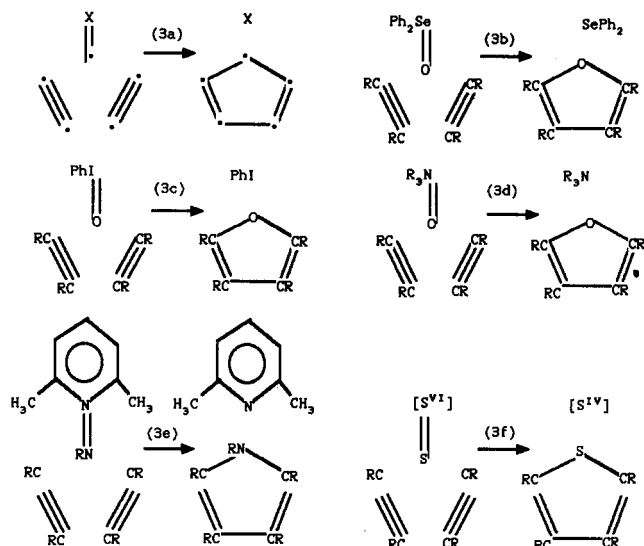


Figure 4. Design of reactions 3b-f from SEQ 3a.

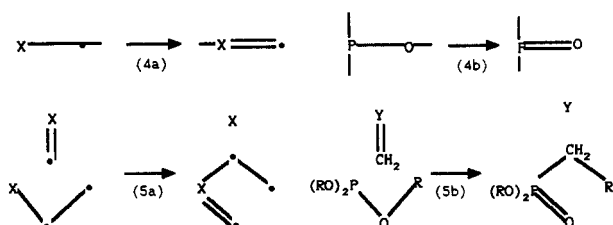


Figure 5. Design of reaction 5b from SEQ 5a by replacement of reaction fragment 4a with its presentation, 4b.

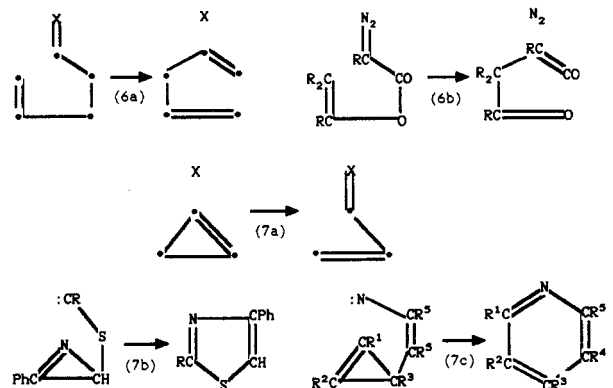
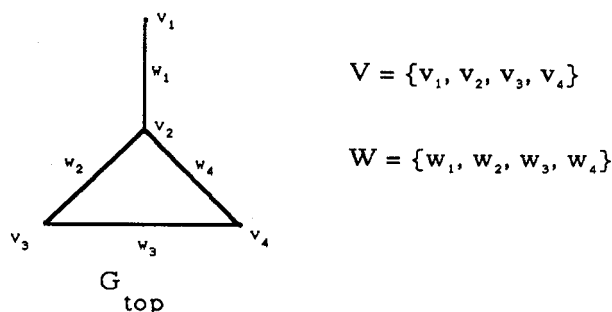


Figure 6. Design of (a) reaction 6b from SEQ 6a and (b) reactions 7b and 7c from SEQ 7a.

acid; the high stability of the carbonyl group and molecular nitrogen is likely to promote this reaction. The formation of a stable aromatic system must promote intramolecular cyclization of carbenes (eq 7b) or nitrenes (eq 7c) (or carbenoids and nitrenoids) in accordance with SEQ 7a.

In the design of novel organic reactions, one can proceed from the proposition that for any SEQ it is possible to find such substituents and heteroatoms in the structures of reagents, such catalysts, and such experimental conditions that the suggested reactions would actually take place. Therefore, the main problem is *how to make this search*. In the experimental studies, the main obstacles preventing the desired reactions from proceeding in the desired direction should be found and eliminated by changing structures of the reagents and experimental conditions. In theoretical investigations, it is necessary to represent the desired reactions as sequences of elemental stages (or as one elemental stage), for each of which exact theoretical computations can be made.

Figure 7. Topology identifier G_{top} and the sets of its vertices and edges.

EXPERIMENTAL VERIFICATION

The reactions 3b-d predicted using the first version of SYMBEQ¹⁰ were experimentally investigated in our laboratory.²⁸ The ester of acetylenedicarboxylic acid was taken as the acetylene component. It was shown that dimethyl selenoxide gives the ester of furan tetracarboxylic acid, in accordance with the eq 3b, with yield of 40%; iodosobenzene gives the same product with substantially less yield [the major product is iodobenzene bis(methyl oxalate)^{29,30}], but the oxide of trimethylamine reacts in a different manner³¹ to give substituted dimethyl oxalacetate via a 1,3-dipolar cycloaddition reaction followed by cleavage of the isoxazoline ring of the adduct. Therefore, even the study of only one of the symbolic equations generated using the SYMBEQ program has allowed us to find the novel reaction 3b.

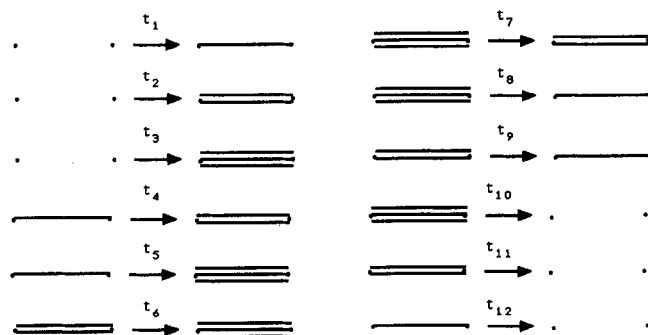
CONCLUSION

On the basis of the formal-logical approach, we have created the SYMBEQ program for searching for novel reactions of organic compounds. This program allows one to generate complete lists of SEQs, which can serve as an important instrument in searching for novel types of organic reactions. We have also proposed the methodology for the application of SEQs for the design of novel organic reactions. This methodology partially incorporated in the SYMBEQ program has allowed us to propose a number of interesting chemical processes, and, what is more, some of them have been already found in the laboratory.

Nevertheless, serious problems associated with the application of SYMBEQ for reaction design exist. This program still does not include the criteria associated with physical and quantum chemistry; therefore, the selection of novel types of reactions is based on the experience and intuition of a chemist. The selection of reaction systems for the experimental search for novel reactions should involve the estimation of kinetic and thermodynamic parameters for them. The application of quantum chemical calculations for this purpose was recently demonstrated by R. Herges and C. Hoock³⁴ for the discovery of a novel pericyclic reaction (generated by the IGOR program). The solution of related problems for various types of organic reactions would surely make a big impact on the future development of this direction in computer-assisted reaction design.

APPENDIX. MATHEMATICAL BACKGROUND OF THE GENERATING PROBLEM

Let G_{top} be a topology identifier graph with n vertices and m edges (see Figure 7). $V = \{v_1, v_2, \dots, v_n\}$, $|V| = n$, is defined as a set of all vertices of the graph G_{top} and $W = \{w_1, w_2, \dots, w_m\}$, $|W| = m$, as a set of all its edges. A vertex automorphism group ${}^vG = \{^vg\}$ acting on the set V consists of all permutations



$$T = \{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9, t_{10}, t_{11}, t_{12}\}$$

Figure 8. Set of all possible changes of individual bond orders.

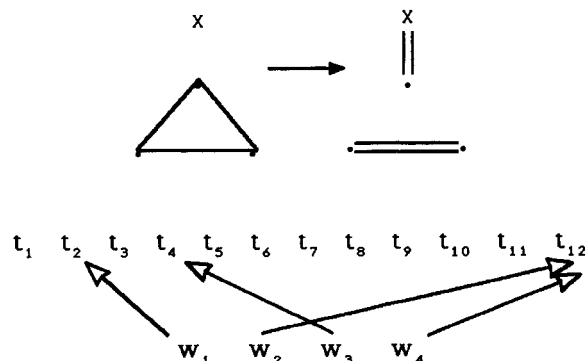


Figure 9. Correspondence between SEQ and the labeling of edges.

preserving adjacency of the graph G_{top} unchanged (this means that if vertices v_i and v_j are connected in the graph G_{top} by an edge, then the vertices ${}^v g(v_i)$ and ${}^v g(v_j)$ must also be connected by it, and vice versa). Permutations ${}^v g$ are searched in the SYMBEQ program using a separate backtrack procedure.³² The group ${}^v G$ is isomorphic to an edge automorphism group ${}^w G$ acting on the set W .

Let a set $T = \{t_1, t_2, \dots, t_{12}\}$ consist of 12 unified labels representing all formally possible changes of individual bond orders (see Figure 8). From the sets W and T , a new set ${}^s \Phi = T^w = \{{}^s \phi: W \rightarrow T | {}^s P({}^s \phi)\}$, consisting of all possible mappings ${}^s \phi$ from W to T which satisfy some predicate system ${}^s P({}^s \phi) = {}^s P_1({}^s \phi) \wedge {}^s P_2({}^s \phi) \wedge \dots \wedge {}^s P_p({}^s \phi)$, can be constructed. While the first predicate ${}^s P_1({}^s \phi)$ checks that all valence changes in canonical SEQs should be even, other predicates correspond to possible restrictions (e.g., restrictions related to maximal reaction numbers of RCs or to allowed numbers of specific RCs in corresponding SEQs) imposed by the user. Each mapping ${}^s \phi$ corresponds to a SEQ (see Figure 9).

It can be seen from Figure 8 that labels t_j and t_{13-j} reflect opposite directions of the same process. This can be described by introducing a group $H = \{h_0, h_1\}$ which acts on the set T and consists of the following permutations: $h_0(t_j) = t_j$, $h_1(t_j) = t_{13-j}$. From the groups ${}^w G$ and H , a new power group ${}^s \Gamma = H {}^w G$ acting on the set ${}^s \Phi = T^w$ can be constructed. Thus, the generating problem for SEQs reduces to a problem of constructive enumeration of a representative system for all orbits of the power group ${}^s \Gamma$ on the set of mappings ${}^s \Phi$. This problem is solved in the SYMBEQ program by using a special combinatorial algorithm³³ for the sequential generation of canonical mappings ${}^s \phi$ having the lexicographically smallest code in their orbits. This algorithm is based on the use of the partial codes of the mappings ${}^s \phi$ corresponding to partial labeling of G_{top} with edges 1, 2, ..., i ($i < m$) labeled. At each step of a backtrack procedure, partial codes are evaluated with respect to each permutation ${}^w g \in {}^w G$, and all partial codes

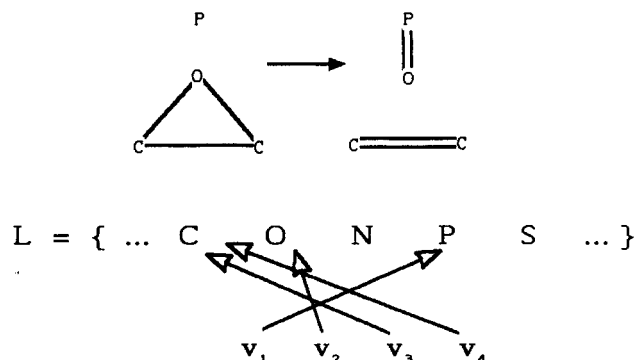


Figure 10. Correspondence between the reaction equation and the labeling of vertices.

to which no complete canonical codes correspond are rejected. Therefore, only canonical mappings are generated in the algorithm.

A SEQ automorphism group ${}^v G' = \{{}^v g'\}$ is a subgroup of ${}^v G$ and acts on the set V . ${}^v G'$ consists of all permutations preserving the adjacency of the graph G_{top} and the labeling ${}^s \phi$ of the edges unchanged (this means that if vertices v_i and v_j are connected in the graph G_{top} by an edge labeled by t_k , then the vertices ${}^v g'(v_i)$ and ${}^v g'(v_j)$ must also be connected by an edge labeled by t_k). This group is evaluated in the SYMBEQ program from the group ${}^v G$ by checking the aforementioned condition for all its permutations.

$L = \{l_1, l_2, \dots, l_s\}$ is defined as a set of all possible atomic symbols (e.g., C, H, N, O, etc.). From the sets V and L , a new set ${}^r \Phi = L^V = \{{}^r \phi: V \rightarrow L | {}^r P({}^r \phi)\}$, consisting of all possible mappings ${}^r \phi$ from V to L which satisfy some predicate system ${}^r P({}^r \phi) = {}^r P_1({}^r \phi) \wedge {}^r P_2({}^r \phi) \wedge \dots \wedge {}^r P_r({}^r \phi)$, can be constructed. The predicates correspond to possible restrictions (e.g., restrictions related to valences of the chemical elements or to allowed numbers of chemical elements of each kind). Each mapping ${}^r \phi$ corresponds to a reaction equation (see Figure 10).

E is a permutation group acting on the set L and consisting of only one identity permutation. From the group E and ${}^v G'$, a new induced power group ${}^r \Gamma = E {}^v G'$ acting on the set ${}^r \Phi = L^V$ can be constructed. Thus, the generating problem for reaction equations reduces to a problem of constructive enumeration of a representative system for all orbits of the power group ${}^r \Gamma$ on the set of mappings ${}^r \Phi$. This problem is solved in the SYMBEQ program by using a special combinatorial algorithm for the sequential generation of canonical mappings ${}^r \phi$ having the lexicographically smallest code in their orbits. The idea of this algorithm is very close to those considered above for the case of generating SEQs.

REFERENCES AND NOTES

- (1) (a) Bauer, J.; Ugi, I. Chemical Reactions and Structures without Precedent, Generated by Computer Programs. *J. Chem. Res.* **1982**, (S) 11, 298; (M) 11, 3101-3196, 3201-3260. (b) Bauer, J.; Herges, R.; Fontain, E.; Ugi, I. IGOR and Computer Assisted Innovation in Chemistry. *Chimia* **1985**, 39, 43-53.
- (2) Bauer, J.; Fontain, E.; Forstmeier, D.; Ugi, I. Interactive Generation of Organic Reactions by IGOR2 and the PC-assisted Discovery of a New Reaction. *Tetrahedron Comput. Methodol.* **1988**, 1, 129-132.
- (3) Bauer, J. IGOR2: a PC-Program for Generating New Reactions and Molecular Structures. *Tetrahedron Comput. Methodol.* **1989**, 2, 269-280.
- (4) Ugi, I.; Bauer, J.; Baumgartner, R.; Fontain, E.; Forstmeier D.; Lohberger, S. Computer Assistance in the Design of Syntheses and a New Generation of Computer Programs for the Solution of Chemical Problems by Molecular Logic. *Pure Appl. Chem.* **1988**, 60, 1573-1586.
- (5) Herges, R. Reaction Planning (Computer Aided Reaction Design). In *Chemical Structures*; Warr, W., Ed.; Springer: Berlin, 1988; p 385.
- (6) Herges, R. Reaction Planning: Computer-Aided Reaction Design. *Tetrahedron Comput. Methodol.* **1988**, 1, 15-26.

- (7) Herges, R. Reaction Planning: Prediction of New Organic Reactions. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 377–383.
- (8) Tratch, S. S.; Baskin, I. I.; Zefirov, N. S. Problems of Molecular Design and Computers. XIII. Systematic Analysis of Organic Processes with Opened Topologies of Bond Redistributions. *Zh. Org. Khim.* **1988**, *24*, 1121–1132.
- (9) Tratch, S. S.; Baskin, I. I.; Zefirov, N. S. Problems of Molecular Design and Computers. XIV. Systematic Analysis of Organic Processes with Linear-Cyclic Topology of Bond Redistributions. *Zh. Org. Khim.* **1989**, *25*, 1585–1606.
- (10) Baskin, I. I.; Tratch, S. S.; Zefirov, N. S. Program of Searching for New Types of Organic Reactions on the Personal Computer "Iskra-226". *Proceedings of the 7th All-Union Conference on Application of Computers in Chemical Studies and Molecular Spectroscopy*; Riga, 1986; pp 88–90.
- (11) Baskin, I. I.; Palyulin, V. A.; Zefirov, N. S. SYMBEQ2—Computer Program for Designing Organic Reactions. *Proceedings of the 9th All-Union Conference on Chemical Informatics*; Chernogolovka, 1992; p 203.
- (12) Zefirov, N. S.; Tratch, S. S. Symbolic Equations and Their Application to Reaction Design. *Anal. Chim. Acta* **1990**, *235*, 115–134.
- (13) Zefirov, N. S.; Tratch, S. S.; Gamziani, G. A. Problems of Molecular Design and Computers. IX. Formal-Logical Approach to Organic Reactions. Main Characteristics of Processes with Linear Electron Transfer. *Zh. Org. Khim.* **1986**, *22*, 1341–1359.
- (14) Zefirov, N. S.; Tratch, S. S. Rearrangements and Cyclizations. X. Formal-Logical Approach to Synchronic Reactions. Formal Technique for Deducing the Complete Set of Possible Multicenter Processes with the Cyclic Electron Transfer. *Zh. Org. Khim.* **1975**, *11*, 225–231.
- (15) Zefirov, N. S.; Tratch, S. S. Rearrangements and Cyclizations. XI. Formal-Logical Approach to Synchronic Reactions. Classification of Multicenter Processes with Cyclic Electron Transfer. *Zh. Org. Khim.* **1975**, *11*, 1785–1800.
- (16) Zefirov, N. S.; Tratch, S. S. Rearrangements and Cyclizations. XII. Formal-Logical Approach to Synchronic Reactions. Complete Sets of Multicenter Processes with Cyclic Electron Transfer on Three, Four, Five and Six Centers. *Zh. Org. Khim.* **1976**, *12*, 7–18.
- (17) Zefirov, N. S.; Tratch, S. S. Rearrangements and Cyclizations. XV. Tautomerism: Common Problems, Classification, the Search for New Topological and Reaction Types. *Zh. Org. Khim.* **1976**, *12*, 697–718.
- (18) Zefirov, N. S.; Tratch, S. S. Systematization of Tautomeric Processes and Formal-Logical Approach to the Search for New Topological and Reaction Types of Tautomerism. *Chem. Scr.* **1980**, *15*, 4–12.
- (19) Zefirov, N. S. An Approach to Systematization and Design of Organic Reactions. *Acc. Chem. Res.* **1987**, *20*, 237.
- (20) Tratch, S. S.; Zefirov, N. S. Problems of Molecular Design and Computers. VII. Formal-Logical Approach to Organic Reactions. Main Notions and Terminology. *Zh. Org. Khim.* **1982**, *18*, 1561–1583.
- (21) Zefirov, N. S.; Tratch, S. S. Problems of Molecular Design and Computers. VIII. Reaction Fragments and Classification Equations in Formal-Logical Approach to Organic Reactions. *Zh. Org. Khim.* **1984**, *20*, 1121–1142.
- (22) Dugundji, J.; Ugi, I. An Algebraic Model of Constitutional Chemistry as a Basis for Chemical Computer Programs. *Top. Curr. Chem.* **1973**, *39*, 19–64.
- (23) (a) Fujita, S. Description of Organic Reactions Based on Imaginary Transition Structures. 1. Introduction of New Concepts. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 205–212. (b) Fujita, S. Description of Organic Reactions Based on Imaginary Transition Structures. 2. Classification of One-String Reactions Having an Even-Membered Cyclic Reaction Graph. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 212–223. (c) Fujita, S. Description of Organic Reactions Based on Imaginary Transition Structures. 3. Classification of One-String Reactions Having an Odd-Membered Cyclic Reaction Graph. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 224–230. (d) Fujita, S. Description of Organic Reactions Based on Imaginary Transition Structures. 4. Three-Nodal and Four-Nodal Subgraphs for a Systematic Characterization of Reactions. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 231–237. (e) Fujita, S. Description of Organic Reactions Based on Imaginary Transition Structures. 5. Recombination of Reaction Strings in a Synthesis Space and Its Application to the Description of Synthetic Pathways. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 238–242. (f) Fujita, S. Formulation of Isomeric Reaction Types and Systematic Enumeration of Six-Electron Pericyclic Reactions. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 22–30.
- (24) Bissing, D. E.; Speziale, A. J. Reactions of Phosphorus Compounds. IX. The Opening of Epoxides with Tertiary Phosphines. *J. Am. Chem. Soc.* **1965**, *87*, 2683–2690.
- (25) Harary, F. *Graph Theory*; Addison-Wesley: Reading, MA, 1969.
- (26) Baskin, I. I. Systematic Generation and Analysis of Equations Expressing the Reaction Types of Organic Compounds. Ph.D. Thesis, Moscow State University, Moscow, 1990.
- (27) The graphic editor GIRL (graph input with regular lengths) has been developed by S. S. Tratch and O. L. Slovokhotova.
- (28) Zefirov, N. S.; Safronov, S. O.; Kozhushkov, S. I.; Chovnikova, N. G.; Baskin, I. I.; Magdesieva, N. N. Oxygen-containing Organic Derivatives of the V–VII Groups Elements in Reactions with Electronodeficient Acetylenes. *Proceedings of the 5th All-Union Symposium on Organic Synthesis*; Moscow, 1988; p 3.
- (29) Zefirov, N. S.; Kozhushkov, S. I.; Safronov, S. O.; Zhdankin, V. V. The Structure of the Product of the Reaction of Iodosobenzene with Acetylenedicarboxylate. *Zh. Org. Khim.* **1989**, *25*, 1109.
- (30) Zefirov, N. S.; Kozhushkov, S. I.; Yufit, D. S.; Struchkov, Yu. T.; Safronov, S. O. Molecular and Crystal Structure of the Product of the Reaction of Iodobenzene with Acetylenedicarboxylate. *Dokl. Akad. Nauk SSSR* **1989**, *309*, 627–630.
- (31) Takahashi, S.; Kano, H. 1,3-Dipolar Cycloaddition Reaction with 1-Methylbenzimidazole 3-Oxide. *Tetrahedron Lett.* **1963**, *25*, 1687–1691.
- (32) Stankevitch, M. I.; Tratch, S. S.; Zefirov, N. S. Combinatorial Models and Algorithms in Chemistry. Search for Isomorphisms and Automorphisms of Molecular Graphs. *J. Comput. Chem.* **1988**, *9*, 303–314.
- (33) Tratch, S. S.; Podymova, E. V.; Zefirov, N. S. An Algorithm for Generating the Complete Lists of Equations. Describing the Possible Reaction Types of Organic Compounds. *Proceedings of the 6th All-Union Conference on the Application of Computers in Molecular Spectroscopy and Chemical Studies*; Novosibirsk, 1983; p 205.
- (34) Herges, R.; Hoock, C. Reaction planning: computer-aided discovery of a novel elimination reaction. *Science (Washington, D.C.)* **1992**, *255*, 711–713.
- (35) The SYMBEQ program (executable file for IBM PC compatibles, source code for Microsoft QuickBasic and instruction for work) is available on either 3 1/2-in. or 5 1/4-in. disks (please specify which size disk or 5 1/4 will be sent as the default size) for a nominal cost of \$20 (check or money order only, do not send cash) from: PHM, Inc., 9520 Linden Ave., Bethesda, MD 20814. Phone: 301-493-6595. FAX: 301-897-3487.