GRAPH THEORY APPROACH*

Vladimír KVASNIČKA

Deparment of Mathematics,
Slovak Institute of Technology, 812 37 Bratislava

A graph-theory formalism of the organic chemistry is suggested. The molecular system is considered as a multigraph with loops, the vertices are evaluated by their mapping onto the vocabulary of vertex labels (*e.g.* atomic symbols). A multiedge of multiplicity *t* corresponds to a *t*-tuple (single, double, triple, *etc*) bond. The chemical reaction of molecular systems is treated by the transformation of graphs. The suggested graph-theory approach allows to formalize many notions and concepts that are naturally emerging in the computer simulation of organic chemistry.

Until now the graph theory[1,2] has been used in organic chemistry mainly in the following two fields: *1*) The enumeration problem of organic molecules with prescribed structure is solved by making use of the efficient tools of graph theory[1]. *2*) the theory of HMO is simulated by simple concepts of graphs, many results of HMO can be obtained directly from the structural formula of molecules with conjugated bonds[2].

The purpose of the present communication is to suggest a mathematical model of organic chemistry following the concepts and notions of the graph theory[3,4]. The obtained model allows to formulate exactly many problems emerging during the algorithmic stage of studies of the computer simulation or ogranic chemistry[5-6]. Many ideas and concepts of Ugi and Dugundji algebraic model of constitutional chemistry[8,9] are implanted in the present formalism; in particular, their **R** matrices are treated now as the so-called $\Re$ transformation of graphs. We believe that the present approach offers very effective formal "machinery" unifying many concepts of organic chemistry of rather diverse nature on the basis of simple graph-theory considerations. In order to keep the theory simplest as possible we postulate that the molecules contain even number of electrons and all bonds are realized by two electrons (in some extent this is identical with the restricted chemistry of Ugi and Dugundji[8,9]).

In physics the graphs are used very extensively, in particular, for the visualization of perturbation terms in quantum electrodynamics[10] and many-body theories[11,12]. Here, the graphs (or diagrams, following the physical terminology) are used as the efficient tool not only for the above mentioned visualization of perturbation terms

---

with complicated algebraic structure, but also as classification scheme. The perturbation terms are classified by topological notions (connected − disconnected, linked − unlined, ladder diagrams, *etc.*), this introduces the powerful heuristic essentially simplifying many considerations in forming new theoretical approaches. Recently, Rosensteel and coworkers[13] and Kvasnička[14] have formulated graph-theory approach for the unambiguous representation of the so-called Feynman-Goldstone diagrams. The present stage of graph theory in physics and chemistry is given in review article of Slanina[15].

In the forthcoming sections we shall use the physically oriented graph-theory terminology summarized by Essam and Fischer[4].

This terminology is partially modified to be more appropriate for our purposes. The molecular graph is defined as a multigraph with loops the vertices of which are evaluated by atomic symbols, all the concepts of graph theory (*e.g.* automorphism, isomorphism, *etc.*) should be now determined in such a way that the evaluation of vertices is properly taken into account.

Basic Concepts

A *vertex st* $V = \{v_1, v_2, ..., v_N\}$ is a set composed of N vertices $v_1, v_2, ..., v_N$. A vocabulary $\mathfrak{B} = \{\alpha_1, \alpha_2, ..., \alpha_p\}$ is a set composed of p vertex labels $\alpha_1, \alpha_2, ~..., \alpha_p$. The vertex set V is surjectively[17] mapped onto the vocabulary $\mathfrak{B}$,

$$\varphi : V \to \mathfrak{B} . \qquad (1)$$

It means that each vertex is uniquely evaluated by a vertex label from the vocabulary $\mathfrak{B}$.

An edge is an unordered pair of distinct vertices from the vertex set V. The edge $[i, j]$ is incident with the vertices v_i and v_j and connects them. Two distinct edges are adjacent if they have a vertex in common. Two distinct vertices are adjacent if they are incident with the same edge. A multiedge of multiplicity t is a set of t edges incident with the same pair of distinct vertices, $\{[i, j]_1, [i, j]_2, ..., [i, j]_t\}$, and are said to be parallel edges. A valence of a vertex is the number of edges incident with that vertex. An edge set $E = \{e_1, e_2, ..., e_M\}$, where $e_i = [i_1, i_2]$, associated with the vertex set V is a set of M edges with vertices in V.

A loop is the pair obtained by taking the same vertex twice from the vertex set V. The loop $[i, i]$ is incident with the vertex v_i. A multiloop of multiplicity u is a set of u loops incident with the same vertex, $\{[i, i]_1, [i, i]_2, ..., [i, i]_u\}$. A loop set $L = \{l_1, l_2, ..., l_P\}$, where $l_i = [i_1, i_1]$, associated with the vertex set V, is a set of P loops with vertices in V.

A *molecular graph* (or simply *graph*) is an ordered 5-tuple $G = (V, E, L, \varphi, \mathfrak{B})$, where V is a non-empty vertex set, E and L are edge and loop sets both associated

with the vertex set V, and finally, φ is a surjective mapping (evaluation) (1) of the vertex set onto the vocabulary $\mathfrak{B}$. The concepts introduced are illustrated in Fig. 1. We see that an arbitrary molecular system belonging to the restricted chemistry may be simply represented by a graph, this representation is of one-one character.

An *automorphism* of the graph G is a one-one correspondence ω between its vertices,

$$\omega : V \to V \qquad (2)$$

which induces one-one correspondences χ and χ' between its edges and loops, respectively,

$$\chi : E \to E, \qquad (3a)$$

$$\chi' : L \to L. \qquad (3b)$$

The correspondence ω conserves the evaluation of the vertices,

$$\omega(v_i) = v_j \quad \text{implies} \quad \varphi(v_i) = \varphi(v_j). \qquad (4)$$

A pair of distinct vertices v_i and v_j satisfying $\omega(v_i) = v_j$ and $\omega(v_j) = v_i$ are called topologically equivalent. If a graph has only a trivial automorphism realized *via* the identity correspondence, then the given graph contains the vertices that are all topologically nonequivalent. The concept of automorphism is illustrated in Fig. 2.

Two vertex sets $V = \{v_1, v_2, ..., v_N\}$ and $V' = \{v'_1, v'_2, ..., v'_N\}$ with mappings φ and φ' onto the same vocabulary $\mathfrak{B}$ of vertex labels, $\varphi : V \to \mathfrak{B}$ and $\varphi' : V' \to \mathfrak{B}$, are called similar if and only if (iff) there exists a one-one correspondence ψ between
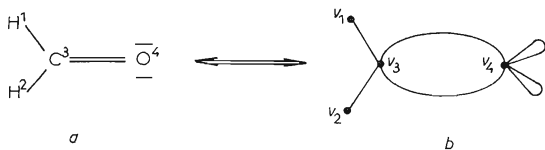

FIG. 1

The formaldehyde molecule is represented by the graph $G = (V, E, L, \varphi, \mathfrak{B})$, where the vertex set $V = \{v_1, v_2, v_3, v_4\}$, the edge set $E = \{e_1 = [1,3], e_2 = [2,3], e_3 = [3,4], e_4 = [3,4]\}$, and finally, the loop set $L = \{l_1 = [4,4], l_2 = [4,4]\}$. The mapping φ evaluates the vertices by the atomic symbols, $\varphi(v_1) = H$, $\varphi(v_2) = H$, $\varphi(v_3) = C$, and $\varphi(v_4) = 0$, the vocabulary $\mathfrak{B} = \{H, C, O\}$

them, $\psi : V \to V'$, which conserves the evaluation of vertices, *i.e.* $\psi(v_i) = v'_j$ implies $\varphi(v_i) = \varphi'(v'_j)$.

Two graphs $G = (V, E, L, \varphi, \mathfrak{B})$ and $G' = (V', E', L', \varphi', \mathfrak{B})$ are isomorphic iff the vertex sets V and V' are similar, and the one-one correspondence ψ induces a one-one correspondence between their edge and loop sets, respectively, see Fig. 3.

Two graphs $G = (V, E, L, \varphi, \mathfrak{B})$ and $G' = (V', E', L', \varphi', \mathfrak{B})$ are isohypsic iff the vertex sets V and V' are similar, and furthermore, these graphs satisfy

$$M + P = M' + P', \tag{5}$$

that is, the graphs G and G' have the same sum of numbers of edges and loops. If two graphs are isomorphic, then they are automatically isohypsic, but the converse is not necessarily true.

In order to illustrate these two concepts let us consider a pair of molecular systems obtained over the same set of atoms. If both the molecules are determined by the same structural formulas that are differing only in the indexing of atoms, then the corresponding graphs are isomorphic. Hence, with respect to a parent graph composed of N vertices (atoms or atomic cores) we can construct $N!$ graphs that are mutually isomorphic. This number is decreased for the parent graph containing topologically equivalent atoms. For example, the graph in Fig. 1 contains four vertices, in general we get 24 isomorphic graphs. Since the given graph has two topologically equivalent atoms (see Fig. 2) the total number of isomorphic graphs is $4!/2! = 12$. The concept of isohypsic graphs will represent in our forthcoming considerations the conservation of atomic cores and number of valence electrons.
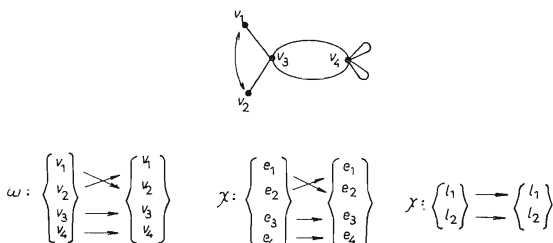

Fig. 2

A nontrivial automorphism of the graph presented in Fig. 1. We see that the vertices v_1 and v_2 are topologically equivalent

That is, a pair of molecules constructed over the same set of atoms (atomic cores) must be represented by a pair of graphs that are, *apriori*, isohypsic.

A *graph* $G' = (V', E', L', \varphi, \mathfrak{B})$ is a *subgraph* of the graph $G = (V, E, L, \varphi, \mathfrak{B})$ iff $V' \subseteq V$, $E' \subseteq E$, $L' \subseteq L$. The subgraph G' is a *spanning graph* of the graph G iff $V = V'$. The property of subgraph is denoted by $G' \subseteq G$.

A union of two graphs $G' = (V', E', L', \varphi', \mathfrak{B})$ and $G'' = (V'', E'', L'', \varphi'', \mathfrak{B})$ determined over the disjoint vertex sets $(V' \cap V'' = \phi)$ is defined by

$$G = G' \cup G'' = (V = V' \cup V'', E = E' \cup E'', L = L' \cup L'', \varphi, \mathfrak{B}), \qquad (6)$$

where the mapping φ is

$$\varphi(v) = \begin{cases} \varphi'(v) & \text{for } v \in V', \\ \varphi''(v) & \text{for } v \in V''. \end{cases} \qquad (7)$$

The resulting graph G is called disconnected, and its components are G' and G''. The difference is determined as $G - G' = G''$. The graphs G' and G'' are subgraphs of G.

The *adjacency matrix* of a graph G is a square symmetric matrix $\mathbf{A}$, the rows and columns of which are labeled by the vertices of the graph, the elements being determined by

$$a_{ii} = \text{the number of loops incident with the vertex } v_i, \qquad (8a)$$

$$a_{ij} = \text{the number of edges incident with both the vertices } v_i \text{ and } v_j. \qquad (8b)$$

In section 3 the graph G was determined as an ordered 5-tuple, *i.e.* $G = (V, E, L, \varphi, \mathfrak{B})$. Now, using the concept of adjacency matrix $\mathbf{A}$, we may give the second alternative determination of G as

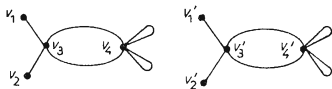$$G = (V, \mathbf{A}, \varphi, \mathfrak{B}), \qquad (9)$$

FIG. 3
The graphs G and G' are isomorphic, the mappings φ and φ' and the vocabulary $\mathfrak{B}$ are determined in Fig. 1

which is fully equivalent with the previous determination of G. In our forthcoming considerations we shall use both these definitions in a dependence on their particular advantages.

The adjacency matrix A of the graph in Fig. 1 looks like this

$$
A = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array}
\begin{array}{c} v_1\ v_2\ v_3\ v_4 \\
\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{pmatrix} \end{array}. \tag{10}
$$

Let us have two isomorphic graphs $G = (V, E, L, \varphi, \mathfrak{B})$ and $G' = (V', E', L', \varphi', \mathfrak{B})$, their adjacency matrices are A and A', respectively. The one-one correspondence ψ between vertex sets V and V' is simply realized by a permutation

$$
P = \begin{pmatrix} 1, & 2, & \dots, N \\ p_1, & p_2, & \dots, p_N \end{pmatrix}, \tag{11}
$$

the vertices of V and V' are realted by $v'_i = v_{p_i}$ and $\varphi'(v'_i) = \varphi(v_{p_i})$. Then the adjacency matrix A' is determined as follows

$$
A' = P^T A P, \tag{12}
$$

where P is an orthogonal $(P^T = P^{-1})$ matrix corresponding to the permutation (11).

The so-called matrix distance between two symmetric matrices X and Y is determined by

$$
l(X, Y) = \sum_{i \leq j} |x_{ij} - y_{ij}|, \tag{13}
$$

where x_j and y_j are the entries of X and Y, respectively.

Let us have two isohypsic graphs $G = (V, A, \varphi, \mathfrak{B})$ and $G' = (V', A', \varphi', \mathfrak{B})$, and let $\tilde{G}'$ be an isomorphic graph of G'. The adjacency matric of $\tilde{G}'$ is $\tilde{\ }A' = P^T A' P$, see Eq. (12). A distance between a pair of isohypsic graphs G and G' is the shortest matrix distance between adjacency matrices of G and an isomorphic graph $\tilde{G}'$ of G',

$$
d(G, G') = \min_P l(A, P^T A). \tag{14}
$$

The distance induces a metrics, the following three fundamental properties are satisfied:

1) The symmetry

$$
d(G, G') = d(G', G), \tag{15a}
$$

2) the triangular inequality

$$d(G, G') \leqq d(G, G'') + d(G'', G'), \qquad (15b)$$

3) the positive semi-definitess

$$d(G, G') \geqq 0 \quad (= 0 \text{ iff } G \text{ and } G' \text{ are isomorphic}). \qquad (15c)$$

One can say, a set of isohypsic graphs forms a metric space.

The notion of distance serves as a "measure" to quantify the similarity and/or dissimilarity of two isohypsic graphs (which are representing the molecular systems obtained over the same set of atoms). For a pair of isomorphic graphs, see Eq. (15c), we get the zero distance, this property corresponds to a situation when we have compared two molecules that are differing only in the indexing of atoms. Ugi and Dugundji[8,9] have introduced the notion of the so-called chemical distance between two ensembles of molecules by making use of the matrix distance (13). It seems that such a straighforward determination of the distance does not offer very suitable framework for the construction of a mathematical model of organic chemistry. Obviously, their determination of the distance depends on an actual indexing of vertices, this enters many additional formal difficulties that must be solved separately.

Reaction Operator

Consider two isohypsic graphs $G = (V, E, L, \varphi, \mathfrak{B}) = (V, \mathbf{A}, \varphi, \mathfrak{B})$ and $G' = (V, E', L', \varphi, \mathfrak{B}) = (V, \mathbf{A}', \varphi, \mathfrak{B})$, both constructed over the same vertex set V. Formally, we can say that the graph G is transformed onto the graph G' by a reaction operator,

$$\mathbf{R}G = G', \qquad (16a)$$

or conversely, the graph G' is transformed onto the graph G by a retro-reaction operator $^-\mathbf{R}$,

$$^-\mathbf{R}G' = G. \qquad (16b)$$

The reaction operator acts only on the edge and loops sets,

$$\mathbf{R}E = E' \quad \text{and} \quad \mathbf{R}L = L'. \qquad (17)$$

It rebuilds the initial sets of edges and loops onto the new sets E' and L' in such a way that the condition (5) is satisfied (*i.e.* the sum of numbers of edges and loops for both G and G' is conserved during the application of $\mathbf{R}$, this directly follows from our requirement that G and G' are isohypsic).

Following Ugi and Dugundji[8,9], the reaction operator **R** is represented by an (N, N) symmetric reaction matrix **R**,

$$\mathbf{R}G = \mathbf{R}(V, \mathbf{A}, \varphi, \mathfrak{B}) =_{\text{def}} (V, \mathbf{R} + \mathbf{A}, \varphi, \mathfrak{B}) \qquad (18)$$
$$= G' = (V, \mathbf{A}', \varphi, \mathfrak{B}),$$

or

$$\mathbf{R} + \mathbf{A} = \mathbf{A}'. \qquad (19)$$

This relation gives, in fact, the reaction matrix **R** in the following explicit form

$$\mathbf{R} = \mathbf{A}' - \mathbf{A}. \qquad (20)$$

The entries of **R** are simply interpreted as

$$r_{ii} = \text{the number of loops created } (r_{ii} > 0) \qquad (21a)$$
$$[\text{annihilated } (r_{ii} < 0)] \text{ incident with}$$
$$\text{the vertex } v_i,$$

$$r_{ij} = \text{the number of edges created } (r_{ij} = > 0) \qquad (21b)$$
$$[\text{annihilated } (r_{ij} < 0)] \text{ incident with}$$
$$\text{both the vertices } v_i \text{ and } v_j.$$

Since the graphs G and G' are isohypsic, the total numbers of created and annihilated edges/loops are exactly balanced, see Eq. (5), the entries of **R** must satisfy

$$\sum_{i \leq j} r_{ij} = 0. \qquad (22)$$

The retro-reaction matrix $^-\mathbf{R}$ corresponding to the retroreaction operator $^-\mathbf{R}$, see Eq. $(16b)$, is simply determined by

$$^-\mathbf{R} = -\mathbf{R}. \qquad (23)$$

Let $\tilde{G} = (\tilde{V}, \tilde{\ }\mathbf{A}, \tilde{\varphi}, \mathfrak{B})$, $\tilde{G}' = (\tilde{V}, \tilde{\ }\mathbf{A}', \tilde{\varphi}, \mathfrak{B})$ be a pair of graphs which are isomorphic of the original pair G, G', and let the one-one correspondence ψ between vertex sets V and $\tilde{V}$ be realized by a permutation (9). Then, following the relation (12), the adjacency matrices $\tilde{\ }\mathbf{A}$, $\tilde{\ }\mathbf{A}'$ are related with their original counterparts by

$$\tilde{\ }\mathbf{A} = \mathbf{P}^T\mathbf{A}\mathbf{P} \quad \text{and} \quad \tilde{\ }\mathbf{A}' = \mathbf{P}^T\mathbf{A}'\mathbf{P}. \qquad (24)$$

The graph $\tilde{G}'$ may be formally considered as a transformation of $\tilde{G}$ performed by a reaction operator $^\sim\!\mathbf{R}$,

$$\tilde{G}' = \,^\sim\!\mathbf{R}\tilde{G} . \tag{25}$$

The reaction matrix $^\sim\!\mathbf{R} = \,^\sim\!\mathbf{A}' - \,^\sim\!\mathbf{A}$, see Eq. (20), corresponding to the reaction operator $^\sim\!\mathbf{R}$, is simply determined as follows

$$^\sim\!\mathbf{R} = \mathbf{P}^\mathsf{T}\mathbf{R}\mathbf{P} . \tag{26}$$

It is easy to prove that the distance between two isohypsic graphs G and G', related by $(16b)$, is equal to the sum of the absolute values of upper-triangle entries of $\mathbf{R}$

$$d(G, G') = \sum_{i \le j} |r_{ij}| . \tag{27}$$

Consider a pair of isohypsic graphs $G = (V, E, L, \varphi, \mathfrak{B}) = (V, \mathbf{A}, \varphi, \mathfrak{B})$ and $G' = (V', E', L', \varphi, \mathfrak{B}) = (V', \mathbf{A}', \varphi, \mathfrak{B})$. In general, the vertex sets V and V' are similar, this directly follows from the assumption that G and G' are isohypsic. The distance between G and G' is determined by (14)

$$d(G, G') = \min_\mathbf{P} l(^\sim\!\mathbf{A}, \,^\sim\!\mathbf{P}^\mathsf{T}{}^\sim\!\mathbf{A}'{}^\sim\!\mathbf{P}) = l(^\sim\!\mathbf{A}, \,^\sim\!\mathbf{P}_0^\mathsf{T}, \,^\sim\!\mathbf{A}'{}^\sim\!\mathbf{P}_0) , \tag{28}$$

where $\mathbf{P}_0$ is the matrix representation of a permutation (11) for which the above extremal property is achieved. The adjacency matrix $^\sim\!\mathbf{A}' = \mathbf{P}_0^\mathsf{T}\mathbf{A}'\mathbf{P}_0$ determines a graph $\tilde{G}'$ which is isomorphic of G'. We introduce a class of all possible graphs $\{\tilde{G}', G'', \ldots\}$ that are isomorphic of G', of course, it contains automatically the graph G', its adjacency matrix $^\sim\!\mathbf{A}'$ has shortest matrix distance from the adjacency matric $\mathbf{A}$ of G.

Now, after these preliminary considerations we are ready to define the reaction operator $\mathbf{R}$ which transformes the graph G onto the class $\{G', G'', \ldots\}$ of graphs that are isomorphic of G'. As a representative of this class we choose the graph $\tilde{G}'$ with shortest matrix distance from G. Hence, treating a transformation of G onto G' we put

$$\mathbf{R}G = \tilde{G}' \in \{G_1', G_2', \ldots\} . \tag{29}$$

The reaction matrix corresponding to this process is determined by [cf. Eqs (20) and (28)]

$$^\sim\!R = \,^\sim\!\mathbf{A}' - \,^\sim\!\mathbf{A} = \,^\sim\!\mathbf{P}_0^\mathsf{T}{}^\sim\!\mathbf{A}'{}^\sim\!\mathbf{P}_0 - \,^\sim\!\mathbf{A} . \tag{30}$$

In order to reduce the result (30) onto (20) we have to assume that the graphs G

and G' are determined over the vertex set V (i.e. $V = V'$), then the permutation P_0 is equal to the identity permutation (the assigned matrix $\boldsymbol{P}_0$ is unity).

Bond, Synthon and Protosynthon

A bond $G(i, j)$ of a graph $G = (V, E, L, \varphi, \mathfrak{B})$ is a connected subgraph determined for a pair of adjacent vertices $v_i, v_j \in V$,

$$G(i, j) = (V_{ij}, E_{ij}, L_{ij}, \varphi, \mathfrak{B}), \qquad (31)$$

where

$$V_{ij} = \{v_i, v_j\} \subseteq V, \qquad (32a)$$

E_{ij} = a multiedge which is incident with both the vertices v_i and $\qquad (32b)$

$$v_j, (E_{ij} \subseteq E),$$

L_{ij} = a union of multiloops that are incident with the vertex v_i and $\qquad (32c)$

$$v_j, \text{ respectively}, (L_{ij} \subseteq L).$$

The bond $G(i, j)$ is t-tuple (single, double, triple, *etc.*) if the edge set E_{ij} is a multiedge of multiplicity t.

A *synthon* $G(Q)$ of the graph G is a subgraph determined for an index set Q,

$$G(Q) = (C_Q, E_Q, L_Q, \varphi, \mathfrak{B}), \qquad (33)$$

where

$$V_Q = \{v_i; i \in Q\} \subseteq V, \qquad (34a)$$

$$E_Q = \bigcap_{i \neq j \in Q} (\text{the multiedge incident with both the vertices } v_i \text{ and } v_j), \qquad (34b)$$

$$L_Q = \bigcup_{i \in Q} (\text{the multiloop incident with the vertex } v_i). \qquad (34c)$$

If $Q = \{i, j\}$ and the vertices v_i and v_j are incident, then the synthon $G(Q)$ is identical with the bond $G(i, j)$. The synthon may be further classified as connected/disconnected, this fully depends on the selection of Q.

Let $G(Q) = (V_Q, E_Q, L_Q, \varphi, \mathfrak{B})$ be a synthon of a graph $G = (V, E, L, \varphi, \mathfrak{B})$. A synthon complement $\tilde{G}(Q) = (\tilde{V}_Q, \tilde{E}_Q, \tilde{L}_Q, \varphi, \mathfrak{B})$ is determined as a subgraph of G, its union with the synthon $G(Q)$ produces the original graph G,

$$G = \tilde{G}(Q) \cup G(Q), \qquad (35)$$

where the following set of conditions must be satisfied

$$V = V_Q \cup \tilde{V}_Q \,, \tag{36a}$$

$$E = E_Q \cup \tilde{E}_Q \quad \text{and} \quad E_Q \cap \tilde{E}_Q = \phi \,, \tag{36b}$$

$$L = L_Q \cup \tilde{L}_Q \quad \text{and} \quad L_Q \cap \tilde{L}_Q = \phi \,. \tag{36c}$$

Hence, for a pair of a graph G and its synthon $G(Q) \subseteq G$ we may construct a synthon complement $\tilde{G}(Q)$ in such a way that the union of $G(Q)$ and $\tilde{G}(Q)$ gives the original graph G [Eq. (35)], and furthermore the synthon and its complement are edge and loop disjoint [Eqs (36b) and (36c)], whereas, in general, they can share a common vertex [Eq. (36a)]. For illustration of these concepts see Fig. 4.

Assuming that a reaction operator $\mathbf{R}$ acting on the graph G affects only the synthon $G(Q)$, the relation (16a) may be rewritten as follows

$$\begin{aligned}
\mathbf{R}G &= \mathbf{R}[G(Q) \cup \tilde{G}(Q)] \tag{37} \\
&= [\mathbf{R}G(Q)] \cup \tilde{G}(Q) \\
&= G'(Q) \cup \tilde{G}(Q) \\
&= G' \,,
\end{aligned}$$

that is, the synthon complement $\tilde{G}(Q)$ is intouched by the reaction operator $\mathbf{R}$, whole transformation is running only on the synthon $G(Q)$.

Let us have a class of graphs $\mathfrak{G} = \{G_1, G_2, \ldots\}$ and let $G(Q)$ be a common synthon of these graphs. This implies that an arbitrary $G_i \in \mathfrak{G}$ is equal to a union of the synthon $G(Q)$ and a synthon complement $\tilde{G}_i(Q)$ [determined with respect to both G_i and $G(Q)$ see above]
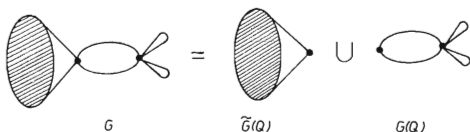
$$G_i = G(Q) \cup \tilde{G}_i(Q) \,. \tag{38}$$

G $\tilde{G}(Q)$ $G(Q)$

Fig. 4

Illustration of the concept of synthon $G(Q)$ and its synthon complement $(\tilde{G})Q$

Then an action of the reaction operator **R** affecting only their common synthon $G(Q)$ is determined by [Eq. (37)]

$$\mathbf{R}G_i = G'(Q) \cup \tilde{G}_i(Q) . \tag{39}$$

Summarizing, the approach of synthons allows to essentially simplify the formal treating of the action of reaction operator **R** on an arbitrary graph from the class $\mathfrak{G}$. In particular, the action of **R** on the graphs from the class $\mathfrak{G}$ is fully determined by $\mathbf{R}G(Q) = G'(Q)$, see Fig. 5.

Consider a disconnected graph $G = (V, E, L, \varphi, \mathfrak{B})$ composed of two components $G_1 = (V_1, E_1, L_1, \varphi, \mathfrak{B})$ and $G_2 = (V_2, E_2, L_2, \varphi, \mathfrak{B})$. Let $G_1(Q_1)$ and $G_2(Q_2)$ be synthons of G_1 and G_2, respectively, their union may be formally considered as a synthon of $G = G_1 \cup G_2$,

$$G(Q) = G_1(Q_1) \cup G_2(Q_2) , \tag{40}$$

where $Q = Q_1 \cup Q_2$. Conversely, the synthon $G(Q)$ of the disconnected graph G is said to be disconnected as well, and it is composed of components $G_1(Q_1)$ and $G_2(Q_2)$ that are again synthons of G_1 and G_2, respectively.

We introduce two classes of graphs $\mathfrak{G}_1 = \{G_1^{(1)}, G_1^{(2)}, \ldots\}$ and $\mathfrak{G}_2 = \{G_2^{(1)}, G_2^{(2)}, \ldots\}$. Assuming that the graphs $G_1^{(i)}$ and $G_2^{(j)}$ are determined over the different vertex sets, then their union $G_{ij} = G_1^{(i)} \cup G_2^{(j)}$ is simply a disconnected graph composed of two components $G_1^{(i)}$ and $G_2^{(j)}$. Let $G_1(Q_1)$ and $G_2(Q_1)$ be the common synthons of graphs

$$G_i =$$

$$\tilde{G}_i(Q)$$
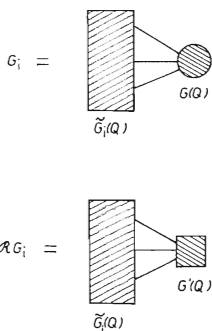
$$G(Q)$$

$$\mathscr{R}G_i =$$

$$\tilde{G}_i(Q)$$

$$G'(Q)$$

FIG. 5

An arbitrary graph $G_i \in \mathfrak{G}$ can be written as the union of the synthon and the synthon complement. The reaction operator **R** acting on the graph G_i affects only the synthon $G(Q)$, and transforms this synthon onto a modified synthon $G'(Q)$

from $\mathfrak{G}_1$ and $\mathfrak{G}_2$, respectively. The union $G(Q) = G_1(Q_1) \cup G_2(Q_2)$ (for $Q = Q_1 \cup \cup Q_2$) may be considered as a synthon of an arbitrary disconnected graph $G_{ij} = = G_1^{(i)} \cup G_2^{(j)}$. All the possible unions G_{ij} form a new class of graphs $\mathfrak{G} = \{G_{ij}\}$, formally $\mathfrak{G} = \mathfrak{G}_1 \cup \mathfrak{G}_2$.

The graphs of $\mathfrak{G}$, $\mathfrak{G}_1$, and $\mathfrak{G}_2$ have the same synthon $G(Q)$, $G_1(Q_1)$, and $G_2(Q_2)$, respectively, where $G(Q) = G_1(Q_1) \cup G_2(Q_2)$, for illustration see Fig. 6.

Let us assume that a reaction operator **R** acting on the graphs from $\mathfrak{G}$ affects only the common synthon $G(Q)$. Then the present formalism allows to treat simultaneously the class of disconnected graphs that have the synthon $G(Q)$ in common. Ghis is the main reason why the concept of synthons was introduced, it essentially simplifies the application of a reaction operator on the disconnected (and not only these) graphs.

Now we turn our attention to the adjacency-matrix representation of a synthon $G(Q) \subseteq G$. Following the relation (35) the graph G is equal to a union of the synthon $G(Q)$ and its synthon complement $\tilde{G}(Q)$, $G = G(Q) \cup \tilde{G}(Q)$. In the "adjacency--matrix" representation, see Eq. (9), we have

$$G(Q) = (V_Q, \mathbf{A}(Q), \varphi, \mathfrak{B}), \qquad (41a)$$

$$\tilde{G}(Q) = (\tilde{V}_Q, {}^{\sim}\mathbf{A}(Q), \varphi, \mathfrak{B}), \qquad (41b)$$

$$G \quad = (V, \mathbf{A}, \varphi, \mathfrak{B}). \qquad (41c)$$

The rows and columns of the adjacency matrix are labeled by the vertices of a given graph, this directly implies that the total adjacency matrix $\mathbf{A}$ may be expressed as a direct sum of $\mathbf{A}(Q)$, ${}^{\sim}\mathbf{A}(Q)$,

$$\mathbf{A} = \mathbf{A}(Q) \oplus {}^{\sim}\mathbf{A}(Q), \qquad (42)$$

FIG. 6

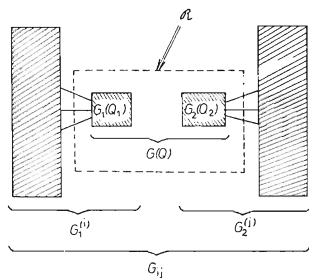The graph $G_{ij} \in \mathfrak{G}$ is composed of two components $G_1^{(i)} \in \mathfrak{G}_1$ and $G_2^{(j)} \in \mathfrak{G}_2$, each of these components contains the synthon $G_1(Q_1)$ and $G_2(Q_2)$, respectively. The total synthon of G_{ij} is $G(Q) = G_1(\omega_1) \cup G_2(Q_2)$. The reaction operator **R** affects only the synthon $G(Q)$, hence the rectangular shaded blocks are irrelevant for this transformation

which is a matrix counterpart of the set-theory relation (35). If a reaction operator $\mathbf{R}$ affects only the synthon $G(Q)$ [see Eq. (37)], then its reaction matrix may be formally considered as a "proto-matrix" reffered to the adjacencyc matrix $A(Q)$, we get

$$
\begin{aligned}
\mathbf{R} + A &= \mathbf{R} + \left[A(Q) \oplus {}^{\sim}A(Q) \right] \qquad\qquad (43) \\
&= \left[\mathbf{R} + A(Q) \right] \oplus {}^{\sim}A(Q) \\
&= A'(Q) \oplus {}_{-}A(Q) = A' .
\end{aligned}
$$

Finally, consider two reaction operators $\mathbf{R}_1$ and $\mathbf{R}_2$ both affecting only a synthon $G(Q)$. Then the adjacency matrix $A'(Q)$ of a transformed synthon $G'(Q) = \mathbf{R}_1\mathbf{R}_2G(Q)$ is given by

$$
A'(Q) = \mathbf{R}_1 + \mathbf{R}_2 + A(Q) \qquad\qquad (44)
$$

it means that, formally, the reaction operators $\mathbf{R}_1$ and $\mathbf{R}_2$ are commuting,

$$
\mathbf{R}_1\mathbf{R}_2 = \mathbf{R}_2\mathbf{R}_1 . \qquad\qquad (45)
$$

An *elementar reaction operator* $\mathbf{R}_{ij}$ transforms a graph $G = (V, E, L, \varphi, \mathfrak{B})$ with adjacent vertices v_i and v_j onto a graph $G' = (V, E', L', \varphi, \mathfrak{B}) = \mathbf{R}_{ij}G$, where

$$
E' = E - \{[i, j]\} , \qquad\qquad (46a)
$$

$$
L' = L \cup \{[j, j]\} . \qquad\qquad (46b)
$$

It means that the edge $[i, j]$ is annihilated in E to form E', and simultaneously, the loop $[j, j]$ is created in L to form L'. For completeness, if G does not contain the edge $[i, j]$, then the action of $\mathbf{R}_{ij}$ is not permitted.

In a reverse way we define an elementar retr-reaction operator $^-\mathbf{R}_{ij}$, it transforms a graph $G = (V, E, L, \varphi, \mathfrak{B})$, containing a loop incident with a vertex v_j, onto a graph $G' = (V, E', L', \varphi, \mathfrak{B})$, where

$$
E' = E \cup \{[i, j]\} , \qquad\qquad (47a)
$$

$$
L' = L - \{[j, j]\} . \qquad\qquad (47b)
$$

Hence, the edge $[i, j]$ is created in E to form E', and and simultaneously, the loop $[j, j]$ is annihilated in L to form L'. If the graph G does not contain the loop $[j, j]$, then the action of $^-\mathbf{R}_{ij}$ on the graph is not permitted. For illustration see Fig. 7.

Following the terminology of Kratochvil[16-19], the elementar reaction operator $\mathbf{R}_{ij}(\mathbf{R}_{ij})$ corresponds to the so-called nucleofug (electrofug) dissociation process

(*i.e.* $A_i - A_j \rightarrow A_i + \overline{A}_j$ and $A_i - A_j \rightarrow \overline{A}_i + A_j$). The elementar retro-reaction operator $^-R_j(^-R_j)$ corresponds to an inverse process called the nucleophilic (electrophilic) association (*i.e.* $A_i + \overline{A}_j \rightarrow A_i - A_j$ and $\overline{A}_i + A_j \rightarrow A_i - A_j$).

The matrix representation of these reaction operators is

$$R_{ij} = \begin{array}{c} \\ v_t \end{array} \begin{array}{c} v_i \quad v_j \\ \begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix} \end{array}, \qquad (48a)$$

$$^-R_{ij} = \begin{array}{c} \\ v_i \\ v_j \end{array} \begin{array}{c} v_i \quad v_j \\ \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \end{array} = -R_{ij}. \qquad (48b)$$

The elementar reaction operator R_{ij} (or $^-R_{ij}$) relates a pair of isomeric graphs G and G' with shortest possible distance

$$d(G, G') = 2, \qquad (49)$$

the total number of edges and loops is determined by

$$M = M' + 1 \quad \text{and} \quad P = P' - 1 \quad \text{(reaction)}, \qquad (50a)$$

$$M = M' - 1 \quad \text{and} \quad P = P' + 1 \quad \text{(retro-reaction)}. \qquad (50b)$$

In general, an arbitrary reaction operator R may be expressed as a product [unordered, see Eq. (45)] of the proper elementar reaction and/or retro-reaction operators. For instance, let us study a reaction operator R acting on a graph $G = (V, E, L, \varphi, \mathfrak{B})$

$$RG = G' = (V, E', L', \varphi, \mathfrak{B}), \qquad (51)$$

Fig. 7
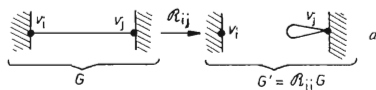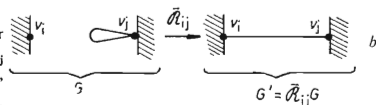Schematical illustration of the elementar reaction and retroreaction operators R_{ij} and $^-R_{ij}$, respectively. *a* Structural formula, *b* molecular graph

where

$$E' = \left(E - \{[i, j]\} \right) \cup \{[i, k]\} \,, \tag{52a}$$

$$L' = \left(L - \{[k, k]\} \right) \cup \{[j, j]\} \,. \tag{52b}$$

The resulting graph G' is isohypsic of G, $M = M'$ and $P = P'$. There is easy to see that the reaction operator $\mathbf{R}$ may be expressed *via* the elementar (retro-) reaction operators,

$$\mathbf{R} = \mathbf{R}_{ij}{}^{-}\mathbf{R}_{ik} = {}^{-}\mathbf{R}_{ik}\mathbf{R}_{ij} \,, \tag{53}$$

the corresponding reaction matrix is determined by

$$
\mathbf{R} = \begin{array}{c} \\ v_i \\ v_j \end{array} \begin{pmatrix} \overset{v_i \quad v_j}{0} & -1 \\ -1 & 1 \end{pmatrix} \oplus \begin{array}{c} \\ v_i \\ v_k \end{array} \begin{pmatrix} \overset{v_i \quad v_k}{0} & 1 \\ 1 & -1 \end{pmatrix}
$$

$$
= \begin{array}{c} v_i \\ v_j \\ v_k \end{array} \begin{pmatrix} \overset{v_i \quad v_j \quad v_k}{0} & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \,. \tag{54}
$$

Assume that the graph has a bond (synthon) $G(i, j) \subseteq G$. Then, following the relation (35), the graph is the union of $G(i, j)$ and its bond complement $\tilde{G}(i, j)$,

$$G = G(i, j) \cup \tilde{G}(i, j) \,. \tag{55}$$

The elementar reaction operator $\mathbf{R}$ (*i.e.* $\mathbf{R}_{ij}$, $\mathbf{R}_{ij}$, ${}^{-}\mathbf{R}_{ij}$, ${}^{-}\mathbf{R}_{ji}$) affects only the bond $G(i, j)$, we arrive at [Eq. (37)]

$$\mathbf{R}G = [\mathbf{R}G(i, j)] \cup \tilde{G}(i, j) = G'(i, j) \cup \tilde{G}(i, j) \,, \tag{56}$$

where an actual form of the resulting bond $G'(i, j) = \mathbf{R}G(i, j)$ is fully determined by the relations (46) and (47), respectively.

A *protosynthon* G_p is determined as a graph $G_p = (V_p, E_p, L_p, \varphi_p, \mathfrak{B})$ and let $\mathbf{R}_p$ be a *protoreaction operator* acting on whole G_p (*i.e.* not only on a subgraph of G_p). The pair $(G_p, \mathbf{R}_p)$ is formally considered as the prototype of a transformation $\mathbf{R}$ of an arbitrary graph $G = (V, E, L, \varphi, \mathfrak{B})$ which contains a synthon $G(Q)$ isomorphic of the protosynthon G_p, the reaction operator $\mathbf{R}$ is unambiguously related to the protoreaction operator $\mathbf{R}_p$.

Let $G(Q) = (V_Q, E_Q, L_Q, \varphi, \mathfrak{B}) \subseteq G$ be a synthon of G and $G(Q)$ be isomorphic

of the preselected protosynthon G_p. This implies that the vertex sets V_Q and V_p are similar, there exists such a one-one correspondence ψ between V_Q and V_p ($\psi : V_p \to V_Q$, saving the evaluation of vertices) which induces a one-one correspondence between their edge and loop sets. The protosynthon G_p may be alternative presented as $G_p(Q_p)$, where the index set $Q_p = \{1, 2, ..., n\}$ corresponds to an indexation of vertices in $V_p = v_1^{(p)}, v_2^{(p)}, ..., v_n^{(p)}\}$. Introducing an index set $Q'_p = \{\psi(1), \psi(2), ..., \psi(n)\}$ the reaction operator $\mathbf{R}$ can be now formally determined as follows

$$\mathbf{R} = \mathbf{R}_p(Q'_p) \,. \tag{57}$$

That is, it acts structurally in the same way as the original protoreaction operator $\mathbf{R}_p$, but now the vertices of modified (by the reaction) edges and loops are indexed according to the modified indexed set Q'_p.

The relation (57) can be rewritten in the matrix form, let us assume that the index set Q'_p differs from the original Q_p only in a permutation (11) of indices, i.e. $\psi(i) = p_i$, then the reaction matrix $\mathbf{R}$ assigned to the reaction operator $\mathbf{R}$ is see Eq. (26)

$$\mathbf{R} = \mathbf{P}^T\mathbf{R}_p\mathbf{P} \,, \tag{58}$$

where $\mathbf{R}_p$ is the reaction matrix of the protoreaction operator $\mathbf{R}_p$. The present concepts are illustrated in Fig. 8.
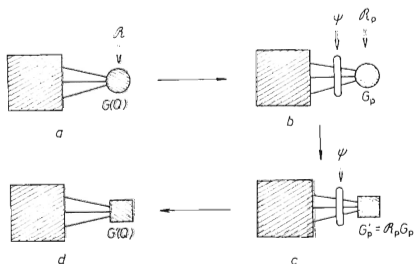

FIG. 8

Illustration of the concept of protosynthon G_p and protoreaction operator $\mathbf{R}_p$. (a) The graph G contains the synthon $G(Q)$ which is modified by the reaction operator $\mathbf{R}$. (b) The synthon $G(Q)$ and the reaction operator $\mathbf{R}$ are substituted by the protosynthon synthon G_p and the protoreaction operator $\mathbf{R}_p$, the pair $(G_p, \mathbf{R}_p)$ represents the so-called protoreaction. The edges are divided by a vertical oval block which represents a new indexation of vertices in G_p, an isomorphism of $G(Q)$ and G_p is represented by the $1-1$ mapping ψ. (c) The action of $\mathbf{R}_p$ on G_p gives a modified protosynthon G'_p. (d) Removing the vertical oval block the modified protosynthon G'_p is changed onto the modified synthon $G'(Q)$

Mechanistic Approach

Recently, Kratochvíl[16-19] has suggested the formal approach how to select for a given molecular system the proper chemical transformations (reactions). We turn our attention to the application of the present formalism developed in the previous section to give the exact graph-theory formulation of Kratochvíl's approach.

Let us study a connected graph $G = (V, E, L, \varphi, \mathfrak{B})$ which will be called the substrate graph. Applying successively a group of heuristic rules (formally expressed by a mapping Φ) on the substrate graph we get a class $\mathfrak{B}$ composed of a pairs of a strategic bond $B = G(i, j) \subseteq G$ and an elementar reaction operator $\mathbf{R}_{ele}$ [equal either to $\mathbf{R}_{ij}$, $\mathbf{R}_{ji}$, $^-\mathbf{R}_{ij}$, or $^-\mathbf{R}_{ji}$]

$$\Phi : G \to \mathfrak{B} = \{(B_i, \mathbf{R}_{ele}^{(i)}) ; \quad i = 1, 2, ..., a\} . \tag{59}$$

Hence, applying the mapping Φ the substrate graph G is evaluated by the class $\mathfrak{B}$, it contains a pairs of strategic bonds of G and assigned elementar reaction operators. In chemistry this concept is related to an initial mechanistic attack of a fixed bond.

Each pair $(B, \mathbf{R}_{ele})$ serves as an argument of the next mapping Ψ, the given pair is evaluated by class $\mathfrak{S}$ composed of b pairs $(G_p, \mathbf{R}_p)$, where G_p is a protosynthon and $\mathbf{R}_p$ is a protoreaction operator,

$$\Psi : (B, \mathbf{R}_{ele}) \in \mathfrak{B} \to \mathfrak{S} = \{(G_p^{(i)}, \mathbf{R}_p^{(i)}) ; \quad i = 1, 2, ..., b\} . \tag{60}$$

The mappings Φ and Ψ (that are realized by two groups of heuristic rules) evaluate the substrate graph G by pairs $(B, \mathbf{R}_{ele})$ and $(G_p, \mathbf{R}_p)$. The first pair determined an initial mechanistic attack of the bond B, whereas the second pair $(G_p, \mathbf{R}_p)$ corresponds (as will be specified below) to a standard transformation of the graph G.

Usually, the protosynthon G_p is a disconnected graph composed of two disjoint components, the first (second) one corresponds to the so-called *substrate* (reagent) subsynthon $G_{p,subs}$ ($G_{p,reag}$),

$$G_p = G_{p,subst} \cup G_{p,reag} . \tag{61}$$

We look for on the graph G a synthon $G(Q) \subseteq G$ which is isomorphic of $G_{p,subst}$. For a given standarr transformation [described by both the pairs $(B, \mathbf{R}_{ele})$ and $(G_p, \mathbf{R}_p)$] we have available a bank of reagent graphs (which is nothing else than a set of connected graphs). The next problem standing before us is to select from the given bank a reagent graph $\hat{G}$ [of course, as was mentioned above, with respect to the graph G and both the pairs $(B, \mathbf{R}_{ele})$ and $(G_p, \mathbf{R}_p)$]; formally, this can be realized by a mapping Ω,

$$\Omega : [G, (B, \mathbf{R}_{ele}), (G_p, \mathbf{R}_p)] \to \hat{G} . \tag{62}$$

In an actual algorithm, the mapping Ω is performed by making use of heuristic rules. Finally, we look for a on the selected reagent graph $\hat{G}$ a synthon $\hat{G}(\hat{Q}) \subseteq \tilde{G}$ which is isomorphic of $G_{\mathrm{p,reag}}$.

We have started in our considerations from a fixed substrate graph (connected) G, applying successively the mappings Φ, Ψ, and Ω we have the reagent graph $\hat{G}$ and two synthons $G(Q) \subseteq G$ and $\hat{G}(\hat{Q}) \subseteq \hat{G}$, where these synthons should be isomorphic of $G_{\mathrm{p,subst}}$ and $G_{\mathrm{p,reag}}$, respectively. The reaction operator $\mathbf{R}$ [constructed from the given protoreaction operator, see Eq. (57)] acts on the union $G \cup \hat{G}$ and it affects only the common synthon $G(Q) \cup \hat{G}(\hat{Q})$, we arrive at

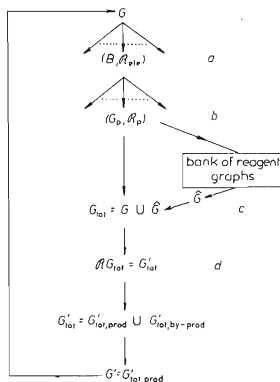$$\mathbf{R}G_{\mathrm{tot}} = G'_{\mathrm{tot}}\,, \tag{63}$$

FIG. 9

The general scheme of the approach described in section 16. The top root of this flowing chart is a preselected graph G. (a) Using the mapping Φ we get a pair $(B, \mathbf{R}_{\mathrm{ele}})$ composed of the so-called strategic bond $B \subseteq G$ and an elementar reaction operator $\mathbf{R}_{\mathrm{ele}}$. (b) The above pair $(B, \mathbf{R}_{\mathrm{ele}})$ is evaluated by the mapping Ψ onto the pair $(G_{\mathrm{p}}, \mathbf{R}_{\mathrm{p}})$ of a protosynthon and a proto-reaction operator, respectively. This second step represents the assignment of a protoreaction to a given preselected graph G. We look for in the bank of reagents the so-called reagent graphs $\hat{G}$, this is performed with respect to the protoreaction described by the pair $(G_{\mathrm{p}}, \mathbf{R}_{\mathrm{p}})$. (c) A chosen reagent graph $\hat{G}$ is unified with the original graph G, the result is denoted by G_{tot}. (d) The graph G'_{tot} is transformed by a reaction operator uniquely determined by the protoreaction operator $\mathbf{R}_{\mathrm{p}}$, the modified graph is denoted by G'_{tot}. (e) The modified graph G'_{tot} is composed (usually) of two components, the first (second) one represents the product (by-product) graph $G'_{\mathrm{tot,prod}}$ $(G_{\mathrm{tot,by-prod}})$. Finally, the graph $G'_{\mathrm{tot,prod}}$ is identified with the original input graph G and whole scheme is repeated

where $G_{tot} = G \cup \hat{G}$. In the case that the resulting graph G'_{tot} is disconnected and composed of two or more components, a first component is called the product graph $\left(G'_{tot,prod}\right)$ while the remaining ones are unified at a graph (connected or disconnected) called the by-product graph $\left(G'_{tot,by\text{-}prod}\right)$,

$$G'_{tot} = G'_{tot,prod} \cup G'_{tot,by\text{-}prod} \, . \tag{64}$$

This process is repeated cyclically (Fig. 9) in such a way that the graph $G'_{tot,prod}$ is identified with the substrate graph G at the beginning of the next cycle, schematically,

$$G = G_{subst}^{(1)} \xrightarrow{\textbf{R}_1} G_{subst}^{(2)} \dots \xrightarrow{\textbf{R}_{n-1}} G_{subst}^{(n)} \, . \tag{65}$$

The starting graph is transformed by a set of successive transformations $\textbf{R}_1, \textbf{R}_2, \dots$ $\dots, \textbf{R}_{n-1}$ onto the resulting (final) graph $G_{subst}^{(n)}$.

CONCLUSIONS

The purpose of this communication is to demonstrate that the graph-theory approach offers fruitful and suitable machinery how to formulate a mathematical model of organic chemistry. The structural formula of an organic molecule is represented by a multigraph with loops, its vertices are evaluated by the so-called vertex labels corresponding, for example, to atomic symbols. The concept of reaction operators allows to treat a transformation of a graph (*i.e.* a molecular system) onto another one, this is very serious aspect to the present theory which covers the chemical transformation of molecules. From the standpoint of the formal languages[20,21], the described method may be formally considered as a grammar generating the product graphs from a preselected substrate graph. In theory of formal languages are elaborated very powerful techniques for the syntactic analysis (parsing), they can be used also in the present graph-theory approach for retrosynthetic simulation of chemical transformations[22].

REFERENCES

1. Balaban A. T. (Ed.): *Chemical Application of Graph Theory*. Academic Press, London 1976.
2. Graovac A., Gutman I., Trinajstic N.: *Topological Approaches to the Chemistry of Conjugated Molecules*, Lecture Notes in Chemistry, Vol. 4. Springer-Verlag, Berlin 1977.
3. Harary F.: *Graph Theory*. Addition-Wesley, Reading, Mass. 1969.
4. J. W., Fisher M. E.: Rev. Mod. Phys. *42*, 271 (1970).
5. Wipke W. T., Heller S. R., Feldmann R. J., Hyde E. (Eds): *Computer Representation and Manipulation of Chemical Information*. Wiley, New York 1974.

6. Wipke W. T., Howe W. J. (Eds): *Computer Assisted Organic Synthesis*. American Chemical Society Symposium Series, No 61, 1977.

7. Wipke W. T.: *Computer Planning of Research in Organic Chemistry*. Plenum Press, New York 1976.

8. Dugundji J., Ugi I.: Topics Current Chem. *39*, 19 (1973).

9. Ugi I., Bauer J., Brandt J., Gasteiger J., Jochum C., Schubert W.: Angew. Chem. Int. Ed. (Engl. *18*, 111 (1979).

10. Schweber S. S.: *An Introduction to Relativistic Quantum Field Theory*. Row and Peterson, Evanson, Ill. 1961.

11. Goldstone J. A.: Proc. Roy. Soc. A *239*, 267 (1957).

12. Hugenholtz N. M.: Physica (Utrecht) *23*, 481 (1957).

13. Rosensteel G., Ihring E., Trainor L. E. H.: Proc. Roy. Soc. A *344*, 387 (1975).

14. Kvasnička V.: Int. J. Quantum. Chem.: *21*, 1003 (1982).

15. Slanina Z.: Chem. Listy *72*, 1 (1978).

16. MacLane S., Birkhoff G.: *Algebra*, p. 8. McMillan, New York 1967.

17. Kratochvíl M.: Chem. Listy *75*, 673 (1981).

18. Kratochvíl M.: Chem. Listy, in press.

19. Kratochvíl M., Koča J., Kvasnička V.: Chem. Listy, in press.

20. Hopcroft J. E., Ullman J. D.: *Formal Languages and their Relation to Automata*. Addison--Wesley, Reading, Mass. 1969.

21. Fu K. S.: *Syntactic Methods in Pattern Recognition*. Academic Press, New York 1974.

22. Kvasnička V.: This Journal *48*, 2118 (1983).

Translated by the author.