

- the Analysis of IR Spectra". *Appl. Spectrosc.* **1985**, 39, 331-333.
- (15) Tomellini, S. A.; Hartwick, R. A.; Stevenson, J. M.; Woodruff, H. B. "Automated Rule Generation for the Program for the Analysis of Infrared Spectra (PAIRS)". *Anal. Chim. Acta* **1984**, 162, 227-240.
- (16) French, P.; Clark, K. L. *LPA MacPROLOG User Guide*; Logic Programming Associates: London, 1985.
- (17) Visser, T.; van der Maas, J. H. "Systematic Computer-Aided Interpretation of Vibrational Spectra". *Anal. Chim. Acta* **1980**, 122, 357-361.
- (18) Cleij, P.; Dijkstra, A. "Information Theory Applied to Qualitative Analysis". *Fresenius' Z. Anal. Chem.* **1979**, 298, 97-109.
- (19) Dupuis, P. F.; Cleij, P.; van't Klooster, H. A.; Dijkstra, A. "Information Theory Applied to Feature Selection of Binary-Coded Infrared Spectra for Automated Interpretation by Retrieval of Reference Data". *Anal. Chim. Acta* **1979**, 112, 83-93.
- (20) Socrates, G. *Infrared Characteristic Group Frequencies*; Wiley: Chichester, UK, 1980.
- (21) Weinberg, F. *Grundlagen der Wahrscheinlichkeitsrechnung und Statistik sowie Anwendungen in Operations Research*; Springer Verlag: Berlin, 1968.
- (22) Shannon, E.; Weaver, W. *The Mathematical Theory of Information*; University of Illinois: Urbana, IL, 1947.

Description of Organic Reactions Based on Imaginary Transition Structures. 6. Classification and Enumeration of Two-String Reactions with One Common Node

SHINSAKU FUJITA

Research Laboratories, Ashigara, Fuji Photo Film Co., Ltd., Minami-Ashigara, Kanagawa, Japan 250-01

Received October 6, 1986

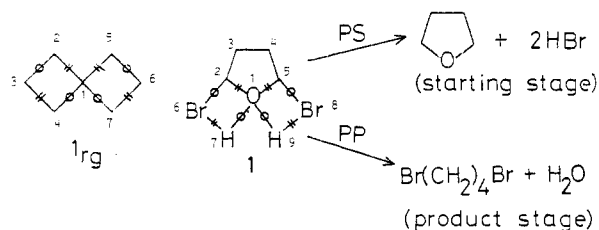
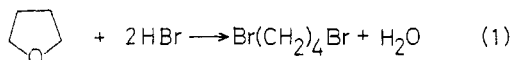
The number of reaction strings is a clue in the classification of organic reactions. Various two-string reactions are classified by their reaction graphs, each of which has two reaction strings sharing one node. The two-string reaction graphs are enumerated by Polyá's theorem.

Classification of organic reactions is an important problem to be solved in order to construct computer systems for retrieval of organic reactions and for synthetic design. Many methods have been reported for this purpose and reviewed from various points of view.^{1,2} We have proposed an imaginary transition structure (ITS) as a comprehensive representation of an individual organic reaction, which involves substrates and products as well as other components such as catalysts.³ In the ITS approach, we have introduced three colored bonds, i.e., an out-bond ($\text{---}\parallel\text{---}$), an in-bond ($\text{---}\ominus\text{---}$), and a par-bond (---).⁴ The ITS is a kind of structural formula, in which all nodes are connected by the three colored bonds in accordance with structural change during a reaction.⁵ From the ITS of an individual reaction, we have abstracted a reaction graph as a subgraph, which represents the corresponding reaction type. The reaction graph contains one or more reaction strings,⁴ each of which has alternate in-bonds and out-bonds and can be modified by par-bonds.

In the previous papers,^{3b,c} I have dealt with one-string reactions and enumerated trigonal, tetragonal, pentagonal, hexagonal, and octagonal reaction graphs. In this paper, I will describe two-string reactions and report that the present ITS approach has several advantages over other methods proposed for description of organic reactions.

ABSTRACTION OF REACTION STRINGS FROM AN ITS OR FROM A REACTION GRAPH

Formation of 1,4-dibromobutane by ring opening of tetrahydrofuran (entry 1)⁷ is represented by ITS 1, which can be

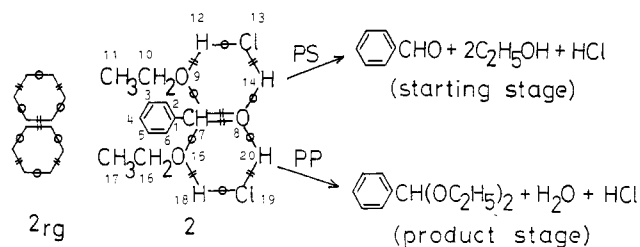
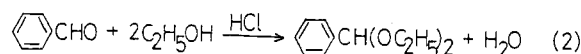


stored as a connection table shown in Table I. Several pieces of information concerned with this reaction can be abstracted

in the light of newly defined operations upon the ITS or ITS connection table.^{3a} For example, the starting and product stages are derived by PS and PP operations, respectively.⁸ The corresponding reaction graph (1_{rg}) represents a generic reaction type involving this reaction.

Two reaction strings, 1-2+6-7+1 and 1-5+8-9+1, can be abstracted graphically from ITS 1 when $-$ and $+$ represent an out-bond and an in-bond, respectively.⁹ These two reaction strings are stored by a connection table shown in Table II or by codes such as '1O'(1-1)2C(0+1)6Br(1-1)7H(0+1)'1O' and '1O'(1-1)5C(0+1)8Br(1-1)9H(0+1)'1O'. In the latter codes, the nodes and the ITS bonds³ are taken up, and the common node(s) shared by two reaction strings is (are) indicated by the single quotes. The same reaction strings are involved in the reaction graph 1_{rg} in more abstract fashion.¹⁰

Acetalization of benzaldehyde (entry 2)¹¹ leads to ITS 2, in which two reaction strings, i.e., 7+9-12+13-14+8-7 and



7+15-18+19-20+8-7, share two nodes (nodes 7 and 8). The reaction graph 2_{rg} abstracted from ITS 2 contains two hexagonal reaction strings sharing two nodes in a similar way.

In the case of multistring reactions, there is some ambiguity in abstracting reaction strings. For example, in the case of entry 1, one can visualize a single reaction string, i.e., 1-2+6-7+1-5+8-9+1, in the manner of a picture drawn with a single stroke of the brush. However, this single string is to be forbidden. For the acetalization of entry 2, a single reaction string, 7-8+14-13+12-9+7-8+20-19+18-15+7, is possible but not permitted in the present method.

To prevent such ambiguity in the adoption of reaction strings, I have established a criterion that *the same colored bonds incident to a node are preferred to form different reaction strings from each other if possible*.¹² In other words,

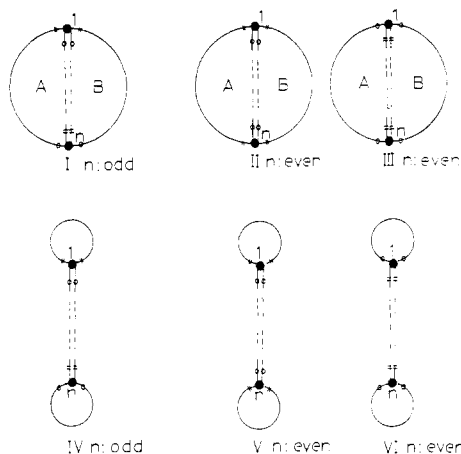
Table I. Connection Table of ITS 1

node	atom or group	coordinate		neighbor 1		neighbor 2		neighbor 3		neighbor 4	
		x	y	node	(a b)	node	(a b)	node	(a b)	node	(a b)
1	O	0	0	2	(1-1)	5	(1-1)	7	(0+1)	9	(0+1)
2	CH	-200	280	1	(1-1)	3	(1+0)	6	(0+1)		
3	CH ₂	-200	480	2	(1+0)	4	(1+0)				
4	CH ₂	200	480	3	(1+0)	5	(1+0)				
5	CH	200	280	1	(1-1)	4	(1+0)	8	(0+1)		
6	Br	-400	0	2	(0+1)	7	(1-1)				
7	H	-200	-280	1	(0+1)	6	(1-1)				
8	Br	400	0	5	(0+1)	9	(1-1)				
9	H	200	-280	1	(0+1)	8	(1-1)				

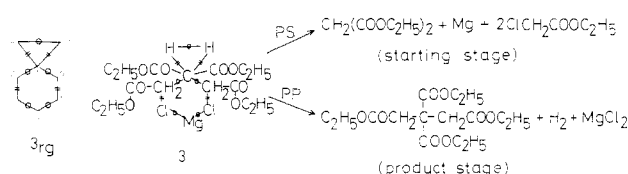
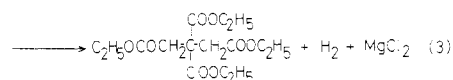
Table II. Connection Table of the Reaction Strings for the Reaction of Entry 1

node	atom	bond (a b)	neighbor 1		
			node	atom	(a b)
String 1					
1	O	(1-1)			
2	C	(0+1)	3	C	(1+0)
6	Br	(1-1)			
7	H	(0+1)			
String 2					
1	O	(1-1)			
5	C	(0+1)	4	C	(1+0)
8	Br	(1-1)			
9	H	(0+1)			
1	O				

a reaction string (RS) is preferred, if possible, not to contain any node twice. This criterion is explained schematically as follows. In the case of reaction graphs of types I, II, and III,



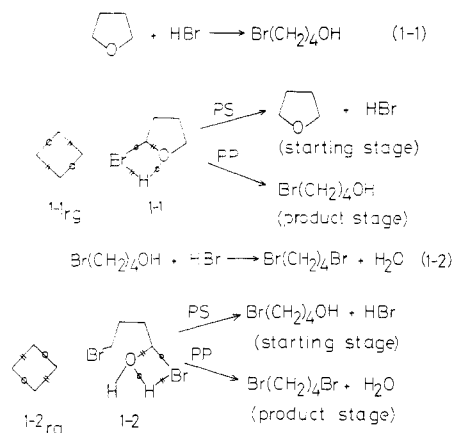
adoption of two reaction strings (A and B) is possible. However, reaction graphs of types IV, V, and VI are of one string. The following reaction (entry 3) is a representative



example.¹³ The reaction graph 3_{rg} contains a reaction string, 4+3-2+1-7+8-1+6-5+4, that cannot be divided in light of the above criterion.

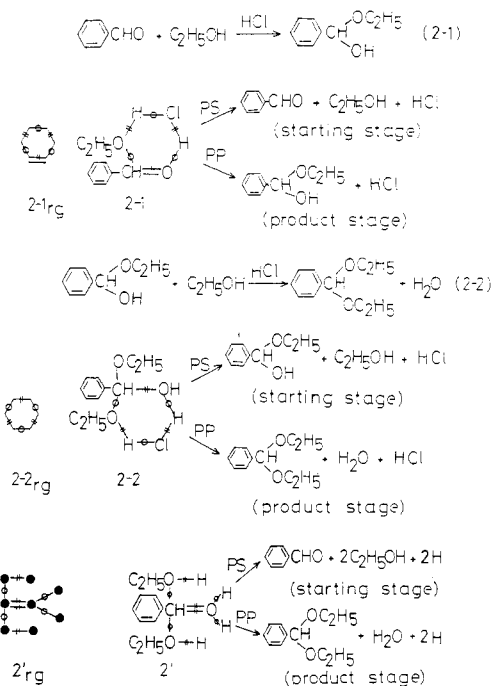
Divisible reaction graphs such as types I-III correspond to the fact that said reactions involve two (or more) primary

one-string reactions. For example, the reaction of entry 1 involves ring opening into 4-bromo-1-butanol (entry 1-1) and



the subsequent bromination of the hydroxyl group (entry 1-2). ITS 1 of the total reaction (entry 1) is the accumulated structure of ITS 1-1 and 1-2. And thus, adoption of two reaction graphs as described above is reasonable chemically.

The acetalization of benzaldehyde (entry 2) involves hemiacetal formation (entry 2-1) and subsequent conversion into the acetal (entry 2-2). The corresponding ITS's, 2-1 and 2-2, are both of one string, and their accumulation provides ITS 2 of two strings.



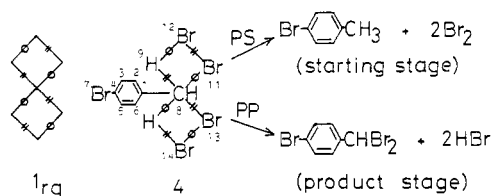
Sometimes, an abbreviated ITS such as 2' is convenient when one does not refer to the kind of catalyst. Two reaction strings, 12-9+7-8+13 and 17-14+7-8+18, are obtained also

in abbreviated ITS 2'. Thus, stringity (the number of reaction strings) should remain invariant in the abbreviation of multistring ITS's if possible. Otherwise, reaction features will be varied.

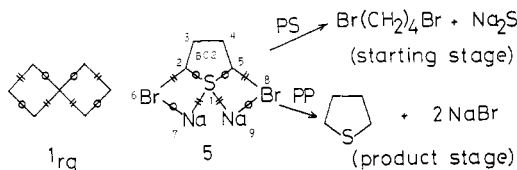
The present method is based upon graphical treatment; multistring reactions as well as one-string reactions are easily manipulated. This fact is important for retrieval of organic reactions and is an advantage of the present ITS approach over other methods proposed previously.¹⁴⁻¹⁷ In the following sections, I will deal with entangled (node-sharing) two-string reactions. The present discussion can be applied also to multistring reactions. Separate multistring reactions are regarded as simple ensembles of one-string reactions and are not referred to here.

EXAMPLES AND ENUMERATION OF TWO-STRING REACTIONS IN WHICH TWO TETRAGONAL REACTION STRINGS SHARE ONE NODE

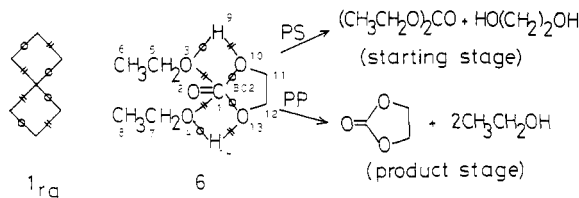
Bromination of a benzylic position (entry 4)¹⁸ provides an ITS (4) that involves two tetragonal reaction strings sharing one node, i.e., 8-9+12-11+8 and 8-10+14-13+8. Hereafter, I will omit reaction schemes but show starting and product stages resulting from PS and PP operations.



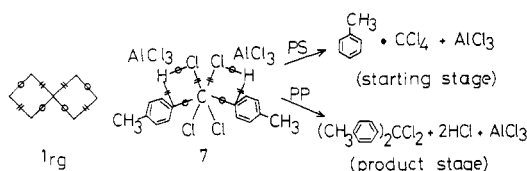
Formation of tetrahydrothiophene (entry 5)¹⁹ is represented by ITS 5, from which the same reaction graph (1rg) is abstracted. The ITS or the reaction graph contains two reaction strings, i.e., 2-6+7-1+2 and 5-8+9-1+5.



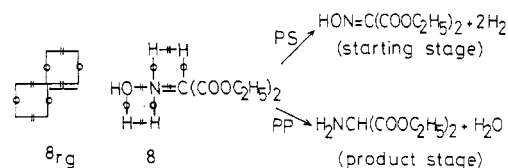
Formation of a cyclic carbonate via ester exchange²⁰ yields ITS 6, which contains reaction graph 1rg. Two reaction strings, 1-3+9-10+1 and 1-4+14-13+1, can be abstracted from ITS 6 or corresponding reaction graph 1rg. This reaction needs no catalysts.



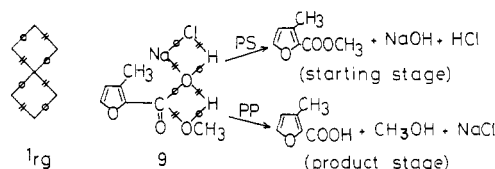
The Friedel-Crafts reaction forming diphenyldichloromethane affords ITS 7, which involves two reaction strings, 1-8+19-17+1 and 9-16+18-17+9. This reaction is catalyzed by aluminum chloride, which is, however, independent from the reaction graph abstracted as above.



Catalytic hydrogenation of oximes to amines²² provides ITS 8, the reaction graph of which is modified by a single par-bond.

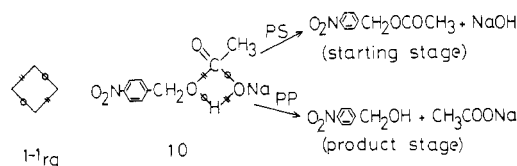


Alkaline hydrolysis of esters gives carboxylate anions. From the viewpoint of retrieval of organic reactions, however, free carboxylic acids are suitable subjects for indexing. For example, formation of 3-methyl-2-furoic acid by hydrolysis of methyl 3-methyl-2-furoate²³ is represented by ITS 9, which



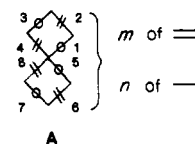
involves the subsequent neutralization with hydrochloric acid in an exact fashion.

When one regards an alkaline hydrolysis as formation of an alcohol, such neutralization as above need not be considered. Thus formation of nitrobenzyl alcohol²⁴ yields ITS 10 of one



string. It is emphasized that a variety of standpoints are present in the retrieval of organic reactions and will be satisfied by the present ITS approach.

The basic reaction graph of this type is represented by formula A. This graph (A) is a self-reaction pair, since it is



superimposable onto its reverse reaction graph, which is obtained by exchanging in-bonds and out-bonds with each other.²⁵ Enumeration of reaction types is translated to counting isomeric reaction graphs, in which the edges of the basic graph A are substituted by *m* double par-bonds (==) and *n* single par-bonds (—). This problem of counting isomers can be solved by use of Polya's theorem.²⁶

The basic graph (A) is superposed on itself by two permutation operations, (1)(2)(3)(4)(5)(6)(7)(8) and (1 5)(2 6)(3 7)(4 8), when the edges are numbered as shown. The cycle index *Z*(*C*₂) is represented by eq 1:

$$Z(C_2) = (1/2)(s_1^8 + s_2^4) \quad (I)$$

According to Polya's theorem, each variable *s_k* in the cycle index is substituted by a figure-counting series:

$$s_k = 1 + x^k + y^k \quad (II)$$

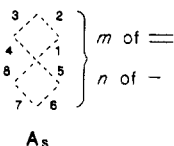
The resulting polynomial series *G*(*x*,*y*) is a reaction-graph-counting series in which the coefficient of *x^myⁿ* is the number of reaction graphs with *m* = bonds and *n* — bonds.

$$G(x,y) = Z(C_2, 1 + x^k + y^k) = \sum_{l+m+n=8} [(1/2)(8!/(l!m!n!)) + 4!/((l/2)!(m/2)!(n/2)!)] x^m y^n = 1 + 4x + 4y + 16x^2 + 28xy + 16y^2 + 28x^3 + 84x^2y + 84xy^2 + 28y^3 + 38x^4 + \dots \quad (III)$$

The summation does not involve the terms when *l*, *m*, and/or

n are odd.

Each reaction graph enumerated by eq III shows the relationship of a reaction pair or a self-reaction pair to an appropriate counterpart. The next problem is enumeration of all reaction pairs of this class. This enumeration is ascribed to counting the par-bond skeleton (A_s) substituted by m double



par-bonds ($=$) and n single par-bonds ($-$) on its edges. The cycle index $Z(D_2)$ and the corresponding polynomial series $P(x,y)$ for counting reaction pairs based upon A_s are calculated as shown in eq IV and V. The coefficient of $x^m y^n$ in $P(x,y)$

$$Z(D_2) = (1/4)(s_1^8 + 3s_2^4) \quad (IV)$$

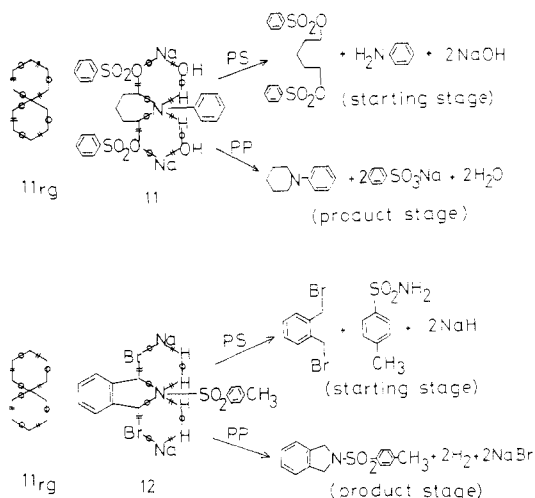
$$P(x,y) = \sum_{l+m+n=8} [(1/4)(8!/(l!m!n!) + 3 \times 4!/(l/2)!(m/2)!(n/2)!)] x^m y^n = 1 + 2x + 2y + 10x^2 + 14xy + 10y^2 + 14x^3 + 42x^2y + 42xy^2 + 14y^3 + 22x^4 + \dots \quad (V)$$

(eq V) is the number of reaction pairs with m double par-bonds and n single par-bonds. The summation is defined as in eq III.

The list of lower members of this class ($m = 0, n = 0-3$) is shown in Table III. Bridges of ring opening (BO_p),²⁷ of ring closure (BC_p),²⁸ and of rearrangement (BR)²⁹ are not indicated in Table III, but can be easily taken up.

EXAMPLES AND ENUMERATION OF TWO-STRING REACTIONS IN WHICH THE TWO HEXAGONAL REACTION STRINGS SHARE ONE NODE

Formation of a piperidine (ITS 11)³⁰ and of an isoindoline (ITS 12)³¹ involves the same reaction graph (11_{rg}), which



consists of two hexagonal reaction strings sharing one node. Reaction graph 11_{rg} constructs a self-reaction pair by itself.

Hydrolysis of an acid anhydride³² yields ITS 13 of two strings. The corresponding reaction graph is equal to 10_{rg} .

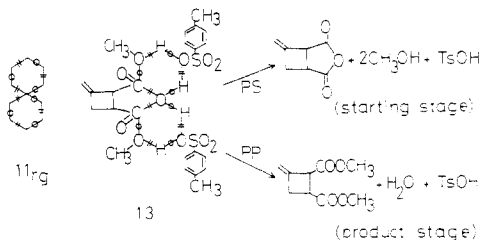


Table III. Reaction Graphs Based on the Graph A ($m = 0, n = 0-3$)^a

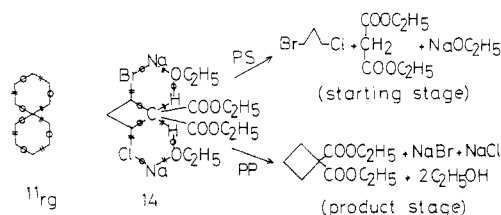
m	n	no. of reacn graphs ^b	no. of reacn pairs ^c	reaction graphs ^d
0	0	1	1	
0	1	4	2	
0	2	16	10	
0	3	28	14	

^a Table III may contain several reaction graphs that are not reasonable chemically, but they are not omitted for completeness of the list.

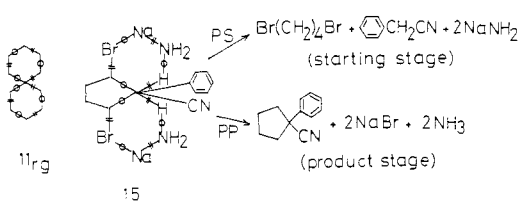
^b The numbers of reaction graphs are the coefficients of $x^m y^n$ in $G(x,y)$ (eq III). ^c The numbers of reaction pairs are the coefficients of $x^m y^n$ in $P(x,y)$ (eq V). ^d Each couple of reaction graphs linked with a bracket is a reaction pair.

Several dialkylation reactions at active methylene compounds provide ITS's having reaction graphs of two hexagonal reaction strings:

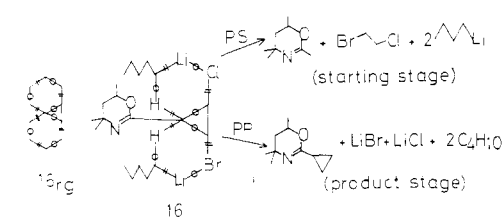
Cyclobutane formation³³



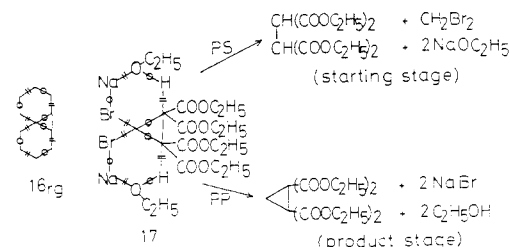
Cyclopentane formation³⁴



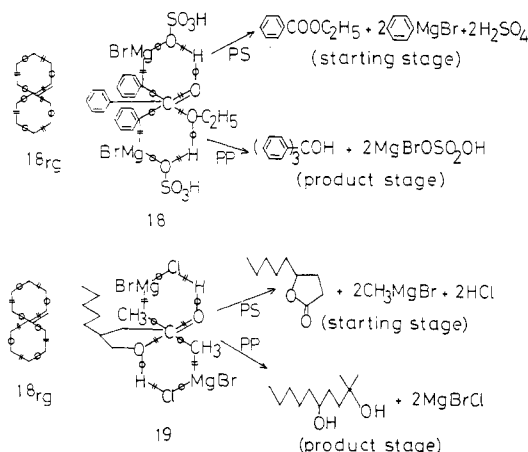
Cyclopropane formation³⁵



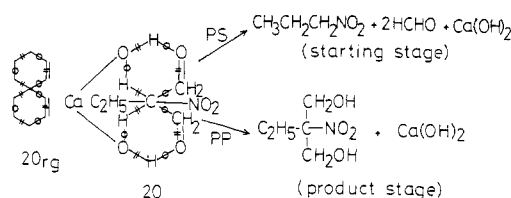
Another cyclopropane formation³⁶



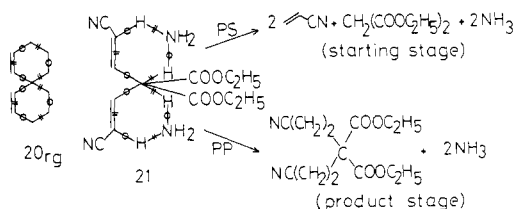
Addition of Grignard reagents to esters^{37,38} provides ITS's 18 and 19 of two strings. The reaction graph is a spiro-type modified by a single par-bond.



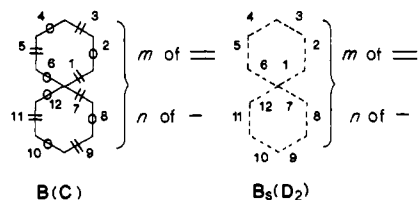
When a carbonyl group contributes to alkylation of an active methylene compound,³⁹ par-bond modification appears in the reaction graph (20_{rg}).



Double Michael addition of an active methylene compound⁴⁰ is represented by ITS 21, which has a spiro-type reaction graph modified by two single par-bonds.



Reaction graphs and reaction pairs of this class are enumerated on the same lines as described above. The cycle index $Z(C_2)$ for the basic graph **B** and the corresponding reaction-graph-counting polynomial series $G(x,y)$ are shown in eq VI and VII, respectively.



$$Z(C_2) = (1/2)(s_1^{12} + s_2^6) \quad (\text{VI})$$

$$G(x,y) = (1/2)((1+x+y)^{12} + (1+x^2+y^2)^6) = \sum_{l+m+n=12} [(1/2)(12!/(l!m!n!) + 6!/(l/2)!(m/2)!(n/2)!)] x^l y^m \quad (\text{VII})$$

For enumeration of reaction pairs based on **B_s**, the cycle index $Z(D_2)$ and the polynomial series $P(x,y)$ for counting reaction pairs are calculated as shown in eq VIII and IX, respectively.

$$Z(D_2) = (1/4)(s_1^{12} + 3s_2^6) \quad (\text{VIII})$$

$$P(x,y) = (1/4)((1+x+y)^{12} + 3(1+x^2+y^2)^6) = \sum_{l+m+n=12} [(1/4)(12!/(l!m!n!) + 3 \times 6!/(l/2)!(m/2)!(n/2)!)] x^l y^m \quad (\text{IX})$$

Table IV. Reaction Graphs Based on the Graph **B** ($m = 0, n = 0-2$)

m	n	no. of reacn graphs ^a	no. of reacn pairs ^b	reaction graphs ^c
0	0	1	1	
0	1	6	3	
0	2	36	21	

^aThe numbers of reaction graphs are the coefficients of $x^m y^n$ in $G(x,y)$ (eq VII). ^bThe numbers of reaction pairs are the coefficients of $x^m y^n$ in $P(x,y)$ (eq IX). ^cEach couple of reaction graphs linked with a bracket is a reaction pair.

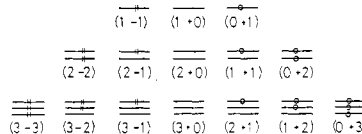
The coefficients of $x^m y^n$ in $G(x,y)$ and in $P(x,y)$ indicate the numbers of reaction graphs and of reaction pairs, respectively, which have m double par-bonds and n single par-bonds. The brief list of this class ($m = 0, n = 0-2$) is shown in Table IV.

CONCLUSION

An imaginary transition structure containing a reaction graph as its subgraph, manipulated as a three-color graph, is valuable in the classification of organic reactions. Various two-string reactions with one common node are classified in terms of their reaction graphs. Enumeration of representative two-string reaction graphs is accomplished by the use of Polya's theorem.

REFERENCES AND NOTES

- Sellow, G. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 249.
- Fujita, S. *J. Synth. Org. Chem., Jpn.* **1986**, *44*, 354.
- (a) Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 205. (b) Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 212. (c) Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 224. (d) Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 231. (e) Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 238.
- An out-bond has bond color of out and means a bond disappearing during a reaction. An in-bond is of in color and represents a bond formed. A par-bond is an invariant bond and has bond color of par.
- Bonds appearing in ITS's are referred to as ITS bonds or imaginary bonds. Each ITS bond is represented by a complex bond number (a/b), wherein integer a is the multiplicity of a bond in the starting stage and b is the difference in bond multiplicity between the product and the starting stage.



- A one-string reaction gives an ITS, or a reaction graph, that contains a reaction string. A two-string reaction provides an ITS, or a reaction graph, having two reaction strings, and so on.
- Fried, S.; Kleen, D. *J. Am. Chem. Soc.* **1940**, *62*, 3258. Entry numbers indicate respective reaction diagrams.
- Deletion of in-bonds of an ITS provides the starting stage of the reaction represented by said ITS. This operation is defined as *projection to starting stage* and is abbreviated PS. *Projection to product stage* is abbreviated PP, which is defined as deletion of out-bonds of an ITS.
- The numbering of nodes appearing in an ITS is arbitrary in this paper. An algorithm for canonical numbering is an important problem to be discussed in the future.
- The numbering of nodes appearing in a reaction graph may be equal to that of the corresponding ITS. Sometimes, it can be canonized for the purpose of classification of reactions.
- Haworth, R. D.; Lapworth, A. *J. Chem. Soc.* **1922**, 121, 76.
- Furthermore, an additional criterion is necessary in some cases. If out-bonds and in-bonds at each node are equal to one another in number, cyclic reaction strings of even members should be preferred to those of odd members.

- (13) Lund, H. *Ber.* **1934**, 67, 935.
 (14) Arens, J. F. *Recl. Trav. Chim. Pays-Bas* **1979**, 98, 155.
 (15) Rogers, D. C. *J. Org. Chem.* **1978**, 43, 1473.
 (16) Hendrickson, J. B. *Angew. Chem., Int. Ed. Engl.* **1974**, 13, 47.
 (17) Ugi, I.; Bauer, J.; Brandt, J.; Friedrich, J.; Gesleiger, J.; Jochum, L.; Shubert, W. *Angew. Chem., Int. Ed. Engl.* **1979**, 18, 111.
 (18) Coleman, G. H.; Honeywell, G. E. *Org. Synth.* **1943**, 2, 89.
 (19) Tarbell, D. S.; Weaver, C. J. *Am. Chem. Soc.* **1941**, 63, 2940.
 (20) Morgan, M. S.; Cretcher, I. H. *J. Am. Chem. Soc.* **1946**, 68, 781.
 (21) Gomberg, M.; Todd, J. D. *J. Am. Chem. Soc.* **1917**, 39, 2392.
 (22) Redemann, Dun J. *Biol. Chem.* **1939**, 130, 345.
 (23) Burness, D. M. *Org. Synth.* **1963**, 4, 628.
 (24) Hartman, W. W.; Rahrs, E. J. *Org. Synth.* **1955**, 3, 652.
 (25) I call this operation a transformation to the reverse reaction. The resulting pair of reaction graphs is referred to as a reaction pair. If a reaction graph is invariant with respect to this operation, it is defined as a self-reaction pair.
 (26) Polyá, G. *Acta Math.* **1937**, 68, 145.
 (27) A ring structure wherein p ITS bonds have complex bond numbers of $a + b = 0$ and all other ITS bonds have (a, b) of $a + b \neq 0$ and $a \neq 0$ is called a bridge of ring opening of order p (BO_p). The presence of BO_p corresponds to cleavage of p bonds.
 (28) A ring structure in which p ITS bonds have complex bond numbers (a, b) of $a = 0$ and all other ITS bonds have (a, b) of $a + b \neq 0$ and $a \neq 0$ is called a bridge of ring closure of order p (BC_p).
 (29) A ring structure that has one ITS bond of $a + b = 0$, another ITS bond of $a = 0$, and other bonds all of $a + b \neq 0$ and $a \neq 0$ is defined as a bridge of rearrangement (BR).
 (30) Reynolds, D. D.; Kenyon, W. O. *J. Am. Chem. Soc.* **1950**, 72, 1597.
 (31) Bornstein, J.; Shields, J. E. *Org. Synth.* **1973**, 5, 1064.
 (32) Stevenson, H. B.; Cripps, H. N.; Williams, J. K. *Org. Synth.* **1973**, 5, 459.
 (33) Mariella, R. P.; Raube, R. *Org. Synth.* **1963**, 4, 288.
 (34) Tilford, C. H.; van Campen, M. G.; Shelton, R. S. *J. Am. Chem. Soc.* **1947**, 69, 2902.
 (35) Meyer, A. I.; Adickes, H. W.; Politzer, I. P.; Geverung, W. N. *J. Am. Chem. Soc.* **1969**, 91, 765.
 (36) Eber, L. *Acta Chem. Scand.* **1958**, 21, 731.
 (37) Bachmann, W. E.; Hetzner, H. P. *Org. Synth.* **1955**, 3, 839.
 (38) Colonge, J.; Marey, R. *Org. Synth.* **1963**, 4, 601.
 (39) Vanderbilt, B. M.; Hass, H. B. *Ind. Eng. Chem.* **1940**, 32, 35.
 (40) Wakamatsu, S. *J. Org. Chem.* **1962**, 27, 1285.

Description of Organic Reactions Based on Imaginary Transition Structures. 7. Classification and Enumeration of Two-String Reactions with Two or More Common Nodes

SHINSAKU FUJITA

Research Laboratories, Ashigara, Fuji Photo Film Co., Ltd., Minami-Ashigara, Kanagawa, Japan 250-01

Received October 6, 1986

Two-string reactions that have two reaction strings with two or more common nodes are classified by their reaction graphs. Several representative two-string reaction graphs are enumerated by Polyá's theorem. Matrix representations of reaction graphs and their canonical forms are discussed.

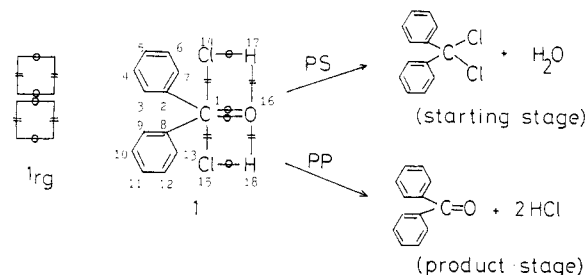
Name reactions have formed an important part of synthetic organic chemistry. However, manipulation of their information has troubled organic and information chemists because many of them contain multistep reactions. For example, the Claisen-Schmidt reaction, the Knoevenagel reaction, aldol condensation, and so on have been retrieved only in terms of their names, since there have been no effective methods for describing structural changes during the reactions.

An imaginary transition structure (ITS) proposed previously is a kind of structural formula with three colored bonds (out-, in-, and par-bonds).^{1,2} This novel formulation has afforded versatile methodology in the manipulation of organic reactions, even if these contain multiple steps. Reaction strings, which are extracted from an ITS as graphs with alternate out- and in-bonds, are indicators for the classification of organic reactions, especially of name reactions.³ Rings involved in ITS's are indices for recognition of ring opening and ring closure and of rearrangement reactions.

As a continuation of the preceding paper,¹ which gives several rules for selection of two or more reaction strings from an ITS, the present paper will discuss their application to further types of organic reactions.

CONNECTION TABLES OF ITS'S, OF RC GRAPHS, AND OF REACTION STRINGS

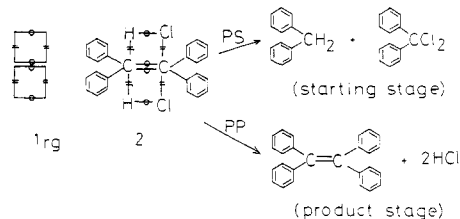
Formation of benzophenone from 1,1-dichloro-1,1-diphenylmethane is represented by ITS 1.^{3,4} The ITS is stored and manipulated by a computer in the form of a connection table of ITS's (Table I). The graph of reaction centers of level 1 (RC graph (1)) is a set of nodes that are incident to in- and out-bonds and the imaginary bonds between them. The corresponding reaction graph, in which values of nodes are ignored in a more abstract sense, is expressed by the formula 1_{rg} . The RC graph (1) corresponds to a connection table as shown in Table II.



Two reaction strings, 1-14+17-16+1 and 1-15+18-16+1, are abstracted graphically and represented by a connection table as shown in Table III.

EXAMPLES AND ENUMERATION OF TWO-STRING REACTIONS IN WHICH THE TWO TETRAGONAL REACTION STRINGS SHARE TWO NODES AND ONE ITS BOND

The condensation of diphenylmethane and dichlorodiphenylmethane⁵ is represented by ITS 2. The corresponding



reaction graph 1_{rg} has two tetragonal reaction strings, which fuse at one edge. The reaction strings are extracted as 1-14+29-16+1 and 1-15+30-16+1, wherein - and + represent out- and in-bonds, respectively.

Formation of a ketone from a *gem*-dichloride⁶ provides ITS 3 of two strings. The two tetragonal reaction strings share