

A Systematic Approach to the Classification and Nomenclature of Reaction Mechanisms

David C. Roberts¹

Contribution No. 3708 from the Department of Chemistry, University of California,
Los Angeles, California 90024

Received February 7, 1977

A universally applicable method for naming organic and inorganic reaction mechanisms is introduced. It is based on simple valence-bond representations of electron flow, and is thus conceptually and intuitively simple and easily grasped. Through its use, relationships between various types of concerted processes may be readily perceived, and their arrangement into classes naturally follows. Aside from its use as a perceptual or pedagogical tool, it is designed to be incorporated into chemical information systems, where it could potentially serve as the basis for a new literature search method.

More than four decades have passed since the introduction by Ingold² of his mechanistic nomenclature scheme and the familiar representational tool which has come to be known as "arrow pushing". His thinking has greatly changed the way we think about organic chemistry; an intuitive grasp of mechanistic principles is now essential to the practicing chemist.

Yet it remains difficult for students to assimilate this intuitive construct mostly because there exists no formalism relating different reactions on a mechanistic level. In an attempt to remedy this while involved in teaching the subject, this author has stressed similarities between concerted reactions with the same overall path of electron flow, hence referred to as *isomorphic* reactions. I found that this approach, which, for example, relates S_N2 , Elcb, nucleophilic addition to carbonyl, and nucleophilic addition to a nitrile group, could effectively impart a sense of cohesiveness to the seemingly random collection of reactions and mechanisms which threatens the beginning student of organic chemistry.

It further became apparent that the isomorphism concept could be extended to serve as the basis for a systematic means of naming, encoding, and classifying reaction mechanisms. With such means available, it would be routinely possible to uncover analogies which would be missed by the usual intuitive approach to mechanism. Information about mechanisms, and hence, reactions, could be concisely incorporated into information storage/retrieval systems such as those of Chemical Abstracts Service. Reactions could then be arranged into easily-searched indices; this would facilitate literature searches and allow new reactions to be placed into perspective with previously known mechanisms. The resulting broad overview of mechanism would lead to unexplored areas of chemistry, in much the same way as Mendeleev's overview of the properties of the elements, formalized as the periodic table, led to the discovery of new elements.

Some recent attempts to systematize organic mechanism should be mentioned. The work of Mathieu et al.,³ which was recently reworked by Guthrie,⁴ is basically a generalization and refinement of Ingold's original work, does not recognize the concept of isomorphic reactions, and can incorporate only simple mechanisms such as dissociation or displacement. Several departures from Ingold's work have been developed. Hendrickson⁵ has systematized and enumerated most of the known or hypothetical six-electron pericyclic reactions. He recognizes that these fall into several (five-center, six-center, seven-center) broad isomorphic classes, but his scheme cannot conveniently be expanded to include the remaining wide variety of reaction types. A similar approach has been developed independently by Stevens.⁶

Woodward and Hoffmann's classic treatment of orbital symmetry⁷ is similarly restricted to a limited set of reactions; here, the emphasis is on systems in which the molecular orbital makeup of a reaction influences its outcome rather than on developing a view of organic mechanism as a whole.

The only truly universal and systematic scheme that has appeared is that which Ugi and his co-workers developed for the encoding of chemical transformations.⁸ In this system, reactants and products are represented in the form of bond/electron (BE) matrices, and a transformation may be represented as an operator (R) matrix, which converts the reactant BE matrix to the product BE matrix by simple addition. It may be algebraically possible to extract concise symbols from R matrices which would represent general reaction types, but as it stands, the system is far too cumbersome to be used either as a teaching aid or in a manual information retrieval system.

In this paper, a prototype notation scheme⁹ will be introduced, and its application to a variety of types of organic reaction mechanisms will be described. This scheme is based on simple valence-bond representations of structure and mech-

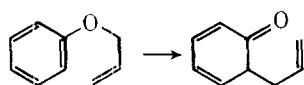
anism, and is easily grasped on an intuitive level. Yet it is capable of representing a mechanism in a manner which reveals on inspection the general and specific classes to which it belongs, its relationship to mechanisms in other classes, and its elements of symmetry.

The Concerted Process and Its Skeleton

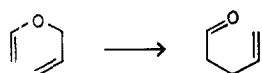
Any chemical transformation involving covalent bonds may be written as a linear sequence of one or more irreducibly synchronous or "concerted" electronic processes interspersed with processes of a translational or conformational nature. A concerted process (CP) may be defined as a single event mapping one discrete set of intermediates into another by a net flow of electrons. Although in the strictest sense mechanistic description of a concerted process must include a quantum mechanical description of this electron flow, the word "mechanism" as routinely used by experimental chemists does not include such descriptions and, partly because the scheme would otherwise be too cumbersome, a CP will hence be defined in terms of its initial and final states.

The problem of naming a stepwise reaction mechanism clearly reduces to the problem of finding a unique, systematic, and descriptive name for each of its component CPs. The overall mechanism may then be named simply by concatenation of the CP names into the proper sequence.

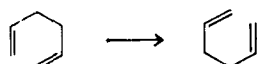
Consider now how we might dissect a CP in order to assess the features by which it might be named and classified. As an example let's use the first step of the familiar Claisen rearrangement.



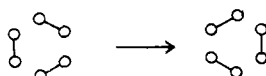
Starting with the standard valence-bond representation, we will follow a process of pruning away characteristics of increasing relevance to the nature of the CP. First we may eliminate substituents,



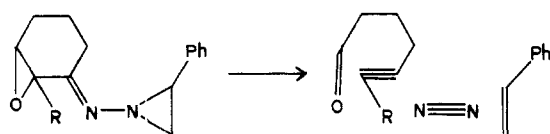
next we may eliminate heteroatoms,



and finally we may eliminate electrons which do not participate in the reaction.



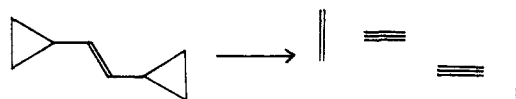
What we are left with is the skeleton of the CP, the dynamic electrons superimposed over a static array of nuclei, which for visual clarity will be represented as open circles.¹⁰ Note the emergence of symmetry during the dissection process. For the purposes of illustration, we may similarly treat a rather complex example, the recently reported Eschenmoser epoxyhydrazone fragmentation.¹¹



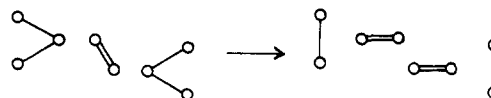
Eliminating substituents,



heteroatoms,



and nonparticipating electrons.



yields a CP skeleton again possessing a beautiful symmetry. Of course, a reaction such as the above fragmentation need not necessarily be concerted; the point in using it was to show how we might look at one particular mechanistic representation of the reaction.

Levels of Mechanistic Information

Looking at this dissection process in reverse, we may now set up a hierarchy of types of information about an individual CP. On the most fundamental level we have the CP skeleton, an array of nuclei whose topological disposition is defined by a superimposed array of dynamic (i.e., participating) electrons (Level 1). Of lesser importance in influencing the nature of the CP is the pattern of static bonding among the atomic centers (Level 2), the identity of the nuclei in the skeleton (Level 3), orbital symmetry factors (Level 4), and the substitution pattern onto the path of electron flow (Level 5). Exerting a still weaker influence are such factors as geometry of the molecules involved, as dictated by conformational preferences (Level 6), and the chemical environment, as well as remote structural features of the molecules which might exert some minor influence (Level 7). These categories might not always fall in this exact order, but the hierarchy remains useful for our purposes of classification. Categories of CPs corresponding to certain levels do already exist; for example, "electrocyclic" and "associative" correspond to Level 1, "extrusion" and "Alder-ene" correspond to Level 2, and "hydroboration" and "acyl transfer" correspond to Level 3. Yet these names clearly are useless in any systematic sense, and many classes in each level remain unnamed and unnoticed.

These levels of information further suggest how we might compile a reaction index for routine manual use. A CP skeleton may be viewed as an archetype, representing a broad class of isomorphic CPs. Within each such broad class there would be divisions at Level 2, each representing a subclass of reactions which are, in a loose sense, isoelectronic; within each such division there would be categories representing different patterns of atomic substitution. From this point on it would probably prove most convenient to list individual CPs as they actually appear, with reference to the overall reaction if the CP is an element in a multistep reaction. An abstractor compiling entries for the index would be expected to exercise common sense in breaking a reaction down into its individual CPs, but this introduces no human element that is not already present in today's abstracting/indexing systems.

To develop such a manual index, we will need a notation scheme which embodies the following characteristics: (a) it must be universally applicable; (b) encoding and decoding must be convenient and easily learned; (c) it must be adaptable to computer use; and (d) it must allow successive incorporation of information, so that, for example, a Level 3 name can be immediately identified as a member of a Level 2 class by reduction to the Level 2 name on inspection, and similarly down to Level 1.

In the next section, a notation scheme at Level 1, i.e., a means of generating a unique name for each CP skeleton (archetype), will be described. Following this will be a dis-

Table I. Two-Center CPs

Notation	CP Skeleton	Example
(10)	$\text{O} \cdot \quad \text{O} \longleftrightarrow \text{O} \cdots \text{O}$	$\text{H}^+ + \text{H}^+ \longrightarrow \text{H}_2^+$
(20)	$\text{O} : \quad \text{O} \longleftrightarrow \text{O} \text{---} \text{O}$	HETEROLYSIS
(11)	$\text{O} \cdot \quad \cdot \text{O} \longleftrightarrow \text{O} \text{---} \text{O}$	HOMOLYSIS
(22)	$\text{O} : \quad : \text{O} \longleftrightarrow \text{O} \equiv \text{O}$	CARBENE DIMERIZATION
(33)	$\text{O} : \quad : \text{O} \longleftrightarrow \text{O} \equiv \text{O}$	$2 \text{ N} \longrightarrow \text{N}_2$

discussion of ways in which these names may be modified to incorporate Level 2 and Level 3 information.

Level 1 Notation

If the process of dissection described above is applied to a variety of the most commonly occurring mechanisms, a relatively small number of archetypes are obtained.¹² These archetypes are of two distinct kinds: (a) those in which only bond formation or bond cleavage alone occurs, and (b) those involving simultaneous bond formation and cleavage. We might refer to type (a) archetypes as "primitive" CPs and type (b) archetypes might be called "compound" CPs. As will be explained later, the primitive CPs may be used as elements from which the compound CPs may be constructed by several kinds of homologation. The principle behind this homologation is that a primitive CP, taken in the "retro" sense, has a certain number of dynamic electrons available for participation in another CP. Starting with a primitive CP, we may replace one or more atomic centers with retro-CPs, or fragments thereof, having the same number of available dynamic electrons, and thus obtain a compound CP. A retro-CP functioning in this way as an element of a compound CP will hence be referred to as a "sub-CP".

For the purposes of classification, no distinction is made between forward and reverse directions; however, it is always assumed that a sub-CP proceeds in a direction formally opposite that of the compound CP in which it functions as an element.¹³

A. Symbols. 1. Parentheses. The basic form of the notation is that of a pair of parentheses containing a sequence of terms. A term is defined as either a numeral or a closed pair of parentheses. Simple ligation is represented by only two terms; more than two designates the joining of the components into a bonded ring.¹⁴ The order of these terms corresponds to the arrangement of the designated atoms or sub-CPs into the ring; larger numerals are placed to the left.

2. Numerals. An atomic center gaining or losing n electrons in a CP is represented by the numeral n . In general, $(mn) \equiv (nm)$, but it is useful to retain the convention of placing the larger numeral to the left.

3. Equals Sign. An atomic center or sub-CP gaining a different number of electrons than it loses in a CP is undergoing a valence change. This is represented by two terms representing its capacity in each of the two valence states separated by an equals sign. Its use will become clearer when demonstrated with examples (see section D, below).

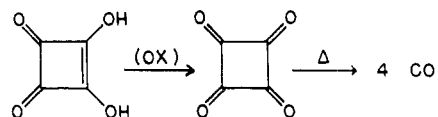
B. Primitive CPs. Some Examples. 1. Two-Center CPs. The symbols introduced above for the representation of CP skeletons (open circle = atomic center, dot = free electron, dotted line = one-electron bond, unbroken line = two-electron bond) will continue to be used in the tables that follow. Table I contains examples of known or possible two-center CPs. These will function not only in the capacity of sub-CPs, but also as models for the assembly of compound CPs from sub-CPs. The notation symbols are self-explanatory.

The list presented is not necessarily complete; other two-center CPs are, at least in principle, possible. The rarity of processes involving odd numbers of dynamic electrons or many dynamic electrons per atomic center is worthy of note.

2. Primitive Cyclic CPs. The simple ligation of more than two atomic centers at one time is assumed for our purposes to proceed in a cyclic fashion, resulting in the interconnection of the centers into a ring. The number of bonds to a given center should equal, or approximate as closely as possible, the number of dynamic electrons on that center at the outset. Table II contains an enumeration of some of the simpler primitive cyclic CPs. Moving from left to right in the table one finds an increase in the number of electrons per center; the CPs are arranged vertically into three major groups, those with two, three, and four centers. Clearly, the two-center CPs constitute special cases of primitive cyclic CPs and are shown in the table as such.

The notation symbols provide useful information on inspection. Note that the number of numerals between the parentheses always represents the number of atomic centers and that the sum of the numerals always equals the number of electrons in a CP.

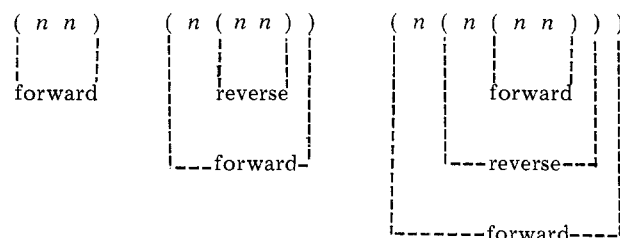
This table is included primarily for illustrative purposes. While the author is aware of no documented examples of cyclic reactions such as those in the table, one might imagine that reactions such as (2222) would occur in the retro sense upon pyrolysis of appropriate precursors:



The usefulness of the primitive cyclic CPs will become clear when they are used as prototypes from which the common cyclic CPs can be constructed by homologation, as illustrated in Table III. The sequence (22), (222), and (2222), illustrated by the diagonals in Table II, will serve in this capacity.

C. Compound CPs. Some Examples. These fall into three broad classes: (1) simple cyclic CPs, where one or more atomic centers is replaced by an entire two-center retro-CP; (2) linear CPs, where one or more atomic centers is replaced by only one center of a two-center retro-CP; and (3) complex CPs, where the role of sub-CP is filled by polyatomic CPs.

1. Simple Cyclic CPs. The top row of Table III contains the sequence of primitive cyclic CPs possessing two dynamic electrons per center. A two-electron center participates in a divalent sense in primitive CPs; a (11) CP, taken in the retro sense, exhibits one electron on each center and may function analogously to a two-electron center. The familiar cyclic CPs may be generated by substituting "(11)" for "2" in the notation and O-O for $\ddot{\text{O}}$ in the graph representation, as illustrated in Table III. Note that the inner parentheses imply a reaction occurring in the reverse sense from that of the overall reaction; if the latter were to function as a sub-CP in a CP of greater complexity, a further reversal would be implied:



Note also that the number of numerals encountered between two complementary parentheses represents the ring size of

Table II. "Primitive" Cyclic CPs

(10)	(20)	(30)	(40)				
	(11)	(21)	(31)	(41)	(42)		
			(22)	(32)	(33)		
	(111)	(211)	(311)	(411)	(421)	(431)	(441)
			(221)	(321)	(331)	(332)	(432)
				(222)	(322)	(422)	(333)
	(1111)	(2111)	(3111)	(4111)	(4211)	(4311)	(4411)
			(2211)	(3211)	(4121)	(4131)	(4141)
			(2121)	(3121)	(3221)	(4221)	(4321)
			(2221)	(3212)	(4212)	(4312)	(4412)
				(2222)	(3321)	(4231)	(4413)
					(3231)	(4222)	(4414)
					(3222)	(3331)	(4415)
						(3322)	(4416)
						(3232)	(4417)

a cyclic process and that the total number of electrons in a CP may still be obtained as the sum of the numerals.

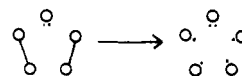
Some examples will clarify the symbols in Table III. In the leftmost column, we have a two-electron series; examples are heterolysis, (20), and attack of Br^+ on a double bond, ((11)0). In the four-electron series in the next column, carbene dimerization, (22), carbene addition to olefins, (2(11)), and 2 + 2 cycloadditions, ((11)(11)), serve as examples; known examples in the six-electron series are the reaction of phosphines with ozone, (22(11)), reaction of butadiene with sulfur dioxide, (2(11)(11)), Diels-Alder and many others, ((11)(11)(11)), and allyl + diene cycloadditions ((11)(11)(11)0). Examples exist of photochemical ((11)(11)(11)(11)) reactions, but the eight-electron series remains for the most part undeveloped. Several ten-electron reactions are known.

Hypothetical CPs may be constructed using other primitive CPs from Table II; for example, one might formulate several series of odd-electron CPs to complete Table III. But for the most part, known chemistry of purely cyclic CPs is limited to the examples given.

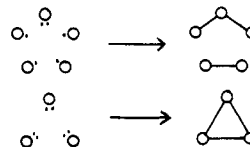
It is useful, in understanding the use of the notation, to consider how we might regenerate the graph representation of a CP from the notation symbols. Each term within the outermost pair of parentheses represents an element in a ring; these elements are so arranged (terms representing retro-CPs are represented in their bonded state) to give an initial state, as, for example, in a (2(11)(11)) reaction:



The retro-(11) operations are then performed, affording an entirely nonbonded state (which should not be interpreted as having any physical significance):



Finally, the binding of these elements into a fully bonded ring (except, of course, where prior bonds existed) affords the final state, in an operation analogous to the primitive CP containing the same number of terms (shown below for comparison).



One ambiguity of this system is that cyclic CPs with fewer than one electron per center result in cyclic arrangements with partial bonding. One-electron bonds are shown in the accompanying skeletal symbols; these are not necessarily implied by the notation but are fairly representative of reactions in these classes. It should be mentioned that ((11)(11)(11)0) can equally well represent both allyl + diene and pentadienyl + olefin type cycloadditions; the former is as shown in the table, but the latter could be thought to possess the skeleton which is represented by the same notation.

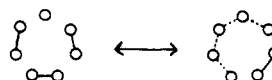
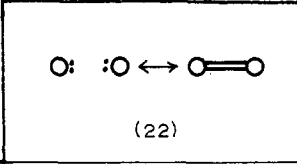
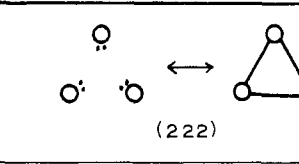
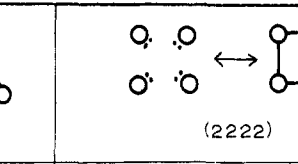
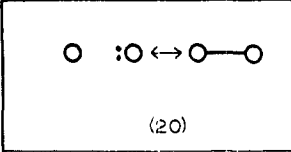
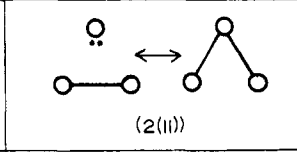
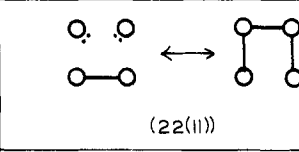
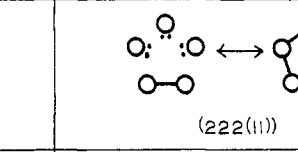
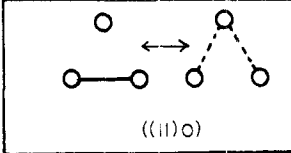
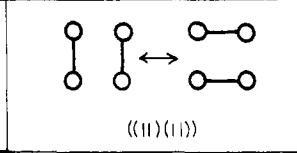
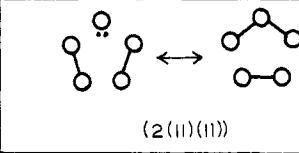
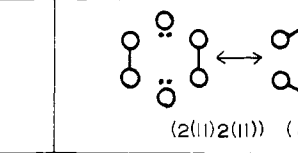
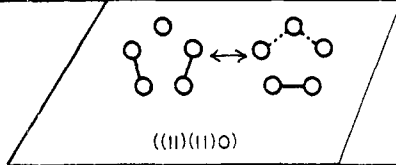
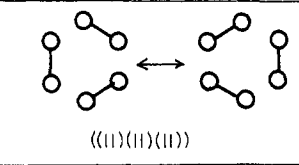
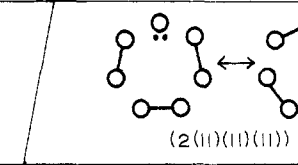
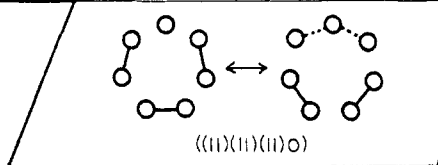
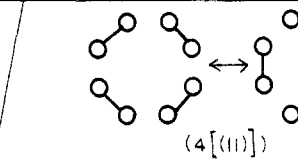
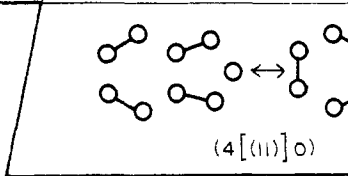


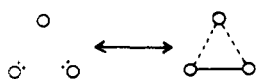
Table III. Simple Cyclic CPs

2. Linear CPs. Another series of well-known reactions may be generated by allowing only one center of a retro-CP (rather than both) to fulfill the function of an individual center of a two-center primitive CP. For example, a retro-(20) exhibits two dynamic electrons on one center and none on the other; each of these centers can combine as an element in another CP proceeding in the forward direction. A (2(0)2) reaction is thus formally defined as a 2 reacting with the 0 of a retro-(02); the notation is generated in a parallel fashion by placing a set of parentheses around the two components involved in the "forward" reaction:



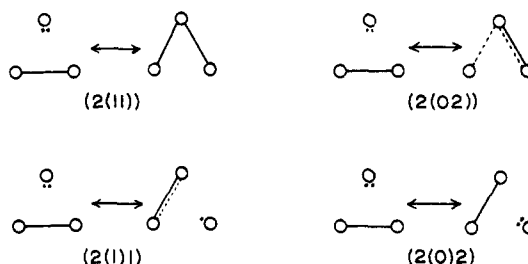
It should be noted that in order to represent a sub-CP a closed set of parentheses must contain two terms; hence, (2(0)2) cannot represent a cycloaddition such as



which, in fact, is properly represented by the notation (220). Two parentheses containing only one term rather represents the transfer of an entity between two others and does not itself represent a term since the parentheses are not a complementary pair. Again, the number of numerals contained within a complementary pair (in this case, two) represents the ring size of the sub-CP.

For the purpose of comparing the way linear and cyclic CPs

are generated from their component parts, the following examples are useful:



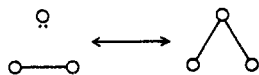
A (2(11)) CP is exemplified by carbene addition to an olefin, (2(0)2) represents reactions such as S_N2 , and the other two have no obvious counterparts in chemistry but are included for the purpose of demonstrating the use of the notation.¹⁵

Table IV enumerates some of the more commonly encountered linear CPs. It can be seen that these fall into several sequences of linear homologues, in which the notation symbols are built up in the manner prescribed above. It is also worth noting that the symmetry of a CP is preserved in the symbols and is thus immediately seen on inspection. Starting with the (11) reaction, a sequence of linear radical reactions is built up: (1(1)1), the radical abstraction, is followed by (1(1)(1)1), exemplified by fragmentation of an azo compound, and so on to cover radical additions to conjugated systems and multibond radical fragmentations. Another series beginning with (2(1)1) is, in principle, possible, although no real examples come to mind. The (20) series is quite fruitful, generating counterparts to a great deal of known ionic chemistry: nucleophilic and

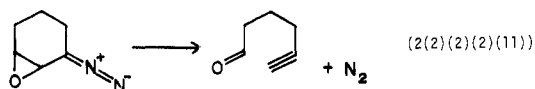
Table IV. Some Representative Linear CPs

electrophilic additions and substitutions correspond to (2(0)2) and (0(2)0), respectively; (2(0)(2)0) would represent protonation of an enamine; further out we have such reactions as protonation of a diene, general acid catalysis of carbonyl addition, ionic fragmentations, S_N2' , and so on. The (22) series calls to mind a variety of obscure and intriguing reactions involving carbenes and other amphiphilic atoms, such as carbene-carbene rearrangements and azine pyrolysis. More hypothetical possibilities arise by successive substitution of zero-electron centers; these are perfectly plausible but constitute little, if any, of known chemistry.

3. Linear-Cyclic, Spirocyclic, and Bicyclic CPs. These are constructed in an analogous fashion to simple linear CPs except that the role of sub-CP is filled by cyclic as well as diatomic CPs. Consider again the familiar (2(11) CP:

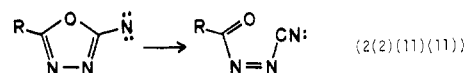


Taken in a retro sense, we have available a divalent center and a (11) sub-CP (also divalent) which may participate as components of a forward CP; some simple combinations which result are shown in Table V. One can generate from a linear-cyclic CP an entire series of homologous CPs in the same fashion as the linear series arise. An example from known chemistry is the fragmentation due to Eschenmoser,

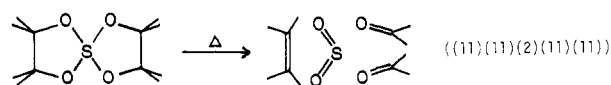
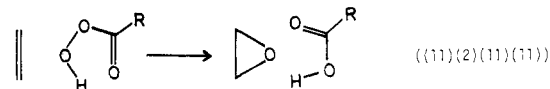


as well as ((11)(2)(2)(2)(2)(11)), which is the corresponding aziridine fragmentation introduced above as a CP skeleton.

Also known are linear cyclics such as



Bicyclic reactions are an essentially unexplored class; a few examples of higher spirocyclic reactions are known, such as



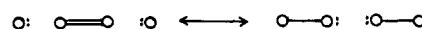
Clearly, a great deal of potentially useful chemistry lies waiting in these categories.

D. Valence Change. An assumption underlying the construction of complex CPs from sub-CPs is that an atomic center which is shared by two sub-CPs functions in the same capacity within each of the sub-CPs.

This is sometimes not the case. For example, consider the dimerization of a nitroso compound



whose skeleton may be written as



This is formally analogous to a (2(2)(2)2) CP,

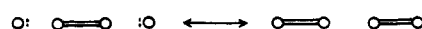
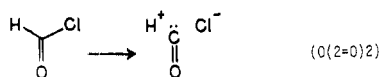


Table V. Complex CPs

<p>LINEAR-CYCLIC</p> <p>(2(2)) — (2(11)) (2(2)(11))</p>	<p>EXAMPLE:</p> <p><chem>C1CO1 + S(PR3)3 >> C1CO1=S(PR3)3</chem></p>
<p>BICYCLIC</p> <p>(2(11)) — ((11)2) (2((11)2))</p>	<p>HYPOTHETICAL:</p> <p><chem>C12OC1C2 + O=C >> C12OC1C2=S</chem></p>
<p>SPIROCYCLIC</p> <p>((11)2) — (2(11)) ((11)2(11))</p>	<p>HYPOTHETICAL:</p> <p><chem>C1OC2OC12 + O=C=O >> C1OC2OC12=S</chem></p>

but it differs in that nitrogen functions both as a nucleophile and an electrophile; it participates as a 2 in a (22) and as a 0 in a retro-(20). In a situation such as this, both numerals will appear in the notation, separated by an equals sign, and this combination will fulfill the function of a single numeral. The notation for this reaction is therefore $2(0=2)(2=0)2$. The term "valence change" is appropriate if we consider that in the context of mechanism the word valence would logically refer to the number of dynamic electrons on a given center (the numeral of the notation). A useful interpretation of the equals sign is as a combined right and left parenthesis, so that in evaluating and comparing notations the sub-CPs may be algorithmically extracted. As we read from right to left we have as sub-CPs $2(0)$, $(2)(2)$, and $(0)2$; the extraneous parentheses are preserved since these indicate the capacity in which a sub-CP functions. $2(0)$ indicates a retro-(20) CP in which the "0" center participates further. This unit may also be found in $2(0)2$, which is representative of, inter alia, addition of a nucleophile to a carbonyl group. The nitroso group does, in fact, function in this manner in nitroso dimerization. Likewise, $(2)(2)$ represents a retro-(22) in which each center is involved in another CP.

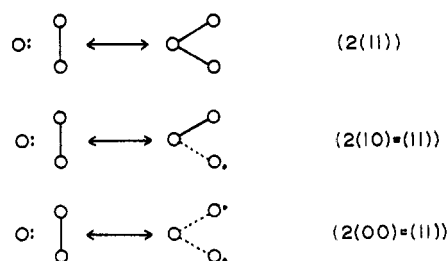
1,1-Eliminations and additions comprise another group of CPs embodying valence change:



Indeed, this is a primitive CP, in that it embodies only bond cleavage in the forward direction. As mentioned previously,¹⁴ these reduce to linear homologations of two-center CPs by interpretation of the equals sign and thus will not be confused with the corresponding primitive cyclic CP [in this case (220)].

The equals sign may be used to indicate valence change in a sub-CP as well as at an individual atomic center. For example, retro-(11) may function at least in principle as a bis-

univalent (11), univalent-zerovalent (10), or bis-zerovalent (00) component of a complex CP:

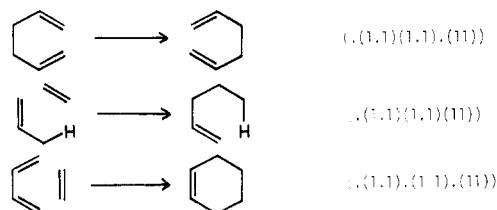


The equals sign is read, as before, as a combined right and left parenthesis. It should be evident that the Level 1 scheme as presented here may be adapted to fit CPs of almost any degree of complexity.¹⁶

Higher Level Notation

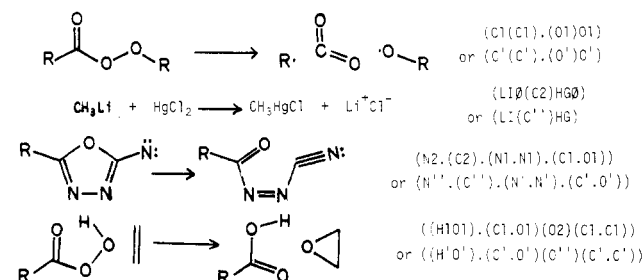
One major advantage of Level 1 notation as introduced above is that it can readily be made to incorporate further information about the CP. A Level 1-2 notation may be generated if we introduce the symbols "." and ":", which are to represent the fixed bonding pattern within the CP, and denote single and double static bonds, respectively. These symbols, when placed between two terms, denote the presence of fixed bonding between the corresponding atomic centers.

Consider, for example, a Cope rearrangement, an Alder-ene reaction, and a Diels-Alder reaction, which are all members of the ((11)(11)(11)) class. They are named as follows:



It is advantageous to assume a convention which places this punctuation as far to the left as possible (colon taking precedence over period); in this way ambiguity is avoided. Since, in the case of a cyclic CP, connection between first and last terms is assumed, a fixed bond connecting the corresponding atomic centers may be indicated by punctuation prior to the first term.

Further refinement, with incorporation of Level 3 information, is possible with the incorporation of symbols representing the elements. For this purpose, we may use either the internationally accepted symbols for the elements, converted to upper case, or otherwise the elemental symbols used by Wiswesser Line-Formula Notation,¹⁷ thus opening the possibility of incorporating even more structural information. These symbols might be incorporated immediately adjacent to the numeral representing the corresponding atomic center. An alternative would be to remove the numerals entirely, modifying elemental symbols with accent marks to correspond to the number of dynamic electrons on the corresponding center. Examples of both are given below:



It will again be advantageous in cases of ambiguity to adopt a convention which places lower atomic numbers to the left. Note that Level 3 information may be incorporated independently of Level 2 information, if this is desired.

The basic format of notation introduced in this paper is amenable to modification which incorporates essentially any desired amount of static structural information. Information about the specific electronic, vibrational, or conformational states of the components of a CP is much more difficult to incorporate, and this can probably be done only by adding a large variety of additional symbols and modifying the syntax. Yet the notation as it stands is capable of being quite specific; at this point it would seem practical to leave the rest to intuition.

The system as presented here may well contain flaws which give rise to ambiguities or redundancies, causing it to fail under certain circumstances. These can be uncovered only with use; it is, nevertheless, a useful tool in its present imperfect and informal state. It is adaptable to computer use as such, or with very slight modification. It has been pointed out that the scheme as presented here does not lend itself to oral rendition and therefore would be of limited value for teaching purposes. The systematic nomenclature of organic compounds suffers similarly, but we willingly tolerate such difficulties so that we may use it as a tool for searching the literature.

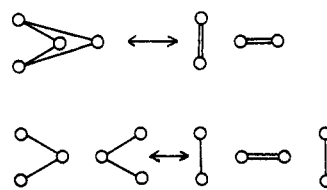
At present, this investigation will concentrate on compiling as broad a list of known CPs as possible. We wish to understand what rules permit a given CP to exist in the physical world; we will attempt to test new possibilities with this goal in mind, hoping that our results will lead us to new reactions of practical significance.

Acknowledgment. The author wishes to thank Professors D. S. Kemp and O. L. Chapman for financial support during the development of this project. The work was also supported in part by a National Institutes of Health Fellowship (No. 1 F32 CA05666-01) from the National Cancer Institute, DHEW.

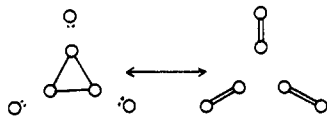
References and Notes

- (1) Address correspondence to this author at the Department of Chemistry, Rutgers-The State University, New Brunswick, N.J. 08903.
- (2) For a review, see C. K. Ingold, "Structure and Mechanism in Organic Chemistry", 2nd ed, Cornell University Press, Ithaca, N.Y., 1969, Chapters 5-15.
- (3) J. Mathieu, A. Allais, and J. Valls, *Angew. Chem.*, **72**, 71 (1960).
- (4) R. D. Guthrie, *J. Org. Chem.*, **40**, 402 (1975).
- (5) J. B. Hendrickson, *Angew. Chem., Int. Ed. Engl.*, **13**, 47 (1974).
- (6) R. V. Stevens, private communication.
- (7) R. B. Woodward and R. Hoffmann, *Angew. Chem., Int. Ed. Engl.*, **8**, 781 (1969).
- (8) J. Blair, J. Gasteiger, C. Gillespie, and I. Ugi, *Tetrahedron*, **30**, 1845 (1974).
- (9) The name NUMEN, for NUMerical MEchanistic Notation, is suggested. Presented here are versions NUMEN-1, NUMEN-12, and NUMEN-123, incorporating successively three different levels of information (see section on this topic).

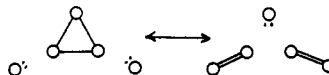
- (10) Graphs of this kind will be used throughout this paper; aside from the open circles as symbols for nuclei, all symbols will retain the same meanings as when used in traditional Lewis structures (dot = nonbonding electron, solid line = bonding pair, broken line = one-electron or fractional bond).
- (11) D. Felix, R. Muller, U. Horn, R. Joos, J. Schreiber, and A. Eschenmoser, *Helv. Chim. Acta*, **55**, 1276 (1972).
- (12) It is not within the scope of the present paper (nor within the power of the author) to make any statements concerning what kinds of CP archetypes are presently known or unknown; at this point the author is aware of about 30 types, but the list is certainly not complete. Ugi and co-workers⁸ claim 38 general *R*-matrix types with a maximum of three bonds broken during the transformation, but this refers to overall transformations rather than strictly to CPs.
- (13) It has been suggested that the distinction between forward and reverse directions be retained, the notation thus providing more information about the reaction in question. If this is desired, the notation as it stands may be designated as representing a reaction proceeding from an overall less bonded to a more bonded state, and a minus sign (-) preceding the notation would reverse this meaning.
- (14) Noncyclic multicenter primitive CPs will always reduce to linear homologations of two-center CPs; see section D in text.
- (15) The representation of the Diels-Alder reaction as ((11)(11)(11)) rather than ((20)(20)(20)) and similar representations of other reactions have been criticized by a referee on the grounds that the former notation implies an unpairing of electrons, which is in contrast with the molecular orbital picture of such reactions in which such pairing is maintained. The author recommends that the ((11)(11)(11)) notation be retained; first, because it correctly does not imply any preferred direction of electron flow around the ring, and second, because the substitution of (20) for (11) in all positions is valid only when no other terms are present, as illustrated by the comparison of (2(11)) and (2(02)).
- (16) Two further types of CPs are, except in special cases, not capable of being represented by the notation as it stands. Tricyclic or higher polycyclic CPs are one such type, and situations where a sub-CP shares each of more than two atomic centers with another sub-CP comprise the other. Both situations are highly unlikely, and it seems unnecessary to complicate the notation scheme just to accommodate them. The simplest example of the first type, the cycloreversion of tetrahedrane to two acetylene molecules, may still be represented by noting its similarity to a ((11)(2)(2)(11)) CP:



For the time being, it can be included in the latter class. Likewise, a CP such as



may be included in the (2(2)2(2)2) class:



Of course, any reaction for which it is impossible to write a traditional valence-bond mechanism (such as transition-metal catalyzed hydrocarbon isomerizations) is incapable of being represented within this scheme.

- (17) E. G. Smith and P. A. Baker, Ed., "The Wiswesser Line-Formula Chemical Notation (WLN)", 3rd ed, Chemical Information Management, Cherry Hill, N.J., 1975.