

VLGS ESTER OR VLGS THIOLESTER)
 #2950—JOURNAL NAME—J ORG CHEM
 TYPE—COMMUNICATION
 VOLUME 44 YEAR 79 PAGE 3437

KEYWORDS:

- | | |
|---------------------|-------------------------|
| 1. TIN ANION | 2. MIXED |
| 3. CUPRATE | 4. ADDITION-ELIMINATION |
| 5. VLGS ACID HALIDE | 6. BETA-X ENONE |
| 7. X=STANNYL | 8. * PIERSE |

#1732—JOURNAL NAME—J ORG CHEM
 TYPE—COMMUNICATION
 VOLUME 40 YEAR 75 PAGE 2694

KEYWORDS:

- | | |
|-------------------------|---------------------|
| 1. 6 RING | 2. VLGS ACID HALIDE |
| 3. MIXED | 4. CUPRATE |
| 5. ADDITION-ELIMINATION | 6. ENONE |
| 7. * PIERSE | |

This particular search found nine matches (two of which were selected to be displayed by the operator). The entries are always displayed in the order of decreasing entry number. Since this is a dynamic system where the keywords for current literature articles are constantly being added, this feature will generally display the most current information first.

A specific problem in using a keyword-based system for literature retrieval in organic chemistry is associated with the variability of organic nomenclature. There is much personal latitude in the use of hyphens and spaces in designating organic functional groups. For example, BETA-KETO-ESTER, BETA-KETOESTER, BETA KETO ESTER, and BETA KETOESTER may all be used to represent a single functional

group. Since the computer would normally perceive each of these as a distinct keyword, an additional routine was developed which removes all hyphens and spaces from keywords during addition, search, and editing procedures. However, the keyword is retained and displayed in the form first entered into the data base. Although the "compacted" form is never seen, the net effect is that the program prevents accidental redundancies from occurring.

In conclusion, it is felt that the PULSAR system provides a highly versatile tool for the organization and retrieval of literature information and should be also applicable in many areas outside of organic chemistry.

REFERENCES AND NOTES

- (1) Purdue University Undergraduate Research Associate.
- (2) A. P. Sloan Fellow, 1979-1981; Camille and Henry Dreyfus Teacher-Scholar, 1978-1983.
- (3) A. P. Sloan Fellow, 1977-1979.
- (4) Milne, G. W. A., *J. Chem. Inf. Comput. Sci.* **1980**, *20*, 204.
- (5) Roush, P. F.; Seitz, J. T.; Young, L. F. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 73.
- (6) Ziegler, H. J. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 141.
- (7) McCarn, D. B.; Leiter, J. *Science (Washington, DC)* **1973**, *181*, 318.
- (8) Van Ree, T., *J. Chem. Inf. Comput. Sci.* **1976**, *16*, 152.
- (9) Wilde, D. V.; Starke, A. C. *J. Chem. Doc.* **1974**, *14*, 41.
- (10) Copyright 1981, Purdue Research Foundation. The PULSAR program is commercially available from Litindex, Inc., P. O. Box 2274, West Lafayette, IN 47906.
- (11) The practice of one of the authors is to simply write keywords on the specific article as the current journal is being read. The article is then copied, a number stamped on it with a consecutive numbering machine, and it is stored in its entirety in a 3-ring binder. Entry of new articles to the data base can be performed by anyone familiar with the operation of a typewriter keyboard.
- (12) This designates a four-drive system where three drives are used for storage and one drive is available for copy purposes.
- (13) Vlgs is an abbreviation for vinylogous.

Present Status of Inorganic Chemical Nomenclature†

W. CONARD FERNELIUS

Kent State University, Kent, Ohio 44242

Received June 26, 1981

A systematic chemical name is one which portrays the essential structural features of a chemical compound by some general pattern. For most purposes it is unnecessary to write nomenclature rules in such detail as to provide a single name for each compound. Like all human activities nomenclature patterns change with time. This is essential to meet new conditions, to secure greater generality, or to obtain simpler names. While nomenclature specialists must be attuned to the needs of nomenclature users, their suggested solutions to be successful must be acceptable to the users. The presentation details with particular reference to inorganic chemistry (1) the committee-commission structure in this country and internationally, (2) significant accomplishments in the past half-century, (3) developments in progress, and (4) areas where nomenclature developments are needed.

Although the nomenclature of chemical compounds may appear confusing, if not meaningless, to the uninitiated, it is amazingly exact and simple in concept. Chemistry has benefited throughout its history by the nomenclature principles adopted by the early pioneers of the science.¹ They realized the importance of systematically relating names to the com-

position of the individual compounds. Their system with relatively minor modification served inorganic chemists well into the present century. Organic chemists early encountered marked inadequacies in the original systematic nomenclature and shifted to structure, rather than simple stoichiometry, as the basis for nomenclature. The adoption of structure as the basic consideration for names of inorganic compounds followed later but is now firmly established.

Many have asked, "Since structure is the important criterion for naming chemical compounds, why not dispense with names

† Adapted from an address on the occasion of the presentation of the Patterson-Crane Award of the Dayton and Columbus Sections of the American Chemical Society, May 22, 1981, Dayton, OH.

altogether and use structural formulas?". Aside from our historical desire for names and the habitual use of them firmly embedded in our behavior patterns, consider the difficulties of information retrieval if we relied solely on structures. So then the question "Why not rely on line notations and use these in place of names?". Once again, useful as such devices are in many ways, we are still conditioned to names and are deeply attached to them.

The major problem of chemical nomenclature is that of systematically relating the names of over five million compounds to their structures and doing this in such a way that each name uniquely indicates just one structure. Ideally, given a structure, one skilled in the art can produce a name and conversely given a name one automatically knows the structure. While this may seem very simple, there are a number of problems. There are two major patterns for accomplishing the objective: substitutive, most highly developed by organic chemists, and additive (including coordinative), most used by inorganic chemists. The very number of organic compounds caused organic chemists to become cognizant of the problems of systematic nomenclature before their inorganic compatriots. So those of us who labor largely in the inorganic area of chemistry have had greater difficulty convincing our colleagues of the need for systematic nomenclature and seem always to be playing a catch-up game with our organic brothers.

At the very outset of any discussion of names, one must deal with the question of whether, for a given compound, there be only one permitted name or several. In everyday society, we are rather ambivalent on this subject. If we discover that a person we know as John Smith is known as William Brown elsewhere or previously, we are immediately suspicious. However, we rather take for granted that actors will have a screen name other than their official name and we accept that writers may adopt a pseudonym. In chemical circles, any name is satisfactory provided it reproduces *unambiguously* the major structural features of the compound. Only in the case of large indexes and major compilations of chemical names is it essential that nomenclature rules be written in such detail as to result in one, and only one, name.

Beyond this point, nomenclature development is quite different from other chemical activities. Habits established by education and usage are deeply ingrained in chemists and are not easily changed by external pressure, although one does observe that nomenclature habits do change slowly with the times as topics of current interest change and as the chemical public tend to follow the usage of an outstanding investigator or that expressed in the more popular journals of the era. In order for changes in nomenclature to take place, there must be a desire on the part of users of names to make the change. Often this desire is lacking: e.g., names in use in the industry, trade names, advertizing directed toward the nonchemical public, and names widely used by those in disciplines other than chemistry.

There have been many proposals for new systems. Indeed, for the most part, these offer helpful suggestions. Yet they are seldom adopted in their entirety. Further, if they were, much of the earlier literature would become less available, and information retrieval systems would have to be revamped. So essentially a worker in nomenclature must deal with modification and improvement without loss of recognition of published work. The possibility of burning down and rebuilding cannot be seriously entertained. Further, the path for reform must be carefully chosen. To listen to and evaluate complaints is essential but not in itself sufficient. It is essential that proposed reforms be test marketed before becoming official. Reform is of no value if not used.

Thus nomenclature as an area of scholarly endeavor is quite different from the activities of most workers in chemistry.

While the true scholar is needed, he does not stand out among his colleagues as would a research worker, a textbook writer, a brilliant teacher, a research director, or even an historian of our science or technology. The developer of nomenclature not only must have input from colleagues to completely understand the problems and search out potential methods for solving them but also must test the acceptability of his solution in the market place. The ultimate test of success is acceptability, and the time span between recognition of a problem and the development of an acceptable solution is often so long that the scholar who initiated the reform may be lost among all of the other participants. Whatever publication results eventually is almost certain to be a joint effort or the report of a committee or commission. Now this is not the way in which one normally practices chemistry nor is it compatible with the way in which the reward system in Academia operates. If one reading an account of research deems that the facts are in error, one can go into the laboratory, repeat the experiment, and either confirm the facts or find the error. On the other hand, if one considers a theory unacceptable, one can derive a better one. In neither case must one pay great attention to human preference.

So in order to give you a review of what is going on in inorganic nomenclature today, I must discuss the subject in general terms. In so doing, I do not imply that more than a small fraction of the total is in any part mine. In reviewing what has been accomplished during my professional life and indicating what is going on today, I am only a reporter.

Since names are related to structures in systematic ways, it is essential that there be agreement on these ways. Further, insofar as possible, names in one language should be clearly recognizable in other languages. How then does a pattern of nomenclature arise, become adopted in one language or nation, and finally obtain international standing?

Several of the Divisions of the American Chemical Society have nomenclature committees. Each of these is represented on the overall nomenclature committee of the National Society. This committee reviews, modifies, and rejects reports from the divisional committees and passes them on to the appropriate nomenclature commission of the International Union of Pure and Applied Chemistry (IUPAC). At the very beginning of the International Union of Pure and Applied Chemistry, there were commissions on nomenclature. The Commission on the Nomenclature of Inorganic Chemistry (CNIC) has not fared so well as the Commission on the Nomenclature of Organic Chemistry (CNOC). CNIC was slow getting started and has not, in my opinion, maintained a satisfactory rate of performance since. Among the reasons for this have been the lack of care in getting a representative group of participants (e.g., there was no American representative until 1946), inability to sustain progress between meetings, somewhat imprecise wording of rules, and failure to recruit those with a driving interest in nomenclature. (An authority in a field of research is not necessarily an authority of nomenclature.) In recent years, Americans have carried a proportion of the work of the Commission well in excess of their number on the Commission. The biggest concern today is whether the IUPAC can be counted on to be an effective overall coordinating body. The size of its nomenclature commissions was reduced from 10 to 8 and may even be reduced to 6 in the near future and the permitted term of service reduced significantly. Certainly, nomenclature development will suffer unless somehow these limitations can be overcome.

We are now in a position to review what has happened in inorganic nomenclature during the last 40 years. The year 1941 is an important year for inorganic nomenclature because of two events: (1) a symposium on nomenclature at the St. Louis meeting of the American Chemical Society² and (2) the

appearance of a report by CNIC whose ongoing work had been brought to a sudden end by the outbreak of World War II.³ The symposium brought inorganic chemists face to face with their nomenclature problems, and the CNIC report answered at least a few of these problems.

- Some important topics covered by the report were
- order of listing components in binary compounds and mixed salts
 - preference for the Stock systems for compounds of metals in variable oxidation states
 - preference for coordination type names rather than those derived from double salt types: fluosilicate vs. silicofluoride, etc.
 - systematic treatment of salts of iso- and heteropolyacids.

The names for the elements are the base from which the names of most inorganic compounds are constructed. Yet throughout the 19th century and nearly half of the 20th, there were elements for which two names were used to about the same extent. In 1947 CNIC made some hard choices and put itself on record as favoring beryllium over glucinium, niobium over columbium, wolfram over tungsten, and (a bit later) promethium over illinium or florentium. The French and Americans have never accepted wolfram, and American metallurgists continue to use columbium. While the Commission sought to avoid any implication of settling controversies about priority or validity of discovery, this was hardly possible in the case of promethium. More on this aspect of nomenclature later.

Other items covered by the 1957 edition of the Red Book include names for groups of elements, sequence for listing constituents where several metals or cations or nonmetals or anions are present in the same salt, e.g., intermetallic compounds, names for compound ion and radicals, functional derivatives of acids, addition compounds, and considerable extension of the treatment of coordination compounds (this topic was often discussed in the period 1946–1956⁵).

The members of CNIC did not regard the 1957 rules as a job accomplished for they began immediately to plan a revision.

The significance of this first CNIC report far exceeded its substance. Chemists generally were happy to have a standard to which to refer; they followed the rules quite well and extrapolated them for their immediate needs as well as they could in the spirit of an agreed upon standard.

The most outstanding development during the past 40 years is the unmistakable evidence that there is a unity among all chemistry and not a clear separation between organic and inorganic. The significant advances in the border areas between these two major fields only emphasizes the necessity of finding solutions to nomenclature problems which are satisfactory to both groups.

Closely following the end of World War II, the ACS Division of Physical and Inorganic Chemistry established a Committee on Inorganic Nomenclature which became the focal point for the discussion of nomenclature problems. This committee became part of the committee structure of the Division of Inorganic Chemistry when it was established. CNIC was reactivated in 1946 with largely new personnel and in 1957 issued a set of rules (Red Book)⁴ which were significantly broader than the initial ones of 1941. These were given widespread distribution and were generally followed quite well. However, they were by no means all inclusive. They were a good starting point for expansion and simplification. For this revision CNIC had the benefit of a number of publications which appeared in the interim between the two sets of rules.^{2,5} In addition, the Commission began working with its counterpart in organic chemistry to prepare a section of the organic

nomenclature rules treating organometallic compounds, chains and rings with regular patterns of heteroatoms, organic compounds containing phosphorus, arsenic, antimony, or bismuth, organosilicon compounds, and organoboron compounds. This effort was published as Section D of the IUPAC Organic Nomenclature Rules in 1973.⁶

The second edition of the Red Book appeared in 1971. Its significant features included

- acceptance of alphabetic arrangement for cations and/or anions in double salts, etc., and for ligands
- acceptance of Ewens–Bassett numbers (charge on ions) as an alternative to Stock numbers
- use of “-ane” as the termination for all molecular hydrides except for PH_3 , AsH_3 , SbH_3 , and BiH_3 where “-ine” is acceptable
- considerable expansion of the treatment of iso- and heteropolyanions and also coordination compounds. The latter chapter included many more examples, organic ligands (neutral and charged, also compounds in which π bonds serve as coordination sites), designation of active coordination sites, assignment of locants and designation of isomerism, di- and polynuclear compounds (with and without bridging groups), extended structures, homoatomic aggregates, and designation of absolute configurations
- a table (Appendix, Table II) listing, on the basis of composition, names for stoichiometric entities when functioning as cations, anions, radicals, or ligands.

Altogether, the second edition of the Red Book (1971) was a bit more than twice the size of the original edition.

Did you ever start rejoicing over the completion of a major task only to find that one big segment of the total job had been overlooked and still had to be done? That is what happened to CNIC in 1971. The Commission members were congratulating themselves on completing the proofreading of the new edition when someone pointed out that Table II of the Appendix should be expanded. Now Table II was a listing of ions, radicals, and ligands by formula and was a convenient way to quickly find answers to specific questions about nomenclature practices. By 1975 the major disagreements of policy and substance had been resolved and the document was ready for final typing. The resulting publication included helpful advice on how to apply the rules to name a specific compound.⁸

As early as 1952, a group of American chemists began to discuss systematic procedures for nomenclature in the burgeoning field of boron hydrides and their derivatives. However, it was impossible to get agreement among the active investigators on some important details. By the middle 60's there was still no agreement, and CNIC had begun some feeble attempts to deal with these compounds without really understanding the problems. In view of the urgency a small working party in this country formulated rules and collected reactions. There were some battles, hurt feelings, and revisions but the rules are now IUPAC approved.^{9,10}

There are problems not altogether unique to boron but encountered to a maximum degree with its compounds: wide variety of boron–hydrogen compounds which apparently follow no simple stoichiometric relationships, prevalence of three-atom bonds which prevents the equating of a line between nearest neighbors as a simple shared pair bond, existence of many ionic hydrides, replacement of boron by carbon and other atoms, included among the clusters several of high symmetry and many with minimal symmetry. Unfortunately, the field has grown so rapidly that the development of nomenclature has not been able to keep pace with the expansion. There presently are emerging a series of papers carefully examining all aspects of the problems encountered and suggesting general solutions.¹¹ The resulting nomenclature system should stand fast against

Table I. Elements of Atomic Number >105

0 = nil	5 = pent
1 = un	6 = hex
2 = bi	7 = sept
3 = tri	8 = oct
4 = quad	9 = enn
107 = un-nil-sept-ium	Uns
236 = bi-tri-hex-ium	Bth

any likely experimental challenge in the future.

After the initial effort on the part of CNIC to have a single name and symbol for each element, noted above, those who sought new transuranium elements adopted the practice of submitting their names and symbols to the Commission for adoption. The Commission approved the names and symbols with only occasional question or suggestion. When it was evident that the original announcement of nobelium, element 102, was premature, CNIC adopted a policy of waiting for 5 years after the original announcement to insure the validity of the claim. (Only the acceptance of the name nobelium for the element by Ghiorso, after his conclusive evidence for its reality, saved the Commission a good deal of embarrassment.) However, when the announcement of elements 104 and 105 were vigorously challenged and the passage of 5 years did not resolve the challenge, CNIC modified its policy to provide for confirmation of a new element in a separate laboratory before accepting a name. Still the problem would not go away. Even after elements 104 and 105 had been prepared in two separate laboratories, questions remained: "was the evidence supporting the original announcement valid?" and "if not, who should have the privilege of naming the element?". CNIC had no expertise which would permit it to answer this question when the experts disagreed. (IUPAC appointed a neutral group of experts to decide the question of priority of discovery. After 8 years it has not reached a decision and seems to be inactive.¹²

It is never pleasant to be caught in cross fire. Even before 1970, CNIC was discussing possible systematic names for the heavy elements so that each element would have a name even before discovery. In 1969 a decision was made that systematic names of the heavy elements should be based on morphemes representing atomic numbers. Hence, by 1971 a set of names had evolved which was debated extensively in 1973 and actually fully resolved only after the formal meetings of the Commission had adjourned, by circulating a petition. Briefly, a set of 10 short numerical morphemes—each one beginning with a different letter—are used to express the digits of the atomic number; the initial letters of the morphemes constitute the symbol for the element (see Table I). The resulting names were ridiculed mainly because the sounds are strange on initial hearing and many parliamentary devices were employed to delay the adoption of the system. However, full IUPAC approval has been obtained.¹³ There are still rumblings and physicists have rejected the system,¹⁴ but no one has ever challenged the validity of a system incorporating atomic numbers in the names.

Immediately following the end of World War II efforts were begun to develop a satisfactory procedure for naming isotopically modified compounds—both pure compounds having only a single isotope of an element at isotopically modified positions and "labeled" or "tagged" molecules. Sets of rules for both organic and inorganic compounds have been published. The final report of the inorganic rules bears the names of three Americans.¹⁵

The development of a field of research is often accompanied by the need for suitable nomenclature devices. Such has been the case with recent research on the coordination compounds of molecular nitrogen and various hydrides to nitrogen. Special rules to cover these compounds, prepared under the leadership

of the British, have been issued.¹⁶ Occasionally, even though the basic principles involved in a special field exist elsewhere, it is desirable to issue a report to insure that everyone applies these principles in the same manner.

A publication dealing with simple inorganic and coordination polymers will appear in the near future. It is a document which originated in the Nomenclature Committee of the ACS Polymer Division and has been under intense study and debate for nearly 10 years.

A number of other areas needing nomenclature development are under active study. Many could be finalized if those taking leadership were not so busy with details of other urgent nomenclature matters.

Closely related to the nomenclature of polyboron compounds is that of clusters in general. There are certain variations here which are not encountered among boron compounds such as polyhedra with an atom inside the polyhedral surface, nontriangular polyhedral faces, and capped triangular faces of polyhedra. It appears, however, that only slight modifications of the principles developed for polyboron compounds will be required to handle most of these other clusters.

The nomenclature of chains and rings is a basic part of organic nomenclature. In recent years, the number of chains and rings containing atoms other than carbon, with or without oxygen, nitrogen, or sulfur, has increased rapidly. The nomenclature of homogeneous chains and rings or heterogeneous chains and rings with repeating patterns is handled quite nicely by the pattern of nomenclature which evolved for siloxanes and related compounds.¹⁷ However, the number of inorganic chain and ring compounds, including many which do not conform to the criteria given above, has mounted rapidly, and there is a strong desire for a completely general pattern, one that is not based on replacement of atoms in a carbon chain or ring or that is limited by constant valency. Several systems have been proposed and these are being critically studied. German chemists are playing a prominent role here.

Many of the transition elements in their higher states of oxidation form anions containing 2–18 or more atoms of the same or different elements—the so-called iso- and heteropolyanions (acids). We are familiar with polyacids in dealing with the compounds of silicon, phosphorus, and boron. They form chains, rings, and ladder chains as well as both two- and three-dimensional networks. With these, we are dealing primarily with tetrahedra in which one or more peripheral oxygen atoms surrounding a central atom of S or P are shared by two or more tetrahedra. However, with the transition elements, we have octahedra of six oxygen atoms surrounding a central atom of Mo, W, V, Nb, etc., and very frequently the oxygen sharing between two units involves not one oxygen but two or even three and also involves tetrahedral units as well as octahedral ones. The result is a most amazing group of aggregates: very few chains, some simple rings, bridged rings, several planar and three-dimensional aggregates, a large number of super tetrahedral assemblies like the molybdophosphate ion composed of interconnected groups of three octahedral units surrounding a tetrahedral center, and a few assemblies which simply boggle the mind. Yet all of these are built of simple units assembled in regular patterns. Typical examples are shown in Table II. Further, one or more of the central atoms in one of the octahedral units may be in a reduced state or replaced by a heteroatom. The possibilities for isomerism are legion. In many instances, the assemblies of four groups of three edge-sharing octahedra are incomplete in that either one group of three is missing or one octahedron from each of the groups of three are missing. Finally, both the complete and incomplete units may link together directly or through a bridging unit. It remains for chemists to discover a system of nomenclature reflecting the simplicity and natural

Table II. Iso- and Heteropolyanions

$W_6O_{19}^{2-}$	$Mo_6O_{19}^{2-}$	$Nb_6O_{19}^{8-}$
$W_5NbO_{19}^{3-}$	$W_6-nNb_nO_{19}^{(n+2)-}$	$WNb_5O_{19}^{7-}$
	$W_6O_{18}S^{2-}$	
$Mo_8O_{26}^{4-}$		$V_{10}O_{28}^{6-}$
$TeMo_6O_{24}^{3-}$	$IMo_6O_{24}^{5-}$	$H_6CoMo_6O_{24}^{3-}$
$As_2Mo_6O_{26}^{6-}$		$(C_6H_5As)_2Mo_6O_{24}^{4-}$
$P_2Mo_5O_{23}^{6-}$		$S_2Mo_5O_{21}^{4-}$
$SiW_{12}O_{40}^{4-}$	$PMo_{12}O_{40}^{3-}$	$PV_2Mo_{10}O_{40}^{5-}$
	$Al_{13}(H_2O)_{12}(OH)_{24}O_4^{7+}$	
$PW_{11}O_{37}^{7-}$		
$SiW_{10}O_{34}^{10-}$		
$SiMo_9O_{31}^{10-}$	$PMo_9O_{31}H_6^{3-}$	
$P_2W_{18}O_{62}^{6-}$		
$H_4As_4Mo_{12}O_{50}^{4-}$		

Table III. Illustrative Chemical Morphemes

sod-ium chlor-ide
tetr-a-sulf-ur tetr-a-nitr-ide
di-brom-o-meth-ane
di-iso-prop-yl-benz-ene-di-carb-oxyl-ic acid
hex-a-chlor-o-ferr-ate-(3-)
carb-ox-imid-ic acid

order underlying the apparent complexity. The French have made real progress in doing just this. [Anyone desiring to gain some idea of the types of structures involved should consult ref 18.]

The greatest need in chemical nomenclature is simplification of the rules: general statements instead of repetition in a half dozen separate areas where the generality is applied, omission rather than detailed presentation of exceptions to rules, and similar situations. Then there need be fewer occasions to consult 100–400 pages to find the answer to a simple question. There is need for devices to enable the searcher to more quickly locate the desired rule. Then finally, there should also be a short version of the rules—only a few pages. Obviously many details would have to be omitted but students and many others with only casual need to consult the details would learn the general principles much more readily. Such an abbreviated version of the inorganic rules is in preparation.

Little has been said about the construction of chemical names. Just like all words, chemical names consist of units of a few letters each of which has meaning. These are known as morphemes. They comprise stems, prefixes, and suffixes. Nomenclature rules are nothing but recipes for putting these morphemes together (in many instances with locants and structural modifiers) to indicate structure. A suffix or a prefix or even a stem may contain several morphemes. One can find tables of prefixes and suffixes but not of morphemes. One is being compiled (see Table III). It reveals some interesting things. In many instances, the same morpheme has more than one meaning, depending upon where and how it is used. In other instances, perhaps equally numerous, two or more different morphemes are used to convey the same meaning. Nomenclature reform ultimately must begin by simplifying this multiple usage and duplication. The great interest in computer retrieval by means of nomenclature may be just the drawing force to speed up this process.

Finally, just a word about acronyms which are in such common use everywhere today. Attempts to control this widespread habit in chemistry so far have not met with much success. The Interdivisional Committee on Nomenclature and Symbols¹⁹ has urged that once in every article, lecture, or book, the author define all of the acronyms and abbreviations that he is using. But should this apply to DNA, or en, or XPS? Does not everyone know what they mean? Elementary as it

may seem, everyone does not know what they mean. Personally, I do not think pleas, threats, or legislation will do the trick. We are going to have to devise some general rules about making acronyms and abbreviated names.

When all of the topics mentioned so far have been disposed of, will those interested in nomenclature have anything to occupy their time? I am sure they have no need to worry about technological unemployment. Problems of nomenclature in solid-state chemistry await study. To be sure, the Red Book has a chapter on non-Daltonian or Berthodide (nonstoichiometric) compounds. Not only is extension of this required but many workers in the field disagree with the fundamental basis of the recommended practice. Reactions at the surface of solids is extremely important today for a variety of reasons: catalysis, protective coatings, mineral flotation, etc. There is no nomenclature whatever to designate surface compounds. Then there are many compounds for which no single formula can be written. While nomenclature seems to have surmounted resonance and electron delocalization, there remain tautomers and fluxional compounds. With NMR techniques, it is relatively easy to detect fluxional character and even freeze the equilibrium between two or more forms. The question is how to designate this in names. Fortunately, this is another place where cooperation between CNOC and CNIC will be necessary. Then there is always the chance that someone will come across a completely new and unexpected type of combination. Although the chance of this happening has been steadily decreasing during the present century, the possibility of its happening again cannot be ruled out.

So far we have confined our attention to names for single species. There is another chemical nomenclature to designate types of compound, general structural features, types of behavior and reactions, etc. As yet little attention has been devoted to these aspects of our system of communication.

You may have guessed that I regard chemical nomenclature development as important and worthy of the time required to solve the problems and gain consumer approval. A few words about future prospects for continued progress. The interest exhibited by chemists generally is heartening. One only needs to attend a meeting of the ACS Nomenclature Committee to realize how real and active the interest is. The existence of national committees in Germany, France, and Britain gives evidence of similar interest elsewhere. However, it is no idle boast to note that the key to nomenclature development and dissemination of information rests with the ACS through its Office of Nomenclature and especially Chemical Abstracts Service. This role becomes all the more important with the diminishing role of IUPAC because of reduced finances. In these days of cost cutting and emphasis on short-range rather than long-term goals, I hope very much that the ACS can continue to play its leading role in nomenclature matters. This is one aspect of world leadership of which all American chemists can be very proud.

REFERENCES AND NOTES

- (a) Guyton de Morveau, L. B. *J. Phys.* **1782**, 19, 310, 382. *Ann. Chim. Phys.* **1778**, 25 (1), 205. (b) Guyton de Morveau, L. B.; Lavoisier, A. L.; Berthollet, C. L.; de Fourcroy, A. F. "Méthode de Nomenclature Chimique"; Cuchet: Paris, 1787.
- Symposium on Inorganic Chemical Nomenclature, 1941. *Chem. Rev.* **1943**, 32, 73–97, 99–108.
- Jorissen, W. P.; Bassett, H.; Damiens, A.; Fichter, F.; Rémy, H. *J. Am. Chem. Soc.* **1941**, 63, 889–897.
- IUPAC Commission on the Nomenclature of Inorganic Chemistry, "Nomenclature of Inorganic Chemistry"; Butterworths: London **1959**. *J. Am. Chem. Soc.* **1960**, 82, 5523–5544.
- (a) Fernelius, W. C. *Chem. Eng. News* **1948**, 26, 161–163. (b) Fernelius, W. C.; Larsen, L. M.; Marchi, L. E.; Rollinson, C. L. *Ibid.* **1948**, 26, 520–523. (c) Ewens, R. V. G.; Bassett, H. *Chem. Ind.* **1949**, 131–139. (d) Patterson, A. M., Ed. "Chemical Nomenclature". *Adv. Chem. Ser.* **1953**, No. 8.
- IUPAC Commission on the Nomenclature of Organic Chemistry, "Nomenclature of Organic Chemistry, Sec. D"; Information

- Bulletin—Appendices on Tentative Nomenclature, Symbols, Units, and Standards, No. 31, IUPAC Secretariat, Oxford, 1973.
- (7) IUPAC Commission on the Nomenclature of Inorganic Chemistry, "Nomenclature of Inorganic Chemistry", 2nd ed.; Butterworths: London, 1971, *Pure Appl. Chem.* 1971, 28, 1-110.
 - (8) IUPAC Commission on the Nomenclature of Inorganic Chemistry, "How to Name an Inorganic Substance," Pergamon Press: Oxford, 1977.
 - (9) American Chemical Society, "Nomenclature of Boron Compounds". *Inorg. Chem.* 1968, 7, 1945-1964.
 - (10) IUPAC Commission on the Nomenclature of Inorganic Chemistry, "Nomenclature of Inorganic Boron Compounds". *Pure Appl. Chem.* 1972, 30, 683-710.
 - (11) (a) Casey, J. B.; Evans, W. J.; Powell, W. H. *Inorg. Chem.* 1981, 20 (5), 1333-1341. (b) *Ibid.* 1981, 20, 3556-3561.
 - (12) IUPAC Executive Committee Minutes, April 2-3, 1981.
 - (13) IUPAC Commission on the Nomenclature of Inorganic Chemistry, "Naming of Elements of Atomic Numbers Greater than 105"; IUPAC Secretariat, Oxford, 1976.
 - (14) Barber, R. C. Secretary of IUPAC Commission on Symbols, Units, Nomenclature, Atomic Masses, and Fundamental Constants. Letter to M. A. Paul, Secretary IDCNS, reporting recent action of SUN-AMCO, Sept 15, 1979.
 - (15) IUPAC Commission on the Nomenclature of Inorganic Chemistry, "Isotopically Modified Compounds". *Pure Appl. Chem.* 1979, 51, 1981-1994.
 - (16) IUPAC Commission on the Nomenclature of Inorganic Chemistry, "Nomenclature of Hydrides of Nitrogen and Derived Cations, Anions and Ligands". *IUPAC Inf. Bull.* 1978 (2), 151-60.
 - (17) ACS Committee on Nomenclature, Spelling and Pronunciation, *Chem. Eng. News.* 1952, 30, 4517.
 - (18) (a) Kepert, D. L. *Prog. Inorg. Chem.* 1962, 4, 199-274. (b) Evans, H. T., Jr. *Perspect. Struct. Chem.* 1971, 4, 1-59. (c) Weakley, T. J. R. *Struct. Bonding (Berlin)* 1974, 18, 131-176. (d) Tytko, H. K.; Glemser, O. *Adv. Inorg. Chem. Radiochem.* 1976, 19, 239-315.
 - (19) IUPAC Interdivisional Committee on Nomenclature and Symbols, Appendices on Provisional Nomenclature, Symbols, Terminology, and Conventions, No. 58, IUPAC Secretariat, Oxford, 1977.

Structure Generation on the Basis of BCT Representation of Chemical Structures

TAKASHI NAKAYAMA and YUZURU FUJIWARA*

Institute of Information Sciences and Electronics, University of Tsukuba, Sakura-mura, Niihari-gun, Ibaraki 305, Japan

Received April 19, 1981

A method of structure generation based on BCT (block-cutpoint tree) representation of chemical structures has been developed. The generation program is a part of the automatic structure analysis system of mass spectra (ASASMAS) and is used when a set of the inferred substructures are given as input data. The input substructures are represented by means of BCT.

INTRODUCTION

The major steps in the process of chemical structure analysis are inference of constituent substructures and structure generation by combining those inferred substructures. A method of structure generation from substructures already inferred is described in this paper. Various schemes for structure generation have been devised;¹⁻¹⁰ each of these methods is based on a method of representation of chemical structures, and it may be said that the method of representation of chemical structures determines the method of structure generation. All of these methods including the present one pursue the governing principle of reducing the number of combinations. In this paper, the chemical structure is represented in terms of BCT,¹¹ which clarifies the hierarchy in chemical structures completely so that the idea of the connectivity stack³ and the superatom⁵ are included naturally in our method. Combinatorial problems that occur in the process of structure generation are thus partitioned into subproblems and are classified into stages.

Structure generation is regarded as a problem in combinatorial analysis. There must be neither duplication nor omission in the generated structures, and the method of structure generation should be efficient; processing time should be short in practice. The processing time depends largely on the method of representation of chemical structures. The constituent unit of the BCT representation of chemical structures is a block (a biconnected component of a graph), and this makes it possible to reduce the number of combinations greatly by omitting atom-by-atom processing. Another feature of the method of structure generation described here is the availability of a graph data base. The various graphs which appear in the process of structure generation are not generated for each case

but instead are retrieved from the graph data base.

REPRESENTATION OF CHEMICAL STRUCTURES

Chemical structures are represented by means of graphs, and atoms and bonds of a chemical structure correspond to vertexes and edges of a graph, respectively. However, it is inefficient to process chemical structures at the atomic level, and sometimes a superatom/ring assembly is used as the processing unit. In ASASMAS, the concept of the superatom is extended to represent the chemical structures in a way that is consistent with BCT. The following is a brief description of BCT representation of chemical structures.¹²

Let $u \in V$ be a vertex of a connected graph $G = (V, E)$. A vertex u is a "cutpoint" if the removal of u yields the disconnectedness of the graph G . A "block" of a graph G is a maximal subgraph of G which contains no cutpoints. Now, $bc(G)$ (block-cutpoint graph of G) is defined: $T = (W, F)$ is a $bc(G)$ if (1) $W = A \cup B$ is a set of vertexes where $A = \{a_1, \dots, a_n\}$ is a set of all cutpoints of G , and $B = \{b_1, \dots, b_m\}$ is a set of all blocks of G and (2) $F = \{f_1, \dots, f_l\} = \{(a_i, b_j) | a_i \in B_j, a_i \in A, b_j \in B\}$, where $a_i \in b_j$ means that a cutpoint a_i is a member of a set of vertexes of block b_j . A $bc(G)$ has the following properties: (1) it is a bipartite graph of subsets A and B , (2) it is a tree regardless of the original structure, (3) terminal vertexes correspond to blocks of G , and (4) the distance between any pair of terminal vertexes is an even number. A $bc(G)$ is called a block-cutpoint tree because of property 2. A tree is a BCT if and only if it possesses property 4, and this property is used for generating BCT. Examples of the BCT representation of chemical structures are shown in Figure 1. The internal structure of each block is filed in the block dictionary.