Specific polymers of regular structure obtained from these monomers can be named by a structure-based nomenclature system, but most of these products are known only from the processes by which they were made and by an analysis of their compositions. Thus, a code name based on percentage composition might be "A(30)B(50)-S(20) Resin." Additional descriptors can be used to give process details. For example,

S(20)-A(30)-g-polyB[A(10)B(70)S(20)]

would represent the grafting of a mixture of 20 parts of styrene and 30 parts of acrylonitrile to poly(butadiene) to give an ABS Resin, 70% of which was derived from butadiene, 10% from acrylonitrile, and 20% from styrene. A mixture of two polymers might be labeled "Poly[A(30)-S(20)]-mix-polyB(70)." Such codes are simply a matter of establishing and learning a group of conventions. They are a form of shorthand and as such are more convenient than the paragraph of description that may be required for an individual family member. These codes are useful but they are in no sense nomenclature, and they will never replace language.

SUMMARY

Our language can handle the real world of organic polymers on a rational structural basis if polymer chemists are willing to make the same assumptions they have long accepted as organic chemists. This will allow the kind of flexibility needed for both today and the distant tomorrow, including the requirements of both chemist and computer. Such acceptance also forces the polymer chemist to break with some traditional language habits with which he is very much at home.

At the same time, it must be recognized that no rational language system can handle the irrational. Polymers having irregular structures must be labeled on the basis of information available. The label may be either a modified trivial name, which communicates little, or it can be an artificial code. In the latter case, word confusion is reduced, and if simple and transmittable terms are used, communication can indeed take place.

A New Notation System for Indexing Polymers*†

HERMAN SKOLNIK and J. T. HAYS
Hercules Incorporated, Research Center, Wilmington, Del. 19899

Received April 3, 1970

A new linear notation system, based on carbon, its bonding, and the number of hydrogens attached to it, is evaluated as an indexing mechanism for polymers from the viewpoint of the structural repeating unit for linear polymers and reactants for highly cross-linked polymers or resins. The system yields computer printouts in which structural classes, such as allyl, vinyl, and vinylidene, and amide, urethane, and ester polymers, are grouped together.

Polymer nomenclature problems caused by the increase in number and complexity of polymers make effective communication difficult.[1] As serious as these nomenclature problems are in the communication aspects, they are considerably more intractable in the indexing aspects. In response to these problems, we have considered the feasibility of the various notation systems described in the literature[2-6] for a possible solution.

Although most of the notation systems described yielded unambiguous and unique designations, each required major modifications for computer processing into a reasonably ordered arrangement. In summary, we found the established notation systems to be discordant with the needs of polymer chemists for bringing like things together and to be incompatible with computer processing.

Recently a new notation system was introduced which is based on carbon, carbon bonding, and hydrogens.[7,8] This notation system yields specific and unambiguous designations with the additional advantage of being in harmony with computer processing into chemically related arrangements. The application of this system to polymer structural indexing via computer is illustrated in this paper.

BASIS FOR POLYMER INDEXING

Polymers can be indexed on the basis of the monomers from which they are formed or on the basis of the actual structure of the repeating unit. We have chosen to use the structural repeating unit for all linear polymers: first, because it avoids ambiguities of one monomer forming more than one polymer or of more than one monomer forming the same polymer; and, second, to assure like polymer structures being grouped together. This is in agreement with the structure-based polymer nomenclature system presented in a report by the Nomenclature Committee of the Division of Polymer Chemistry of the American Chemical Society[1] and with the convincing arguments

* Presented before the Middle Atlantic Regional Meeting, ACS, Newark, Delaware, April 3, 1970.

† Hercules Research Center Contribution No. 1511.

presented by Livingston and Fox.[9] *Chemical Abstracts* has indexed all polymers on the basis of their monomers, but beginning with index entries in Volume 66 (January–June, 1967) has also provided structure-based entries,[10] following the rules presented in the Committee report.[1] However, our notation system provides the added advantage of an inherent specificity within a generic structural classification; for example, all vinyl polymers (substituted ethylene repeating units) fall under the general notation, ¬CY(~~~)¬, and are listed without interruption in the order of the substituents indicated in the parentheses. Other generic classifications are similarly grouped. The alphabetical nomenclature does not permit this type of grouping in the structure-based system used by *Chemical Abstracts*.

THE NOTATION SYSTEM

The symbols used in the new notation system[7,8] are listed in Table I.

The steps in representing a structure in the notation system are:

> Write the structural formula in the usual fashion—i.e., from left to right with functional groups assigned to the right end position in acyclics or to the top or right position in rings.
>
> Starting at the right end, or at atom number one, write the indicated notation symbol for each atom or combination of atoms. Show each substituent or attachment to the main chain immediately after the symbol for the atom to which it is attached and separate the substituent group by parentheses.

These steps are shown in Figures 1 and 2.

GROUPING OF POLYMERS

Grouping of all polymers together is accomplished simply by enclosing the structural repeating unit (SRU) between the symbols indicated:

¬SRU¬

The symbol, ¬, the "logical not," placed at each end of the letter notation for the SRU, is in the group of special characters which are sorted before letters by IBM 360 computers. All polymers thus fall between cyclics (indicated by a period) and acyclics (notation beginning with a letter); all polymers designated in this manner by the SRU are printed out together. Copolymers are also designated by the appropriate SRU with the symbol, @ (the "at" sign), between portions of the SRU due to each monomer. Copolymers of the monomer written first will print out following homopolymers of that monomer.

This system is applicable to both addition and condensation polymers which can be specifically designated by the SRU. The SRU is the preferred basis for notations

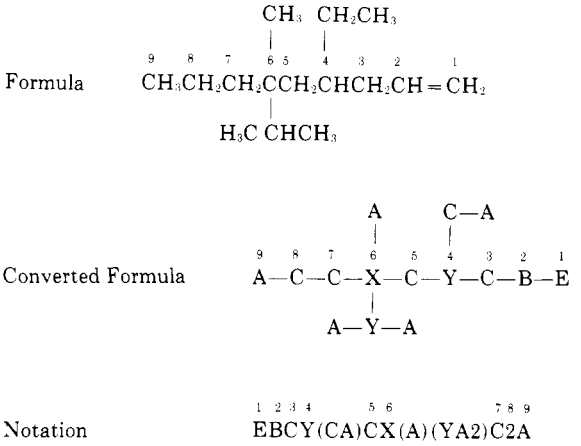


Figure 1. Assignment of notation to 4-ethyl-6-isopropyl-6-methylnonene-1

Table I. Notation Symbols

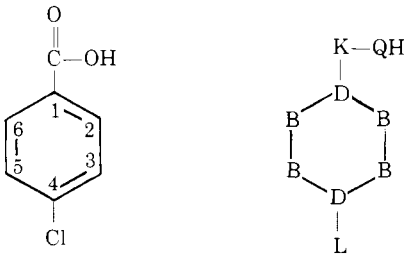
Single-Bonded Carbons			Double-Bonded Carbons		Triple-Bonded Carbons	
—CH3		A	=CH2	E	≡CH	U
—CH2—		C	=CH—	B	≡C—	V
—CH—		Y	>C=	D		
—C—		X	>C= (fused)	R		
>CH—	(fused or bridgehead)	J				
>C<	(fused or bridgehead)	T				

Carbonyl		Halogen		Nitrogen		Oxygen		Other	
>C=O	K	—F	F	>NH	M	—O—	Q	—H	H
		—Br	G	—NH2	MH	—OH	QH	—S—	S
		—I	I	>N—	N	O2	W	=S	
		—Cl	L	≡N or —N=	Z	(NO2 or SO2)		P	P

Character Symbols

* fused or bridgehead atom (other than C)
# ionic form
: substituent between bridgeheads
& Atomic symbol—e.g., & NA for sodium, denotes elements other than already listed

□ condensed ring structure
.~~~. cyclic structure
¬~~~ polymer repeating unit
'~~~' highly cross-linked polymer or resin—e.g., phenol-formaldehyde
@ separates comonomers or reactants



Formula | Converted Formula

Notation: .D(KQH)B2D(L)B2.

Figure 2. Assignment of notation to *p*-chlorobenzoic acid

for linear polymers. Designation of the reactants used offers a simpler basis, however, for notations for highly cross-linked polymers or for resins. In these cases we use the apostrophe, ', to enclose the notations for the reactants separated by the @ sign:

'~~@~~'

Designation of reactants also is preferred in those situations where the nature of the polymeric product is unknown or too complex to be indicated clearly. Thus we have provision for designation of polymers both by the SRU and by the reactant or source materials. The second class will print out after the first on the basis of the order of computer sorting of the enclosing symbols. Application of the notation system to various classes of polymers is illustrated in the following sections.

ADDITION POLYMERS

An essential and distinguishing requirement for application of the notation system to addition polymers is that the repeating unit be written to show the polymerizing radical as the main chain and to show (in parentheses) other portions of the molecule as substituents or pendent groups. Thus the repeating unit for a polymer from 1-butene is written as

$$\left(\begin{array}{l} -CH-CH_2- \\ \quad | \\ \quad CH_2CH_3 \end{array} \right)$$

or by notation as ¬CY(CA)¬.

Table II. Olefin Polymers

Type	Notation
Polyethylene	¬C¬
Vinylidene polymers	¬CX(~~)(~~)¬
Polystyrene	¬CY(.DB5.)¬
Polypropylene	¬CY(A)¬
Allyl polymers	¬CY(C~~)¬
Other vinyl polymers	¬CY(~~)¬
Poly(tetra-substituted ethylenes)	¬X(~~)(~~)X(~~)(~~)¬
Poly(tri-substituted ethylenes)	¬Y(~~)X(~~)(~~)¬
Poly(di-substituted sym. ethylenes)	¬Y(~~)Y(~~)¬

The exact structure of the repeating unit of polymers and copolymers generally is not known or not rigorously established. In writing the notation, the most reasonable type of combination is assumed, such as head-to-tail in vinyl-type polymers.

The specific notation depends on the order of the atoms used in the repeating unit. For olefin polymers, the carbon with the most hydrogens is written as the terminal atom, giving the repeating units:

$$-CH_2-CH_2-,\ -\underset{\underset{H}{|}}{C}-CH_2-,$$

$$-\underset{|}{\overset{|}{C}}-CH_2,\ -\overset{|}{\underset{|}{CH}}-\overset{|}{\underset{|}{CH}}-,\ -\underset{|}{\overset{|}{C}}-\overset{|}{\underset{|}{CH}}-$$

with the corresponding notations: ¬C¬, CY(~~), CX(~~)(~~), Y(~~)Y(~~), and YX(~~)(~~). This deviation from the general procedure[4] of placing atom number one on the right, and hence first in the notation, is required to give the notation printout sequence shown in Table II.

References on the broad classes indicated are thus grouped together in a logical progression of closely related polymer structures; the ¬C¬, ¬CX(~~)(~~)¬, and ¬CY(~~)¬ categories cover the important types of polymers from terminal olefinic monomers.

Repeating units from cyclic monomers containing oxygen, nitrogen, or sulfur are written to show these atoms as terminal, thus making the hetero atom the basis for the indexing sequence in Table III.

$-CH_2CH_2NH-,\ -CH_2O-,\ -CH_2CH_2O-,$

$> CHCH_2O-,\ -CH_2CH_2S-$

In Table III, use of the symbols for nitrogen, oxygen, or sulfur atoms gives notations which separate these polymers from most of the olefin polymers. These polymers are inserted into the listing of carbon chain polymers between those with repeating units with terminal $-CH_2-$ and higher substituted units. This is not considered serious inasmuch as there are so few polymers of tri- and tetra-substituted ethylenes.

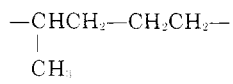
ADDITION COPOLYMERS

It is essential that a polymer notation system be applicable to copolymers. The actual structure of a copolymer, however, is almost never completely established. Nevertheless, a great deal of definite and practical information is available on many copolymers. On others, little more is known other than that the monomers underwent copolymerization. Consequently, there is a qualitative judgment element inherent in the notations for copolymers.

Table III. Poly(Ethers, Imines, and Sulfides)

Type	Notation
Polyethyleneimine	¬MC2¬
Polyformaldehyde	¬QC¬
Poly(ethylene oxide)	¬QC2¬
Poly(propylene oxide)	¬QCY(A)¬
Polyepichlorohydrin	¬QCY(CL)¬
Poly(ethylene sulfide)	¬SC2¬

Notations for copolymers are assigned as follows: Write the structural formula of the repeating unit, with the predominant monomeric unit on the right. For example, an ethylene-propylene copolymer, in which the ethylene is >50%, is considered to be structurally



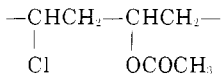
The notation for this copolymer would then be

¬C2@CY(A)¬

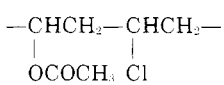
with a cross reference from the minor monomer to the major monomer entry, such as

¬CY(A)@C2¬ See ¬C2@CY(A)¬ for copolymers >50% C2

When the monomers participate essentially equally in a copolymer repeating unit, the copolymer should be indexed under each of the monomeric units. For example, a vinyl chloride-vinyl acetate copolymer, in which each of the monomeric units might be present in equal or nearly equal proportions, is considered to be structurally



or



This copolymer would then be assigned two notations:

¬CY(QKA)@CY(L)¬

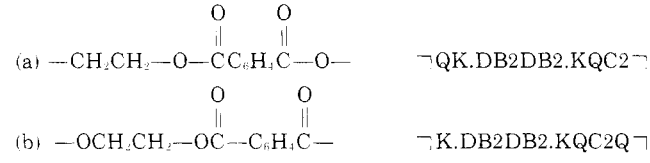
and

¬CY(L)@CY(QKA)¬

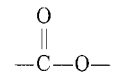
The information needs of the users of the notation index may be such as to require that there be as many notation index entries as there are monomeric units in the copolymer. In this event, it would be valuable to include a composition factor with the notation, or, at least, to specify the composition or relative composition in the abstract for the index entry or entries.

CONDENSATION POLYMERS

The notations for condensation polymers are included with other classes of polymers represented by the SRU. The notation depends on the order in which the structure of the repeating unit is written. The polyester unit for poly(ethylene terephthalate) can be written in various ways which would give different notations; for example:



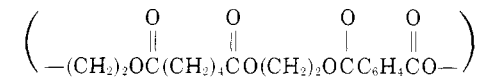
Because it is the structure of the reaction product which is important, (a) must be selected. The QK notation indicates the



grouping characteristic of an ester. Thus, the terminal grouping in (a) permits the polymer to be indexed as an ester; in (b) and other alternatives, it is buried within the notation. Writing repeating units in this fashion for the main classes of condensation polymers, we get the relationships shown in Table IV. These classes of polymers are distinguished by the indexing unit in the computer printout order shown.

The preceding discussion covers linear condensation polymers between two reactants to form a single type of linkage. A more detailed procedure is required for situations involving: additional reactants of the same type—e.g., more than one acid or more than one glycol in a polyester; or more than one type of linkage in a polymer—e.g., both an ester and an amide.

To assign notations for these examples, write the simplest repeating unit which includes all reactants; then write the notation in as many ways as required to show the different reactants and different types of linkage. For example, a polyester from a mixture of terephthalic and adipic acids with ethylene glycol will have the repeating unit:



and the notation:

¬QK.DB2DB2.KQC2QKC4KQC2¬

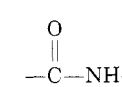
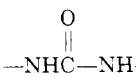
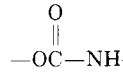
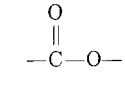
This notation will be listed with polyesters (shown by QK) containing terephthalic acid (QK.DB2DB2.KQ). To enable this product to be listed from the viewpoint of a polyester of adipic acid, the notation accordingly would be rearranged to:

¬QKC4KQC2QK.DB2DB2.KQC2¬

All structural features of the SRU can be emphasized in this way by putting each first in the notation and writing the remainder of the notation in the same relative sequence.

A mixed polyester-polyamide using a diamine as well as a glycol is illustrated in the following repeating unit:

Table IV. Condensation Polymer Groupings

Polymer	Unit Terminal Group	Indexing Notation
Polyamide		MK~~~
Polyurea		MKM~~~
Polyurethane		MKQ~~~
Polyester		QK~~~

The notation would be written to list it with polyesters

¬QK.DB2DB2.KMC2MK.DB2DB2.KQC2¬

and rewritten to list it with polyamides

¬MK.DB2DB2.KQC2QK.DB2DB2.KMC2¬

The advantages in grouping like structures together and the avoidance of arbitrary rules of sequence justify using more than one sequence for a single notation. Indexing flexibility is essential if the needs of users are to be satisfied.

HIGHLY CROSS-LINKED POLYMERS

As indicated earlier, highly cross-linked polymers or resins or other complex polymeric products are preferably designated by reactants rather than by SRU. Phenol-formaldehyde and urea-formaldehyde resins are typical; they cannot be accurately described by a simple repeating unit, because of the variety of structures possible and the dependency of structural variations on reaction conditions. This is true for all resins which can be "cured" or "thermoset." Consequently, we use the notation for each reactant, separated by the sign @, and enclose the combined notations in apostrophes '~~~@~~~'. Thus, a phenol-formaldehyde resin is assigned the notation

'.D(QH)B5.@KH2'

As in the case of copolymers, a judgment factor needs to be resolved in terms of the information needs of the users as to whether the second component should be crossed to the more important component or double entered.

A urea-formaldehyde resin is assigned the notation

'K(MH)2@KH2'

OTHER STRUCTURAL FEATURES

Inasmuch as a polymer is not uniquely characterized by designation of a structural repeating unit, there is a strong temptation to include additional information in the notation. Thus, mole ratio for copolymers and cross-linked polymers, type of combination for copolymers, and molecular weight and molecular weight distribution could be indicated by number and letter combinations. In view of the fact that such information will not be available for all polymers, and certainly not in all references, we concluded that additional terms in the notation would interfere with proper indexing.

Specifically, references on a polymer with no information indexed in addition to the basic notation would be separated from references on the same polymer with such information. For the present, we prefer to rely on the Multiterm concept[11] in conjunction with a suitable abstract system to supply information for polymer characterization and to reserve the notation system for indexing in terms of the structural repeating unit or of the reactants used.

CONCLUSIONS

Our notation system has shown itself readily adaptable to polymer systems. Its outstanding features are its simplicity in application and its capacity to group related polymer structures together in a computer printout. The system can be extended to more complex structures than discussed and such applications will be the subject of further studies.

LITERATURE CITED

(1) "A Structure-Based Nomenclature for Linear Polymers," Committee on Nomenclature, ACS, Division of Polymer Chemistry, *Macromolecules* 1, 193 (1968).

(2) Dyson, G. M., "A New Notation and Enumeration System for Organic Compounds," Longmans, Green and Co., London and New York, 1947, 2nd ed., 1949.

(3) Wiswesser, W. J., "A Line-Formula Chemical Notation," T. Y. Crowell Co., New York, N. Y., 1954.

(4) International Union of Pure and Applied Chemistry, "Rules for I.U.P.A.C. Notation for Organic Compounds," issued by the Commission on Codification, Ciphering, and Punched Card Techniques of the IUPAC, Longmans, Green and Co., London, 1958.

(5) Smith, E. C., "The Wiswesser Line-Formula Chemical Notation," McGraw-Hill, New York, N. Y., 1968.

(6) "Survey of Chemical Notation Systems," a Report of the Committee on Modern Methods of Handling Chemical Information, National Academy of Sciences, Washington, D. C., 1964; also see papers under "Notation" in J. CHEM. Doc. since 1964.

(7) Skolnik, H., "A New Linear Notation System Based on Combinations of Carbon and Hydrogen," *J. Heterocyclic Chem.* **6**, 689 (1969).

(8) Skolnik, H., "A Correlative Notation System for NMR Data," J. CHEM. DOC. **10**, 216–20 (1970).

(9) Livingston, H. K., and R. B. Fox, "Nomenclature of Organic Polymers," J. CHEM. DOC. **9**, 232–4 (1969).

(10) Loening, K. L., W. Metanomski, and W. N. Powell, "The Indexing of Polymers in Chemical Abstracts," J. CHEM. DOC. **9**, 248–51 (1969).

(11) Skolnik, H., "The Multiterm Index: A New Concept in Information Storage and Retrieval," J. CHEM. DOC. **10**, 81–4 (1970).