purchases, approvals, or rejections have been an invaluable aid to the purchasing and manufacturing departments. In addition, the relationship between tablet hardness and disintegration rate has been statistically reviewed using results from samples extending over several years' production.

With the large amount of both chemical and physical data recorded in retrievable card form, other similar comparisons and evaluations of physical and chemical properties, production analyses, and product reviews are being considered and are expected to provide much useful information in the future.

ACKNOWLEDGMENT

The authors wish to acknowledge the assistance and cooperation of members of the data processing staff for the development of the computer programs.

# Indexing Polymers by Formula

PHILIP J. LAUNER
Silicone Products Department, General Electric Company, Waterford, N.Y. 12188

Received August 9, 1967

**Information about polymers is indexed for retrieval by using the formulas of repeating units in polymers. A set of rules is proposed for indexing polymers and copolymers by formula. Examples illustrate each rule. The rules are consistent with the Wyandotte-ASTM codes for indexing spectral data on punched cards. The proposed system can handle organosilicon polymers, which may have fractional atoms in the repeating units.**

Our spectroscopy laboratory must be able to locate reference spectra quickly. It is easy to find spectra of *compounds;* nearly everyone locates them by molecular formula written according to the rules of Chemical Abstracts Service (CAS). It has not been so easy to find *polymers* this way. Homopolymers can be indexed under the formula of the simplest repeating unit. But it is not clear how copolymers are to be indexed, or how to handle organosilicon polymers which often have fractional atoms in the repeating units. This paper describes the system we worked out for indexing polymers by formulas.

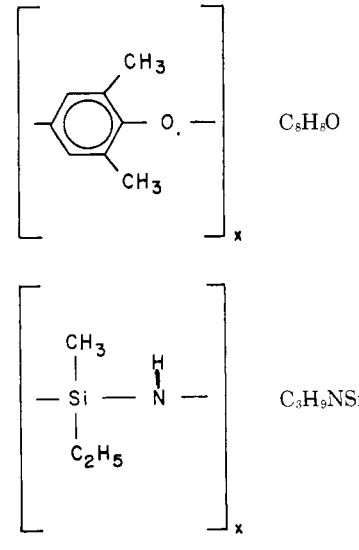
Three major convictions guided the development of our polymer-indexing system:

> The system should use the same order of symbols in formulas as does CAS (*1*).
>
> The system should be a logical extension of the American Society for Testing and Materials' Wyandotte-ASTM rules for indexing spectral data on punched cards (*2, 3*).
>
> The system should be compatible with the shorthand designation for siloxane units in silicone polymers (*4, 5*).

Starting from these principles, we developed a systematic set of rules for indexing polymers by formula. These rules are illustrated in the following examples:

**Homopolymers.** Homopolymers (polymers containing a single type of unit) are indexed under the formula of the smallest repeating unit.

Examples:

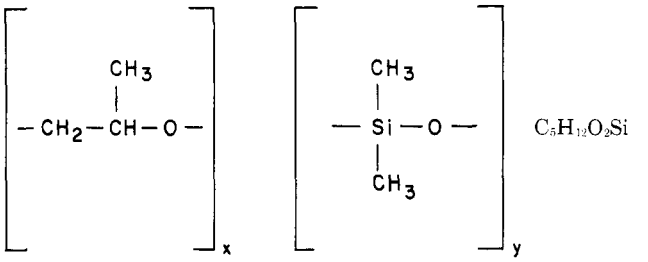


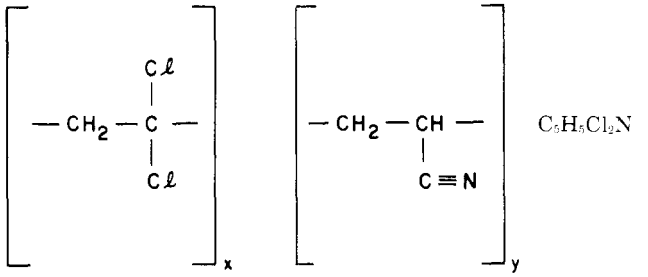

The arrangement of symbols in formulas follows CAS rules (*1*).

**Copolymers Having Two Kinds of Repeating Units.** Copolymers are considered to have a 1:1 molar ratio of the two kinds of units, regardless of the actual ratio in the copolymer. In the examples below, each copolymer

is indexed as the sum of the formulas of the two kinds of repeating units. This coincides with the Wyandotte-ASTM practice (*2*).

Examples:

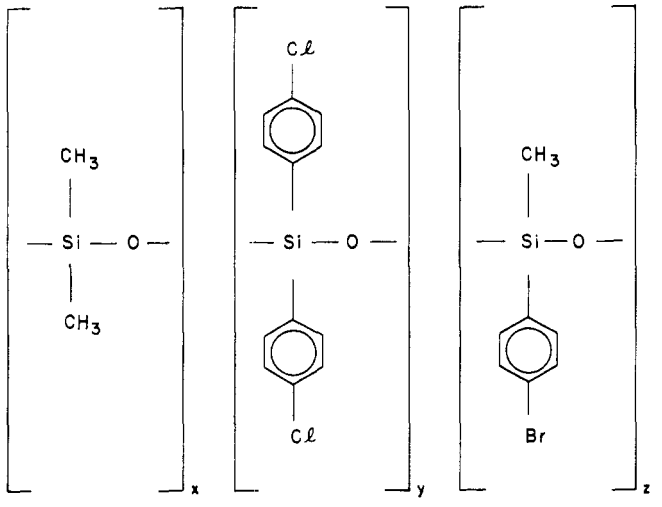



**Copolymers Having Three or More Kinds of Repeating Units.** It is a logical extension of the Wyandotte-ASTM practice to index more complex copolymers by considering them to have a 1:1:1... molar ratio of units, regardless of the number of different kinds of units or the actual ratios in the copolymer.

Example:

A terpolymer containing the units



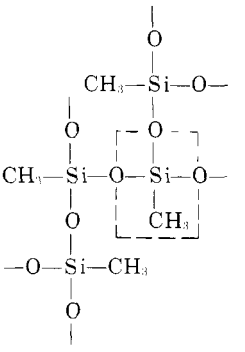

has the formula $C_{21}H_{21}BrCl_2O_3Si_3$, obtained by adding the formulas of the three units:

| | | | | | |
|---|---|---|---|---|---|
| $C_2$ | $H_6$ | | | O | Si |
| $C_{12}$ | $H_8$ | | $Cl_2$ | O | Si |
| $C_7$ | $H_7$ | Br | | O | Si |
| $C_{21}$ | $H_{21}$ | Br | $Cl_2$ | $O_3$ | $Si_3$ |

**Copolymers Having Fractional Atoms in the Repeating Units.** Organosilicon polymers often have fractional atoms in the repeating units. When fractional atoms are present in a repeating unit, we multiply the numbers of atoms by an integer to get the smallest possible whole numbers before counting atoms to determine the formula of the polymer.

Example:

Hydrolysis of $CH_3SiCl_3$ yields a high-molecular-weight white solid having the structure

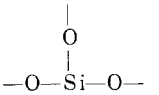
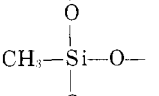
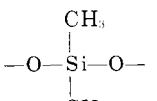
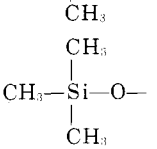



In this structure each oxygen atom is attached to two silicon atoms and is visualized as belonging half to one silicon atom and half to the other; consequently the repeating unit (dashed lines) is $CH_3SiO_{\frac{3}{2}}$. This polymer would be indexed as

$$CH_3SiO_{1.5} \times 2 = C_2H_6O_3Si_2$$

A special shorthand facilitates writing the various kinds of repeating units in siloxane polymers (*4, 5*). This shorthand, which has been adopted by everyone working in silicone chemistry, makes it easy to write complex silicone structures without doing violence to the valences of the atoms. The relation of this shorthand to the units used to determine polymer formulas is shown in Table I.

Table I. Units in Siloxane Polymers

| Structure | Shorthand: Abbreviation | Shorthand: Repeating unit | Unit Used to Determine Polymer Formula |
|---|---|---|---|
|  | Q | $SiO_2$ | $O_2Si$ |
|  | T | $CH_3SiO_{1.5}$ | $C_2H_6O_3Si_2$ |
|  | D | $(CH_3)_2SiO$ | $C_2H_6OSi$ |
|  | M | $(CH_3)_3SiO_{0.5}$ | $C_6H_{18}OSi_2$ |

The silicone shorthand also is used for siloxane units containing organic groups other than methyl by using primes (M′, D′, T′). The organic group must be specified whenever primes are used.

Example:

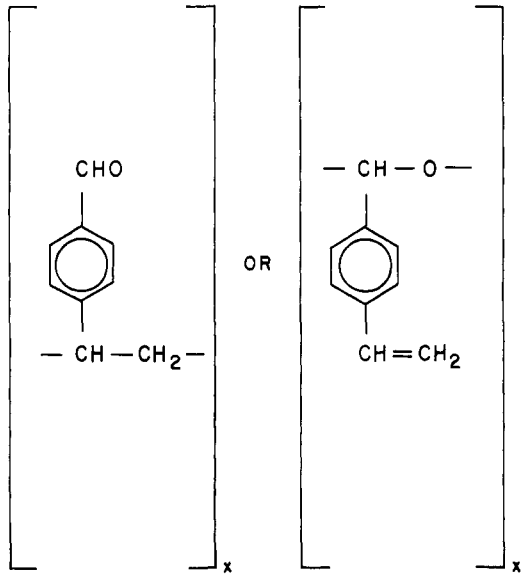
MDT′ copolymer, where T′ = $CH_2=CH-SiO_{1.5}$

To index this copolymer, clear fractions to get the smallest whole numbers and take the units to be in a 1:1:1 ratio, regardless of the actual ratios in the copolymer.

| | | |
|---|---|---|
| M | $(CH_3)_3SiO_{0.5} \times 2$ | $= C_6H_{18}OSi_2$ |
| D | $(CH_3)_2SiO$ | $= C_2H_6OSi$ |
| T′ | $CH_2=CH-SiO_{1.5} \times 2$ | $= C_4H_6O_3Si_2$ |
| | Copolymer indexed as | $C_{12}H_{30}O_5Si_5$ |

## SOME LIMITATIONS AND SPECIAL PROBLEMS

In this formula index, as in all formula indexes, two different substances could have the same formula. An example, taken from a recent discussion (6) of problems in polymer nomenclature, is the polymerization of *p*-vinylbenzaldehyde, which can give two different substances:




Both polymers would be indexed as $C_9H_8O$.

Formula indexing of polymers does not distinguish among random, alternating, block, or graft copolymers containing the same kinds of units. All of these copolymers would have the same formula and would be found together in the file.

A special problem for the indexer is how to handle the ends, the branches, and other minor irregularities in a polymer. It has been our practice to include in the formula of the polymer a branching group or an end group only when the end or branch can be detected readily in the infrared spectrum. For example, in an MQD silicone copolymer, M is an important part of the structure and can be detected readily by infrared spectroscopy; consequently M is included when the formula of the copolymer is determined. On the other hand, in an $MD_xM$ silicone gum where $x \cong 10{,}000$, the end groups are not detected and M is not used in determining the formula. In searching for a polymer by its formula, if there is any question about end groups or branched groups, we search for the formulas obtained both ways—first by neglecting these groups and then by including them.

Finally, there may be certain kinds of polymers whose structures are so complex that it is not sensible to try to index them by formula. This would be the case with some natural polymers like fibroin or keratin. Even in some synthetic polymers, like the silicone-alkyd resins, it is not always clear how the various ingredients combine. For such complex polymers, it has been our practice to file data under a special heading (*e.g.*, "silicone-alkyd") rather than try to work out a molecular formula.

Despite some limitations, we have found that formula-indexing is an efficient way to handle information about polymers. In our laboratory, we file reference spectra of polymers by formula. The arrangement of symbols in formulas and the sequence of formulas in the file are identical with CAS conventions.

## SUMMARY

The following indexing rules are proposed:

Index homopolymers under the formula of the simplest repeating unit.

Index copolymers by considering them to have a 1:1:1... molar ratio of units, regardless of the actual ratios in the copolymer. Add the formulas of the different kinds of units to get the formula of the copolymer.

When repeating units contain fractional atoms, clear the fractions to the smallest whole numbers before counting atoms to determine the formula of the polymer.

Follow CAS rules in the arrangement of symbols in formulas and the sequence of formulas in the file.

## LITERATURE CITED

(1) *CA*, Vol. **64**, Formula Index, January–June, 1966, pp. 1F–2F, The Chemical Abstracts Service, Columbus, Ohio, 1967.
(2) "Codes and Instructions for Wyandotte-ASTM Punched Cards Indexing Spectral Absorption Data," American Society for Testing and Materials, Philadelphia, Pa., 1964.
(3) Kuentzel, L. E., J. CHEM. DOC. **7**, 23 (1967).
(4) Wilcock, D. F., *J. Am. Chem. Soc.* **69**, 477 (1947).
(5) Rochow, E. G., "An Introduction to the Chemistry of the Silicones," 2nd ed., Chap. 6, Wiley, New York, 1951.
(6) Fox, R. B., J. CHEM. DOC. **7**, 74 (1967).