ACKNOWLEDGMENT

W. C. Broad is thanked for his assistance and constructive criticism at various stages of the work reported in this paper.

LITERATURE CITED

(1) Wiswesser, W. J., "Literature Sources of Mammalian Toxicity Data, with Special Emphasis on Tabulating Machinery Applications," Advances in Chemistry Series, No. 16, American Chemical Society, Washington, D. C., 1956, p 64.
(2) Barnard, A. J., Jr., Kleppinger, C. T., Wiswesser, W. J., *J. Chem. Doc.*, **6**, 48 (1966); presented at the 150th National Meeting of the American Chemical Society, Atlantic City, N. J., Sept 12, 1965.
(3) Heumann, K. F., Dale, E., "Statistical Survey of Chemical Structures," in "Progress Report in Chemical Literature Retrieval," G. L. Peakes, A. Kent, and J. W. Perry, Ed., Interscience Publishers, Inc., New York, N. Y., 1957, pp 201–214.
(4) Wiswesser, W. J., "A Line-Formula Chemical Notation," Thomas Y. Crowell Co., New York, N. Y., 1954, 149 pp; Smith, E. G., Wiswesser, W. J., *ibid.*, 2nd ed, in preparation.
(5) Brasie, W. C., Liou, D. W., *Chem. Eng. Progr.*, **61**, No. 5, 102 (1965).
(6) Wiswesser, W. J., *ibid.*, **61**, No. 6, 19 (1965).
(7) Gibson, G. W., *ibid.*, **61**, No. 7, 12 (1965).
(8) Brasie, W. C., Liou, D. W., *ibid.*, **61**, No. 7, 16 (1965).
(9) Wiswesser, W. J., unpublished data.
(10) Benson, F. R., Abstracts, 124th National Meeting of the American Chemical Society, Chicago, Ill., Sept 1953, p 5G.
(11) Sorter, P. F., Granito, C. E., Gilmer, J. C., Gelberg, A., Metcalf, E. A., *J. Chem. Doc.*, **4**, 56 (1964); Granito, C. E., Gelberg, A., Schultz, J. E., Gibson, G. W., Metcalf, E. A., *ibid.*, **5**, 52 (1965); Granito, C. E., Schultz, J. E., Gibson, G. W., Gelberg, A., Williams, R. J., Metcalf, E. A., *ibid.*, **5**, 229 (1965).

Computer-Oriented Chemical Names*

A. J. BARNARD, JR., and C. T. KLEPPINGER
J. T. Baker Chemical Company, Phillipsburg, New Jersey

and W. J. WISWESSER
U. S. Army Biological Laboratories, Fort Detrick, Frederick, Maryland

Received October 29, 1965

A system is described for the transcription of chemical names into capital letters, arabic numerals, and four symbols, namely, & - * /. To allow such transcribed names to be accommodated in a fixed punched-card field, regular practices are introduced for the contraction of frequently cited functional groups, prefixes, and suffixes.

In the management of structural information for small-to-medium size collections, the authors have demonstrated (*1*) the effective use of a readily mastered structural-atomic code, namely, the BATCH Number. The mechanics involved in the preparation of classified structure directories based on the use of this code and the IBM-card layout employed are considered in the previous paper (*1*). One problem that required detailed study was the recording of systematic chemical names in the limited typography allowed by conventional electronic-processing equipment and their contraction, where necessary, so that they could be accommodated in a fixed punched-card field. A fixed field is necessary not only to save invested punched-card capacity but to secure economy of space in any extensive listing. The goal is to achieve transcription and contraction of a chemical name without loss of structural information and with the result still capable of being read by the interested technologist after only brief inspection of the practices adopted. The systematic approaches adopted for the "computer-oriented" expression and contraction of chemical names are the subject of this paper and are employed in the J. T. Baker BATCH Directory (*2*), which considers over 5000 compounds of carbon offered by a single supplier of laboratory chemicals. Certain of the practices adopted in the transcription of chemical names have evolved from the suggestions of Wiswesser presented in 1953 (*3*) and published in 1956 (*4*).

* Based on a paper presented before the Division of Chemical Literature, 150th National Meeting of the American Chemical Society, Atlantic City, N. J., Sept. 13, 1965.

COMPUTER-ORIENTED TRANSCRIPTION OF CHEMICAL NAMES

Basically, all conventional tabulating machinery allows the use of only capital letters and the arabic numerals. It is difficult to envision the economic transcription of systematic organic chemical names with all of their typographical complexities into these 36 characters (and the use of a blank space as a 37th one). Preliminary studies suggested that at least four additional symbols would be required even if a single symbol was assigned different meanings in different positions within a name; that is,

in the prefix, main name, and suffix. The selection of symbols must be based on those most widely available on computing and accounting equipment. Consequently, the selection was limited to the four FORTRAN marks that represent the four operations of arithmetic, namely, the ampersand (&), the hyphen (-), the asterisk (*), and the slash (/).

Special Symbols and Their Use. The adaption of these four symbols immediately prompted the use of a *slash* "polarized" by a *hyphen* preceeding it or following it to designate a left or right parenthesis or bracket, respectively, that is -/ or /-. However, when two expressions within parentheses (or within parentheses and brackets) begin or end at the same point, the intermediate hyphen must be omitted to avoid ambiguity; that is, -// is written in an opening position and //- in a closing position, rather than -/-/ and /-/-, respectively. The distinctive use of *brackets* to set off ring descriptions in present-day organic nomenclature can be maintained by the employment of the slash marks without hyphens. For example, the cyclic hydrocarbon dibenz[*ac*]anthracene would be transcribed as DIBENZ/AC/ANTHRACENE.

A further, almost self-evident, substitution is the use of a *hyphen* in place of a comma as a separator in number and letter prefixes. For example, 1,3,6- would be transcribed as 1-3-6- and *N,N*- as N-N-. Finally, a prime mark on a letter or number could be represented by an *asterisk*. For example, 2′,3″- would be transcribed as 2*-3**-. This assignment seems appropriate since the asterisk is often treated as a superscript mark.

Greek-Letter Transcription. The principal typographical complication remaining to be resolved was the transcription of Greek letters. These could be spelled out, as has been done in some computer-produced indexes, but costly space would be required. It was found that these letters can be denoted readily in computer-oriented names by addition of an *ampersand* to the corresponding Latin or phonetic or pictorial Roman-letter equivalent; thus, *alpha-* is represented by A&-, *beta-* by B&-, and *gamma-* by G&-. The complete listing of the Greek letter transcriptions are summarized in Table I. Note that MU- and PI-, which are already two-letter combinations, are spelled as such and *omicron-*, which is rarely encountered in chemical nomenclature, is abbreviated as OM-. These three departures allow O&-, M&-, and P&- to be used to designate the widely used prefix marks *ortho-*, *meta-*, and *para-*. It may be added that the representation of *psi-*, namely, Y&-, can also be used to designate the "pseudo" prefix, which is encountered in trivial names for natural products. Of course, with any of these transcriptions an asterisk can be appended to denote a prime mark. For example, A&*- and O&*- designate α'- and o'-.

Table I. Computer-Oriented Transcriptions of Greek Letters[a]

A&-	alpha- or 1st	O&-	*ortho-*
B&-	beta- or 2nd	OM-	omicron- or 15th
C&-	chi- or 22nd	P&-	*para-*
D&-	delta- or 4th	PI-	pi- or 16th
E&-	epsilon- or 5th	Q&-	theta- or 8th
F&-	phi- or 21st	R&-	rho- or 17th
G&-	gamma- or 3rd	S&-	sigma- or 18th
H&-	eta- or 7th	T&-	tau- or 19th
I&-	iota- or 9th	U&-	upsilon- or 20th
J&-	(generic)	V&-	(variable or uncertain)
K&-	kappa- or 10th	W&-	omega- or 24th or terminal
L&-	lambda- or 11th	X&-	xi- or 14th
M&-	*meta-*	Y&-	psi- or 23rd or *pseudo-*
MU-	mu- or 12th	Z&-	zeta- or 6th
N&-	nu- or 13th		

[a] Arranged in order of the roman alphabet with an indication of position in the Greek alphabet.

Miscellaneous Practices. Signs for optical rotation are an important part of the systematic treatment of the names of many natural products, notably the saccharides and amino acids. Although plus and minus signs are basic to computer typography, they are not printed as such by many card-printing interpreters. Accordingly, a typographically simpler solution is adopted, namely, the use of DX- for *dextro-*, that is, for -(+)- or *d*-, and LV- for *levo-*, that is, for -(−)- or *l*-. These assignments effectively complement the use of D- and L- for the absolute configurational families. For example, D-(+)-maltose is transcribed as D-DX-MALTOSE and L-(+)-valine as L-DX-VALINE.

Obviously, most prefix terms employed regularly in systematic chemical nomenclature can be transposed to capital letters without loss of meaning or the introduction of ambiguity. Some common terms, thus transposed, include ACI-, ANTI-, CIS-, ENDO-, EXO-, ISO-, MESO-, PERI-, SEC-, TERT-, and TRANS-, and also N-, O-, and S- to designate nitrogen, oxygen, and sulfur substitution, respectively.

Symbols for the Elements. In the Wiswesser Line Notation (*5*), the elements less frequently encountered in organic compounds are all designated by symbols consisting of two capital letters. This practice was retained in the treatment recommended in the previous paper (*1*) for an empirical formula field in punched cards. Consequently, carbon, hydrogen, boron, fluorine, iodine, nitrogen, oxygen, phosphorus, and sulfur retain their approved one-letter atomic symbols; to secure two-letter symbols for all of the remaining elements, augmented atomic symbols are required for potassium, uranium, vanadium, molybdenum (*i.e.*, wolfram), and yttrium, namely, KA, UR, VA, WO, and YT. Thus the name metanilic acid, potassium salt can be denoted by METANILIC ACID -KA SALT. This substitution of a symbol for the name of an element is best viewed as a part of the transcription process, rather than as a contraction.

CONTRACTION OF COMPUTER-ORIENTED TRANSCRIPTIONS

With these special practices, the transcription of most chemical names into capital letters, arabic numerals, and the four selected special symbols, proceeds smoothly and consistently. Indeed, the use of the capital letters facilitates visual recognition of entries of interest in long listings.

Field Size. The use of a fixed punched-card field for these computer-oriented transcriptions is essential in most applications. This requires selection of an appropriate size for the field and contraction of longer names so that they can be accommodated. Benson (*6*) some years ago punched the assigned commodity names of about 3360 compounds then offered by Distillation Products into punched cards, using essentially no contractions or special

symbols and representing the Greek letters and *ortho-*, *meta-*, and *para-* by single letters (*e.g.*, A-, B-, G-, ..., O-, M-, P-). Analysis of the length of the names in Benson's deck revealed that 82% had a length of 25 characters or less (*7*). It was also noted that Benson reserved the first three columns of the field for any prefix to the main name.

Based on Benson's experience, the assigned commodity names of about 1000 representative compounds of the J. T. Baker line of organic laboratory chemicals were keypunched into a field consisting of six columns to accommodate any prefix to the main name, and 28 columns for the main name—a total of 34 columns. Listings immediately revealed the value of a variable length, offset prefix in securing ready recognition of entries. Retention of this feature seemed desirable. However, informal statistical study indicated that a larger number of columns was needed for the prefix if the loss of salient information was to be avoided with a significant proportion of the compounds (notably polyhalogen derivatives of benzenoid compounds). Study revealed that the reservation of nine columns for the name prefix was appropriate. This decision was influenced by two considerations: (1) the recognition of the punched-card capacity that thereby would seldom be used if more columns were reserved, and (2) the need to keep a line of listing in any directory to not more than 46 characters so that reduction to 50% of original size might allow offset printing of three columns on an 8.5 × 11 in. sheet of paper (*1*). In addition, the preliminary study indicated that reduction of the main-name portion of the field to only 23 columns was feasible. This decision suggested that less than one compound in six would require pronounced contraction. The total recommended name field is therefore 32 characters with 9 constituting a variable length offset prefix. Contraction of a name is recommended only where its prefix or main name cannot be accommodated in the assigned space.

Where space limitations do not permit the prefix to be offset, a field of 26 columns may be appropriate for general purposes. As noted in the previous paper (*1*), all but about 1% of the 5030 compounds of carbon in the J. T. Baker deck could be accommodated in such a name field of 24 columns without loss of structural information.

Contraction of the Name Prefix. Obviously, to confine the prefix to the name within nine characters, some contractions must be adopted. Hyphens, used in place of commas as separators, can often be omitted without ambiguity. An extreme example is 33*4*55*7HEXA HYDROXYFLAVONE as the contracted transcription for 3,3′,4′,5,5′,7-hexahydroxyflavone. However, a hyphen should be retained with 10 or a higher number. For example, 1289-10- is preferred, rather than 128910-, for the contracted transcription of 1,2,8,9,10-. In the prefix position, the transcription of *cis-* can be reduced to CS-, *trans-* to TR-, *cis–trans-* (mixture) to C/T-, *sec-* to SC-, and *tert-* to T-.

Use of Hyphen to Separate Contractions in Main Name. In main names, contractions are best set off by hyphens to avoid the confusing appearance of merged contractions. A hyphen separating a locant from its contracted function may, however, be omitted without risk of any such confusion where extreme compression is required.

Contraction of Halogen Substituents. The halogen symbols are so readily recognized that they can be applied unchanged, other than to write them completely in capital letters (*i.e.*, BR, CL, F, I), in order to represent the corresponding halo or halide names. Thus, a 4-bromo substituent can be contracted as 4-BR- or, where extreme compression is necessary, as 4BR-. A contracted transcription for *p*-toluenesulfonyl chloride would be P&-TOLU ENESULFONYL CL. (Iodide as a suffix anion is best represented as a spaced /I.)

The hydrohalide salts of organic bases are frequently encountered. Such salts are readily recognized if the salt portion is written as HBR, HCL, or HI. Where a multiplier is encountered, two possibilities exist: either use of the letter multiplier with a hyphen appended or use of the appropriate arabic numeral as a number multiplier (see below). Thus, ETHYLENEDIAMINE DI-HCL and ETHYLENEDIAMINE HCL2 are acceptable.

Number Multipliers. The prefixed multipliers di-, tri-, tetra-, etc. tend to "scatter" the functions being multiplied and in addition take up considerable space. Accordingly, the substitution of a *suffixed* arabic numeral as the multiplier is an attractive elective when the function is to be contracted. This multiplier *must directly follow* the contraction and should also be separated from any following term, especially a locant number, in order to avoid ambiguity. For example, 2-5-BR2-A&A&A&-F3-TOLUENE is a contracted transcription for 2,5-dibromo-α,α,α-trifluorotoluene. Where necessary, as described above, a hyphen may be omitted in separating a locant number from its contracted function; consequently, 25BR2- would be an acceptable extreme contraction for the 2,5-dibromo- grouping.

Contraction of Alkanes and Alkane Functions. The alkanes, alkenes, alkanols, alkanoic acids and their esters and salts, alkanoamides, and the alkyl, alkoxyl, and alkanoyl groups occur so frequently in systematic organic nomenclature that the adoption of regularly formed, easily recognized contractions for them is essential. It was realized that the two-letter contractions for the lower *n*-alkyl groups, first introduced in 1866 and still popular, should be retained, expressed in capital letters, namely, ME, ET, PR, and BU. In addition, isopropyl and isobutyl groups can be denoted by IPR and IBU. Wiswesser (*3*, *4*) earlier had extended the contractions of *n*-alkyl groups to AM or PE, HX, HP, OC, and DEC for amyl or pentyl, hexyl, heptyl, octyl, and decyl, respectively. These contractions are maintained, but PN is now recommended for pentyl, rather than PE. Since the principle of least effort would dictate that more clues should be provided in the contraction of rarer functions, contractions of three to five letters are recommended for the higher alkyl groups beyond octyl. The complete assignment of contractions for *n*-alkyl and lower branched chain alkyl groups through C_{22} is given in Table II. A hyphen is appended to an alkyl contraction to separate it from following functions, except where a blank space following serves the same purpose, as in esters (*e.g.*, ET HEPTAFLUORO BUTYRATE).

Alkoxyl groups can be indicated by addition of an O to the contraction for the alkyl group, so that ETO- represents the ethoxy group.

An alkane in a terminal position can be represented

Table II. Contraction of Alkanes, Alkyl, Alkoxyl and Acyl Groups, Alkanols, Aldehydes, Alkanones, Alkanoic Acids, and Alkanoamides

Number of carbons	Contraction for Hydrocarbon[a,b]	Alkyl group[b,c,f]	Acyl group[b,e,h]
1	ME*	ME-	FO&
2	ET*	ET-	AC&
3	PR*	PR-	PR&
(iso)	...	IPR-	...
4	BU*	BU-	BU&
4(iso)	(IBU*)	IBU-	IBU&
	2-ME-PR*		
5	PN*	PN-	PN&
		(AM-)	VL&
5(iso)	(IPN*)	IPN-	IVL&
		(IAM-)	
	3-ME-BU*	3ME-BU-	3ME-BU&
5(tert)	(TERT-PN*)	TERT-PN-	...
		(TERT-AM-)	
	(T-PN*)	T-PN-	...
	2-ME-2-BU*	2ME-2-BU	...
5(neo)	(NEOPN*)	NEOPN-	PIV&
	22-ME2-PR*	22ME2-PR-	
6	HX*	HX-	HX&
6(iso)	(IHX*)	(IHX-)	...
	4-ME-PN*	4ME-PN-	4ME-VL&
7	HP*	HP-	HP&
7(iso)	(IHP*)	(IHP-)	...
	5-ME-HX*	5ME-HX-	5ME-HX&
8	OC*	OC-	OC&
9	NON*	NON-	NON&
10	DEC*	DEC-	DEC&
11	UND*	UND-	UND&
12	DODC*	DODC-	DODC&
			LAU&
13	TRDC*	TRDC-	TRDC&
14	TTDC*	TTDC-	TTDC&
			MYR&
15	PNDC*	PNDC-	PNDC&
16	HXDC*	HXDC-	HXDC&
		CET-	PAL&
17	HPDC*	HPDC-	HPDC&
18	OCDC*	OCDC-	OCDC&
			STE&
19	NODC*	NODC-	NODC&
20	EICOS*	EICOS-	EICOS&
			(ARAC&)
21	HNCOS*	HNCOS-	HNCO&
22	DOCOS*	DOCOS-	DOCOS&
			(BEHE&)

[a] Contractions shown within parentheses are for names not recommended by the indexing practices of *Chemical Abstracts* as of 1964–1965. [b] Midname contraction is given. Where hydrocarbon is last term in name, omit the asterisk (*e.g.*, 2-BROMOBU = 2-bromobutane). [c] For corresponding olefin, add terminal E (*e.g.*, 1-BU*E = 1-butene); however, MTN and ETN are preferred for methylene and ethylene, respectively. [d] For corresponding alkanol or alkanone, add terminal OL (*e.g.*, 2-BU*OL = 2-butanol and 2-BU*ONE = 2-butanone). [e] For corresponding alkoxyl group, insert O (*e.g.*, MEO-, ETO-, PRO-, etc.). [f] Omit terminal hyphen in alkyl contraction when a blank space separates it from remainder of name, as for esters (*e.g.*, ET LAU& = ethyl laurate). [g] For corresponding "ic acid," add terminal A (*e.g.*, FO&A and AC&A denote formic and acetic acid, respectively). [h] For acid suffix (-ate), precede indicated contraction with a hyphen or with a space (directly if unsubstituted) (*e.g.*, ET BU& = ethyl butyrate; PR 2-IODO-IBU& = propyl 2-iodoisobutyrate). [i] For corresponding amide, add 'M' to indicated contraction (*e.g.*, BU&M = butyramide) or for "amido" form add M- (*e.g.*, BU&M = butyramido). [j] For corresponding aldehyde, add H to indicated contraction (*e.g.*, BU&H = butyraldehyde or butanal). [k] The names for the *higher* acyl groups and acids not derived from corresponding alkanes under the indexing practices of *Chemical Abstracts*, 1964–1965, are limited to the unsubstituted species only; this restriction applies to pivalate (PV&), laurate (LAU&), myristate (MYR&), palmitate (PAL&), and stearate (STE&).

simply by use of the contraction for the corresponding alkyl group. (In the rare case where an alkyl group is the very last term of a name, the -yl ending is denoted by a slash mark.) In all other positions, it is imperative to distinguish clearly an alkane from an alkyl function. This is accomplished by adding an asterisk to the contraction for the corresponding alkyl; the asterisk here has the force of supplying the "ane" ending. For example, 1-4-BU*DIAMINE would be a contracted transcription for 1,4-butanediamine.

Alkenes, that is, olefins, can be indicated by appending an E to the contraction for the corresponding alkane. Thus, 1-BU*E represents 1-butene or 1-butylene. Because of their high frequency, methylene and ethylene are preferably designated by the shorter contractions MTN and ETN, respectively. Thus, ethylenediamine would preferably be represented as ETN-DIAMINE, rather than as ET*E-DIAMINE.

Alkanols can be contracted by simply adding OL to the contraction for the corresponding alkane. Thus, 3-methyl-2-butanol can be represented as 3-ME-2-BU*OL. Analogously, alkanones can be denoted by adding ONE to the alkane contraction. Thus, 2-BU*ONE represents 2-butanone.

The names assigned to many of the aliphatic monocarboxylic acids are related to the names of the alkane (or alkyl group) from which they are derived by oxidation (*e.g.*, *but*ane or *butyl* — *butyr*ic acid). Consequently, it is desirable to maintain this correspondence in any assigned contractions. The ampersand is selected as an operator that "oxygenates" a contraction or that at least adds an "ate" suffix to it. Addition of an ampersand to an alkyl contraction is taken to represent the corresponding alkano*ate*. Thus, ET BU& denotes ethyl butyrate. In a nonsuffix position an alkyl contraction with an appended ampersand designates an alkanoyl group, RCO-. Thus, 4-BR-BU& CL denotes 4-bromobutyryl chloride. An alkanoyl contraction is followed by either a hyphen or a blank space, as in the example. The use of this alkanoyl contraction with an additional letter allows other functions to be represented; thus, &A denotes an acid, &H an aldehyde, and &M an amide. Consequently, BU&A, BU&H, and BU&M denote butyric acid, butyraldehyde (that is, butanal), and butyramide, respectively. The "amido" form is indicated in a nonterminal position by an appended hyphen; thus, 2-BR-BU&M- is 2-bromobutyramido-.

A special situation exists where the name assigned the

aliphatic acid (and its derived functions) is not related to the corresponding alkane. Here the ampersand-operator is consistently applied to the contracted root supplied by the acid. For example, AC&-, AC&, AC&A, AC&H, AC&M, AC&M- denote acetyl-, acetate, acetic acid, acetaldehyde, acetamide, and acetamido-, respectively.

The full treatment recommended for the systematic contraction of the alkanes and derived functions through C_{22} is summarized in Table II. The analogous treatment applied to the alkene and alkyne carboxylic acids, the hydroxy and oxo alkane carboxylic acids, and the alkane and alkene dicarboxylic acids are summarized in Tables III, IV, and V, respectively. As will be seen from the last two of these tables, two ampersands may be used to indicate a difunctional group where space permits this added clarity. For example, the succinyl group, $-COCH_2CH_2CO-$, is contracted as SUC&&-, and the succinate group as SUC&&. However, succinic acid is represented satisfactorily as SUC&A.

Additional examples of the use of the ampersand as an "oxygenating" operator will be found in Tables VI through IX. This operator is also extended to some non-carbon functions to imply either "oxygenated" or "-ate" or indeed both (see below).

Contractions for the cycloalkanes and cycloalkyl groups are achieved by allowing CY as a prefix to denote "cyclo;" thus, CYHX indicates cyclohexyl or, in a terminal position, cyclohexane. In an intermediate position, CYHX* denotes cyclohexane. For example, ME4-1-3-CYBU*ONE would denote tetramethyl-1,3-cyclobutanone.

Table III. Contractions for the Common Alkene, Arylalkene, and Alkyne Carboxylic Acids

Acid form		"Ate" form	
Name[a]	Contraction[b,c]	Name	Contraction[d]
Acrylic acid	ACR&A	Acrylate	ACR&
Angelic acid	ANGL&A	Angelate	ANGL&
Atropic acid	ATRO&A	Atropate	ATRO&
Brassidic acid	BRASS&A	Brassidate	BRASS&
Cinnamic acid	CINN&A	Cinnamate	CINN&
Crotonic acid	CROT&A	Crotonate	CROT&
Elaidic acid	ELAI&A	Elaidate	ELAI&
Erucic acid	ERUC&A	Erucate	ERUC&
Hydrosorbic acid	H-SORB&A	Hydrosorbate	H-SORB&
Isocrotonic acid	ICROT&A	Isocrotonate	ICROT&
Linoleic acid	LINO&A	Linoleate	LINO&
Methacrylic acid	ME-ACR&A	Methacrylate	ME-ACR&
Oleic acid	OLE&A	Oleate	OLE&
Propiolic acid	PRY&A	Propiolate	PRY&
Tetrolic acid	TETO&A	Tetrolate	TETO&A
Tiglic acid	TIGL&A	Tiglate	TIGL&
Sorbic acid	SORB&A	Sorbate	SORB&
Vaccenic acid	VACC&A	Vaccenate	VACC&

[a] Under the indexing practices of *Chemical Abstracts*, 1964-1965, angelic, isocrotonic, methacrylic, and tiglic acids are retained for the unsubstituted acids only. Also brassidic, erucic, and vaccenic acids are not recommended; instead *trans*-13-docosenoic, *cis*-13-docosenoic, and 11-octadecenoic acids, respectively, are used. [b] For an alkenoic acid (Geneva) name, append &A to the contraction for the alkene group (see Table II). Thus, 2-BU*E&A = 2-butenoic acid. [c] For the corresponding amide, replace the A by M. Thus, ME-ACR&M = methacrylamide. [d] For the "oyl" form, follow the indicated "ate" form by a hyphen or a blank space; for example, CROT& CL = crotonyl chloride.

Table IV. Contractions for the Common Hydroxy and Oxo Alkane and Arylalkane Carboxylic Acids

Acid form		"Ate" form	
Name	Contraction[a,b]	Name	Contraction[c,d]
Acetoacetic acid	AC&AC&A	Acetoacetate	AC&AC&
Benzilic acid	BZIL&A	Benzilate	BZIL&
Citric acid	CIT&A	Citrate	CIT&
Glyceric acid	GLCR&A	Glycerate	GLCR&
Glycolic acid	GLCO&A	Glycolate	GLCO&
Glyoxylic acid	GLYX&A	Glyoxylate	GLYX&
Lactic acid	LAC&A	Lactate	LAC&
Levulinic acid	LEVU&A	Levulinate	LEVU&
Malic acid	MAL&A	Malate	MAL&&
Mandelic acid	MAND&A	Mandelate	MAND&
Mesoxalic acid	MESX&A	Mesoxalate	MESX&&
Pyruvic acid	PYRV&A	Pyruvate	PYRV&
Tartaric acid	TART&A	Tartrate	TAR&&
Tartronic acid	TARTN&A	Tartronate	TARTN&&
Tropic acid	TROP&A	Tropate	TROP&

[a] For a hydroxy or oxo alkanoic acid (Geneva) name, prefix HO- or OXO-, respectively, to the contraction for that acid (see Table II). Thus, 2-OXO-BU&A and ET 2-HO-HP*& denote 2-oxobutyric acid and ethyl 2-hydroxyheptanoate, respectively. [b] For the corresponding amide, replace A by M; where the corresponding aldehyde is similarly named, replace A by H. [c] For the "oyl" form, follow the indicated "ate" form by a hyphen or a blank space. [d] The second ampersand of the "ate" or "oyl" forms from diprotic acids may often be omitted without serious loss of recognition, notably in the case of tartrate (TAR& *vs.* TAR&&).

Extensive contractions are sometimes required for the terms carboxy, carboxylate, carboxylic acid, carboxaldehyde, and carboxamide. These species can be denoted by CBX, CBX&, CBX&A, CBX&H, and CBX&M, respectively. For example, ET 2-CL-CYHX*CBX& denotes ethyl 2-chlorocyclohexanecarboxylate.

Contraction of Aryl Groups and Aromatic Carboxylic Acids. For the phenyl group, the long-established contraction PH is adopted. For the tolyl, xylyl, and naphthyl groups the obvious contractions are TOL, XYL, and NAP, respectively. These contractions are followed by either a hyphen or a blank space. An asterisk may be appended to give the force of "ene" to the last three contractions, that is, TOL*, XYL*, and NAP*, to denote toluene, xylene, and naphthalene. In a terminal position in a name, the asterisk can be omitted.

For benzene, the contraction BZ* is adopted. BZ alone might imply the benzoyl group. In addition, this decision allows BZ to be retained for the infrequently required contraction of benz(o) in the names of fused ring systems. In a terminal position in a name, however, BZN is recommended as the contraction for benzene. Thereby good recognition is achieved and a decision as to either retaining or dropping the asterisk in BZ* is circumvented.

The contractions for the common aromatic and arylalkane carboxylic acids follow closely the pattern established for the aliphatic acids and are summarized in Table VI. The contractions for the common arylalkene carboxylic acids and the hydroxy and oxo arylalkane carboxylic acids are considered in Tables III and IV, respectively.

The prominent aryl-alkyl group benzyl is denoted by BZL. The phenylene radical can be represented by either

Table V. Contractions for the Common Alkane and Alkene Dicarboxylic Acids

Acid form		"Ate" form	
Name[a]	Contraction[b,c]	Name	Contraction[d,e]
Adipic acid	ADI&A	Adipate	ADI&&
Azelaic acid	AZE&A	Azelate	AZE&&
Citraconic acid	CITCO&A	Citraconate	CITCO&&
Decanedioic acid	DC&&A	Decanedioate	DC&&
Fumaric acid	FUM&A	Fumarate	FUM&&
Glutaric acid	GLTR&A	Glutarate	GLTR&&
Glutaconic acid	GLTC&A	Glutaconate	GLTC&&
Heptanedioic acid	HP&&A	Heptanedioate	HP&&
Hexanedioic acid	HX&&A	Hexanedioate	HX&&
Itaconic acid	ITAC&A	Itaconate	ITAC&&
Maleic acid	MLE&A	Maleate	MLE&&
Malonic acid	MLO&A	Malonate	MLO&&
Mesaconic acid	MESC&A	Mesaconate	MESC&&
Nonanedioic acid	NON&&A	Nonanedioate	NON&&
Octanedioic acid	OC&&A	Octanedioate	OC&&
Oxalic acid	OX&&A	Oxalate	OX&&
Pimelic acid	PIM&A	Pimelate	PIM&&
Sebacic acid	SEB&A	Sebacate	SEB&&
Suberic acid	SUB&A	Suberate	SUB&&
Succinic acid	SUC&A	Succinate	SUC&&

[a] Under the indexing practices of *Chemical Abstracts*, 1964–1965, adipic, azelaic, mesaconic, pimelic, sebacic, and suberic acids are retained for the unsubstituted acids only. [b] For an alkenedioic acid (Geneva) name, append &&A to the contraction for the alkene group (see Table II). Thus, 2-OC*E&&A = 2-octenedioic acid. [c] For the corresponding diamide, replace the A by M. Thus, ADI&M = adipamide. [d] For "oyl" form, follow the indicated "ate" form by a hyphen or a blank space. Thus, ADI&& CL = adipoyl chloride. [e] Where extreme compression is required the second ampersand of the "ate" or "oyl" forms may often be omitted without serious loss of recognition. This situation is more commonly encountered with malonate (MLO& *vs.* MLO&&), oxalate (OX& *vs.* OX&&), and succinate (SUC& *vs.* SUC&&).

Table VI. Contractions for Some Aromatic and Arylalkane Carboxylic Acids

Acid form		"Ate" form	
Name[a]	Contraction[b]	Name	Contraction[c]
Anisic acid	ANIS&A	Anisate	ANIS&
Benzoic acid	BZ&A	Benzoate	BZ&
	BENZ&A		BENZ&
Cresotic acid	CRES&A	Cresotate	CRES&
Cumic acid	CUM&A	Cumate	CUM&
Gallic acid	GAL&A	Gallate	GAL&
Hydratropic acid	H-ATRO&A	Hydratropate	H-ATRO&
Hydrocinnamic acid	H-CINN&A	Hydrocinnamate	H-CINN&
Isophthalic acid	IPHT&A	Isophthalate	IPHT&&
Naphthoic acid	NAP&A	Naphthoate	NAP&
Phthalic acid	PHTL&A	Phthalate	PHTL&&
Piperonylic acid	PIPO&A	Piperonylate	PIPO&
Pyrocatechuic acid	PCAT&A	Pyrocatechuate	PCAT&
Salicylic acid	SAL&A	Salicylate	SAL&
Terephthalic acid	TPHT&A	Terephthalate	TPHT&&
Toluic acid	TOL&A	Toluate	TOL&
Vanillic acid	VANL&A	Vanillate	VANL&
Veratric acid	VERA&A	Veratrate	VERA&
Xylic acid	XYL&A	Xylate	XYL&

[a] Cumic and xylic acids are not recommended under *Chemical Abstracts* indexing practices. [b] For the corresponding aldehyde and amide, replace the A by H and M, respectively. Thus, SAL&H and SAL&M denote salicylaldehyde and salicylamide, respectively. [c] For the "oyl" form, follow the indicated "ate" form by a hyphen or a blank space. Thus, PHTL& CL = phthaloyl chloride.

PH-E or PHN; the latter contraction seems preferable in collections where this radical is encountered frequently.

In many chemical names, the phenols occur frequently as a terminal element, which should be contracted rather than some less familiar preceding radical. In such situations, the following contractions are recommended for phenol, cresol, xylenol, and naphthol: PHNL or PHL, CRSL, XYL*OL, and NAP*OL.

The contractions recommended for some additional aromatic functions are summarized in Tables VIII and IX.

Contractions for Amino Acids. The amino acids and their residues may require contraction, especially in the names for biochemicals and peptides. For these acids, the general practices adopted for the alkanoic acids are followed; however, in the case of the α-amino acids, the contraction is terminated by an N, rather than an A, to signal the presence of the α-amino group. Thus, ALA&N, GLY&N, ILU&N, LEU&N, TRP&N, TYR&N, and VAL&N denote alanine, glycine, isoleucine, leucine, tryptophan, tyrosine, and valine, respectively. To designate the suffix ("ate") form of such compounds, the terminal N is dropped and an asterisk is inserted before the ampersand. Thus, ALA*& and GLY*& denote alaninate and glycinate, respectively. Without insertion of this asterisk, the contraction in an intermediate position designates the "yl" form. Thus, ALA& and GYL& designate alanyl and glycyl. N-/N-L-ALA&-GLY&/-GLY&N would represent *N*-(*N*-L-alanylglycyl)glycine. The contractions recommended for the common amino acids and amino acid residues are summarized in Table VII.

In the case of peptides, the alternative is to adopt the practices and abbreviations, transcribed to all capital letters, recommended by the 1963 IUPAC report and employed by many biochemical journals. Thus, in the above example L-ALA-GLY-GLY might be considered as providing sufficient recognition in collections in which biochemicals are prominent.

Contraction of Amine and Ammonium Functions. The regularized contraction of groups containing atoms other than carbon, hydrogen, and oxygen has also been given attention. The amine and amino functions are prominent. In a terminal position, the term "amine" can be contracted to AMIN, AMN, or even AM, depending on the space limitations. The amino group is denoted by AMI. An instructive example is the contraction 2-//2-/ET2-AMI/-ET/-AMI/-ET*OL for 2-[[2-(diethylamino)ethyl]amino]-ethanol. Where space limitations require further compression, hyphens might be omitted to yield 2//2/ET2-AMI/ET/AMI/ET*OL, or even 2//2/ET2AMI/ET/AMI/ET*OL, both of which are still comprehensible. The relatively common aromatic amines aniline, toluidine, and xylidine can be denoted by ANILIN or ANILN, TOLDN, and XYLDN. The imine and imino functions are expressed as IMIN and IMI, respectively, where contraction is necessary (see Tables VIII and IX).

The ammonium ion can be expressed as AMM. Where it is desired to emphasize the cationic character, and space permits, the contraction is followed by a slash, or a slash

Table VII. Contractions for Amino Carboxylic Acids (and Their Derived Groups) Having Trivial Names

Acid form		"Ate" form		Acid "residue"	
Name	Contraction	Name	Contraction	Name	Contraction
Alanine	ALA&N	Alaninate	ALA*&	Alanyl	ALA&
Anserine	ANSE&N	Anserinate	ANSE*&		
Arginine	ARG&N	Argininate	ARG*&	Arginyl	ARG&
Asparagine	ASN&N	Asparaginate	ASN*&	Asparaginyl	ASN&
Aspartic acid	ASP&N	Aspartate	ASP*&	Aspartyl	ASP&
				Aspartoyl	ASP&&
Canaline	CANL&N	Canalinate	CANL*&		
Canavanine	CANV&N	Canavaninate	CANV*&	Canavanyl	CANV&
Carnosine	CARN&N	Carnosinate	CARN*&	Carnosyl	CARN&
Citrulline	CITR&N	Citrullinate	CITR*&		
Cysteic acid	CYSTC&N	Cysteate	CYSTC*&		
Cysteine	CYS&N	Cysteinate	CYS*&	Cysteinyl	CYS&
Cystine	CYSS&N	Cystinate	CYSS*&	Cystyl	CYSS&
Djenkolic acid	DJEN&N	Djenkolate	DJEN*&		
Ergothioneine	ERGO&N	...	...		
Glutamic acid	GLU&N	Glutamate	GLU*&	Glutamyl	GLU&
				Glutamoyl	GLU&&
Glutamine	GLN&N	Glutaminate	GLN*&	Glutaminyl	GLN&
Glycine	GLY&N	Glycinate	GLY*&	Glycyl	GLY&
Hippuric acid	HIPP&A	Hippurate	HIPP&	Hippuroyl	HIPP&
Histidine	HIS&N	Histidinate	HIS*&	Histidyl	HIS&
Hydantoic acid	HYDN&A	Hydantoate	HYDN&	Hydantoyl	HYDN&
Hydroxylysine	HO-LYS&N	Hydroxylysinate	HO-LYS*&	Hydroxylysyl	HO-LYS&
Hydroxyproline	HO-PRO&N	Hydroxyprolinate	HO-PRO*&	Hydroxyprolyl	HO-PRO&
Isoleucine	ILU&N	Isoleucinate	ILU*&	Isoleucyl	ILU&
Lanthionine	LANTH&N	...	...		
Leucine	LEU&N	Leucinate	LEU*&	Leucyl	LEU&
Lysine	LYS&N	Lysinate	LYS*&	Lysyl	LYS&
Methionine	MET&N	Methionate	MET*&	Methionyl	MET&
Norleucine	NLEU&N	Norleucinate	NLEU*&	Norleucyl	NLEU&
Norvaline	NVAL&N	Norvalinate	NVAL*&	Norvalyl	NVAL&
Ornithine	ORN&N	Ornithinate	ORN*&	Ornithyl	ORN&
Phenylalanine	PHE&N	Phenylalaninate	PHE*&	Phenylalanyl	PHE&
Proline	PRO&N	Prolinate	PRO*&	Prolyl	PRO&
Sarcosine	SARC&N	Sarcosinate	SARC*&	Sarcosyl	SARC&
Serine	SER&N	Serinate	SER*&	Seryl	SER&
Threonine	THR&N	Threoninate	THR*&	Threonyl	THR&
Thyroxine	THYR&N	Thyroxinate	THYR*&	Thyronyl	THYR&
Tryptophan	TRP&N	Tryptophanate	TRP*&	Tryptophanyl	TRP&
Tyrosine	TYR&N	Tyrosinate	TYR*&	Tyrosyl	TYR&
Valine	VAL&N	Valinate	VAL*&	Valyl	VAL&

is prefixed to the associated anion (see below). For example, -/2-BR-ET/-ME3-AMM/ BR denotes (2-bromoethyl)trimethylammonium bromide.

Extension of the Ampersand Operator to Other Functions. The ampersand may be used with other functions, notably noncarbon ones, as an operator to denote in a nonterminal position the oxo group or "oxygenated" and in a terminal position the "ate" ending. Thus, the high-frequency nitro group is succinctly denoted by N&. The same contraction in a terminal position denotes a substituted nitrate ion. The nitroso group is denoted by N3&.

Analogously, C&, P&, and S& indicate carbonyl or carbonate, phospho or phosphite, and sulfuryl or sulfate, respectively. Where space permits, the less-condensed contraction SULF& may be preferred for sulfate. A fine distinction is made possible when the phosphorus or sulfur is attached directly to carbon by placing a hyphen between the atomic symbol and the ampersand, thereby signaling this attachment. Thus, in intermediate and terminal positions, respectively, P-& and S-& denote phosphoro or phosphonate, and sulfonyl or sulfonate. For example, 1-BU*S-&A denotes 1-butanesulfonic acid and BZ*S-&M benzenesulfonamide. Where space permits, the longer contraction SULF-&A may be preferred for a sulfonic acid. Further examples of the use of the ampersand operator in the contraction of diverse functions are given in Tables VIII and IX. In some examples, it will be noted that the interposition of an arabic numeral, denoting an oxidation state, allows differentiation of related functions. For example, S4& denotes thionyl or sulfite and S4-& sulfinyl or sulfinate.

Designation of Cations and Ions and Contraction of Inorganic Species. Where recognition is desired of the cationic or anionic nature of a contraction, it is followed or preceeded, respectively, by a slash mark. Where a cation is associated with an anion, the slash need only appear once and the resulting contractions for the two ionic species can be separated by a blank space or, if necessary, be

Table VIII. Contractions for Some Miscellaneous Nonterminal Terms

Name	Contraction	Name	Contraction
Acetimidoyl-	AC&IM-	Lauryl-	LRL-
Acyl form	&-	Naphthyl-	NAP-
Allyl-	ALY-	Nitro-	N&-
Amidino-	AMDN-	Nitrosyl-	N3&-
Amido-	AMD-	Phenyl-	PH-
Amino-	AMI	Phenylene-	PHN-
Ammonium	AMM/	Phosphato-	P&*-
Anthraquinonyl	ANQN-	Phosphino-	P*-
Arsenoso-	AS3&-	Phosphinothioyl-	P*S-
Arsino-	AS*-	Phosphinyl-	P*&-
Arso-	AS&-	Phospho-	P&-
Arsonium	AS/	Phosphonium	P/
Arsono-	AS&A-	Phosphono-	P&A-
Auro-	AU-	Phosphoranyl-	P5-
Azido-	AZD-	Phosphoroso-	P3&-
Azo	AZO-	Pteroyl-	PTER&-
Azoxy-	AZ-O-	Pyrazolinyl-	PYRAZLN-
Benzamido-	BZ&M-	Pyridine-	PY*-
Benzene-	BZ*-	Pyridyl-	PY-
Benzo-	BZ-	Selenino-	SE4&A-
Benzyl-	BZL-	Seleninyl-	SE4&-
Benzylidene	BZLDN-	Selenonium	SE/
Benzyloxy-	BZLO-	Selenono-	SE&A-
Bismuthino-	BI*-	Selenonyl-	SE&-
Borono-	B&A-	Stibino- (or stibyl-)	SB*-
Boryl-	B*-	Stibo-	SB&-
Carbonyl-	C&-	Stibonium	SB/
Carboxyl-	CBX-	Stibono-	SB&A-
Cinnamyl-	CINN-	Stiboso-	SB3&-
Cumenyl-	CUM-	Sulfanilyl-	S&ANIL&-
Cyano-	CN-	Sulfinyl-	S4&-
Cyclo-	CY or CY-	Sulfo-	S&A-
Diacyl form	&&	Sulfonium	S/
Diamino-	DIAMI-	Sulfonyl-	S-&-
	AMI2-	Tauryl-	TAUR&-
Hydro-	H-	Telluronium	TE/
Hydrogen-	H*-	Tetra-	TT-
Hydrogen-(acidic)	H/	Tolyl-	TOL-
Hydroxy-	HO-	Xylyl-	XYL-
Iodo-	I-		
Iodonium	I/		

merged. This slash convention is especially useful for the description of inorganic ions and also of quaternary ammonium salts (see above).

A simple inorganic cation may be represented by the symbol for the metal followed by an arabic numeral designating the oxidation number and then by a slash mark. Thus, FE3/ denotes the iron(III) cation, that is, the ferric ion. It is noteworthy that the use of the slash avoids the arabic numeral being taken for a numeric operator. The oxidation state may, of course, be omitted where no significant ambiguity results. Thus, H/ and KA/ denote the hydrogen and potassium ions, respectively.

Inorganic anions can be treated analogously, using a prefixed slash mark where necessary. The simple oxyanions are indicated by the symbol of the central atom followed by its oxidation state and then the ampersand operator, which here confers the sense of an "ate" ending and of "oxygenated." Thus, AMM/ CL7& would denote ammonium perchlorate. The oxidation state can be omitted from the contraction of an anion where no ambiguity results. The vanadate ion contracted as VA& offers such an example. The contractions for some prominent anions appear in Table IX.

The practices introduced earlier in this paper taken with these contractions for simple cations and the oxyanions permit readily understood contractions for complex ions. For example, CL6-PT4& would denote the hexachloroplatinate(IV) anion, $PtCl_6^{2-}$, and BR2-N&2-PT2& would denote the dibromodinitroplatinate(II) anion, $[PtBr_2(NO_2)_2]^{2-}$. Where in present-day nomenclature an anionic ligand carries an "ato" ending, this can be designated by following the ampersand "ate" operator with an asterisk. Thus, C&*, N&*, P&*, P-&*, S&*, S-&*, SCN&*, AC&*, and OXL&* denote carbonato, nitrato, phosphato, phosphonato, sulfato, sulfonato, thiocyanato, acetato, and oxalato, respectively. For example, the tricarbonatodioxalatoplutonate(IV) anion, $[Pu(CO_3)_3(C_2O_4)_2]^{6-}$, could be expressed as C&*3-OX&*2-PU4&. The neutral ammine ligand is best designated as AMM* to avoid confusion with the contraction for the ammonium ion. The aquo ligand is represented as AQ.

Additional Contractions. In Tables VIII and IX are presented the contractions recommended for some miscellaneous nonterminal and terminal groups, respectively. Most of the entries in these tables occur often enough as to prompt regularized contractions. A few entries that would occur only frequently in specialized collections are included to offer salient examples of the extension of the practices recommended in this paper.

Some Considerations in the Contraction of Transcribed Names. Only where the transcribed name cannot be accommodated in the assigned space should its contraction be undertaken. The contraction of the prefix has already been discussed in full and no additional comments are needed. In the contraction of the main name the following observations may prove of value.

(1) Often a single letter may be dropped from the transcription of a term without serious loss of recognition. For example, the omission of the terminal E is feasible from such endings as ATE, IDE, ITE, INE, OLE, and ONE.

(2) A function that occurs frequently and/or one that has a contraction that is most readily recognized should be contracted before one that occurs infrequently or has a less obvious contraction. Thus, an alkyl group or halo group would be contracted before a nitro group.

(3) Where the use of the regular contractions just fails to provide sufficient contraction, often one or more hyphens used as separators may be omitted or a pair of hyphens associated with slash marks to designate a pair of parentheses or brackets. In addition, it is sometimes feasible to drop a letter from certain of the regular contractions and still have sufficient recognition in the context. For example, AMI can often be reduced to AM and still denote the amino group.

(4) Where a "special" contraction must be devised, as much of the original (transcribed) name should be retained as space permits, and the ampersand and asterisk opera-

Table IX. Contractions for Some Miscellaneous Terminal Terms

Name	Contraction
Acetal	ACTL
Acid	ACD
	A
Acrolein	ACR&H
Alcohol	OH
	ALC
Aldehyde	ALD
	&H
Amic acid	/AM&A
Amide	&M
	/AD
Amine	AMIN
	AMN
	AM
Ammonium salt	-AMM
Anhydride	ANH
	/ANH
Anilide	/ANIL
Aniline	ANILN
Anisidide	/ANIS
Anthranilate	ANTN&
Arsenite	AS3&
Arsinate	AS3-&
Arsonate	AS-&
"Ate" ending	&
Azide	/AZ
Benzaldehyde	BZ&H
Benzene	BZN
Benzoate	BZ&
Benzoic acid	BZ&A
	BENZ&A
Bicarbonate	H/C&
Bipheny	BIPH
Borate	B&
Boronate	B-&
Bromate	BR&
Bromide	/BR
	BR
Bromide (acyl)	&BR
	& BR
Bromite	BR3&
Cacodylate	CACO&
Calcium salt	-CA
Carbamate	CBM&
Carbohydrazide	C&H/AZ
Carbonate	C&
Carboxaldehyde	CBX&H
Carboxamide	CBX&M
Carboxylate	CBX&
Carboxylic acid	CBX&A
Chlorate	CL&
Chloride	/CL
	CL
Chloride (acyl)	&CL
	& CL
Chlorite	CL3&
Cresol	CRSL
Cyanide	/CN
Derivative	DERIV
	DV
Diamine	DIAMIN
	DIAM
Dibromide	DIBR
	/BR2
Dichloride	DICL
	/CL2
Diimide	/IM2
Dioxide	DXD
Dipotassium salt	-KA2
Disodium salt	-NA2
Disulfide	/SS
Dithiocarbamate	DTC&
	S2CMB&
Dithiocarbonate	S2C&
Epoxide	EP/O
Ethoxalate	ET-OX&
Ethoxide	ET/O
	/OET
Ethylenedinitrilo-tetraacetate	EDTA
Ethylenediamine-tetraacetate	
Ethylene oxide	ETN/O
Fluoride	/F
	F
Fluoride (acyl)	&F
	& F
Furfural	FUR&H
Furoate	FUR&
Glyceride	/GLYCR
Halide	/HAL
Hexafluorosilicate	F6-SI&
Hydrate	H&
HYD&	Selenonate SE-&
Hydrazide	H/AZ
Hydride	/H
Hydriodide	HI
Hydrobromide	HBR
Hydrochloride	HCL
Hydrogen sulfate	H/S&
Hydrogen sulfide / Hydrosulfide	/SH
Hydrogensulfide	H/S4&
Hydroxide	/OH
Hypobromite	BR1&
"ic" or "ous acid"	&A
"ide" ending	/
Imide	/IM
Imide chloride	/IM/CL
Imine	IMIN
Iodate	I&
Iodide	/1
Iodide (acyl)	&I
	& /I
Isocyanate	ICN&
Isocyanide	/ICN
	/NC
Isothiocyanate	ISCN&
	NCS&
Lactide	/LAC
Lactone	LACTN
Magnesium salt	-MG
Mercaptan	SH
Metaphosphate	M&P&
Methoxalate	ME-OX&
Methoxide	ME/O
	/OME
Mixture	MIXT
Monomer	MONMR
Naphthalene	NAP
Nitrate	N&
Nitride	/N
Nitrile	NITL
	CN
Nitrite	N3&
Oxamate	OXM&
Oxide	OXD
Oxime	XM
Ozonide	/OZ
Peptide	/PEPT
Percent	PC
	P/C
Perchlorate	CL7&
Peroxide	/OO
	PEROXD
Phenol	PHNL
	PHL
Phenone	PHNON
Phosphate	P&
Phosphide	/P
Phosphinate	P3-&
Phosphonate	P-&
Picrate	PICR&
Polymer	POLMR
Polysulfide	POLY/S
Potassium salt	-KA
Pyridine	PY
Salt	SLT
Selenide	/SE
Seleninate	SE4-&
Selenite	SE4&
Selenocyanate	SE-CN&
Selenoxide	SE/O
Semicarbazide	SCB/AZ
Silicate	SI&
Sodium salt	-NA
Solution	SOLN
Stibinate	SB3-&
Stibonate	SB-&
Succinamate	SUCM&
Sulfanilate	S&ANIL&
Sulfate	SULF&
	S&
Sulfenamide	S2-&M
Sulfenate	S2-&
Sulfide	/S
Sulfinite	S4-&
Sulfite	S4&
Sulfonamide	S-&M
Sulfonate	S-&
Sulfonic acid	S-&A
Sulfoxide	S/O
Sultone	SULTN
Taurinate	TAU&
Thallium(I) salt	-TL/1
Thiocarbamate	S-CBM&
Thiocarbonate	S-C&
Thiocyanate	SCN&
Thioformate	S-FO&
Thiol	SH
Toluidide	/TOL
Toluidine	TOLDN
Tribromide	TRI-BR
	/BR3
Trichloride	TRI-CL
	/CL3
Trithiocarbonate	S3-C&
Ureide	/URE
Vanadate	VA&
Water	H2O
Xylene	XYL
Xylidide	/XYL
Xylidine	XYLDN
Zinc salt	-ZN

tors, introduced in this paper, should be utilized where appropriate. Care should be taken that the resulting contraction is neither identical nor almost identical with a regular contraction and will not be taken as a special contraction for some other group.

(5) No regular contractions for the heterocyclic rings are presented in this paper. Where their contraction is imperative, it appears best to omit as few letters as possible. Thus, QUINOLINE might be progressively reduced to QUINOLIN, QUINOLN, and QUINLN, and even to QINLN; PYRAZOLE might be reduced to PYRAZOL, and PYRAZL, and even to PYRZL.

REMARKS

In this paper a set of practices has been described for regularizing the transcription of chemical names into capital letters, arabic numerals, and four special characters, namely, the ampersand, the asterisk, the hyphen, and the slash, and also for contracting the resulting names so that they can be accommodated within a fixed field. The authors invite criticism of the proposals and will welcome any suggestions as to possible improvements or extensions. (A single alphabetically arranged list of the names and contractions presented in this paper is available from the authors upon request.)

LITERATURE CITED

(1) Barnard, A. J., Jr., Kleppinger, C. T., Wiswesser, W. J., *J. Chem. Doc.*, **6**, 41 (1966).
(2) "J. T. Baker BATCH Directory," J. T. Baker Chemical Co., Phillipsburg, N. J., Sept. 1965, 39 pp.
(3) Wiswesser, W. J., Abstracts, 124th National Meeting of the American Chemical Society, Chicago, Ill., Sept. 1953, p 18G.
(4) Wiswesser, W. J., "Literature Sources of Mammalian Toxicity Data, with Special Emphasis on Tabulating Machinery Applications," Advances in Chemistry Series No. 16, American Chemical Society, Washington, D. C., 1956, p 64.
(5) Wiswesser, W. J., "A Line-Formula Chemical Notation," Thomas Y. Crowell, Co., New York, N. Y., 1954, 149 pp; Smith, E. G., Wisswesser, W. J., *ibid.*, 2nd ed, in preparation.
(6) Benson, F. R., Abstracts, 124th National Meeting of the American Chemical Society, Chicago, Ill., Sept. 1953, p 5G.
(7) Wiswesser, W. J., unpublished data.

Some Information Indexing Techniques in a Real-Time Hospital Computer System

RICHARD A. BOLT
Bolt Beranek and Newman Inc., Cambridge, Massachusetts 02138

SCOTT I. ALLEN
Medical Care Administration Branch, U. S. Public Health Service, Washington, D. C.

and JOHN W. WEBB
Massachusetts General Hospital, Boston, Massachusetts

Received September 15, 1965

A phonetic indexing technique has proved useful for computer retrieval of filed information, especially in those cases where misspellings in the retrieval request can readily occur. Further, the creation of special indexes to filed information can facilitate rapid selection on very large files.

INTRODUCTION

A cooperative research effort to develop a real-time Hospital Computer System is being undertaken by the Massachusetts General Hospital and Bolt Beranek and Newman Inc. The project is supported by the National Institute of General Medical Sciences and the American Hospital Association. The system uses a Time-Shared computer (the computer is a modified PDP-1, by Digital Equipment Corp., Maynard, Mass.), with remote input-output devices allowing rapid real-time collection, storage, retrieval, and dissemination of hospital information. It is hoped that the developing system will prove both a powerful adjunct to the medical and nursing staff in carrying out their tasks, and a significant aid in clinical research.