(25) Gelberg, Alan; Nelson, William; Yee, George S.; Metcalf, E. A. "A Program for Retrieval of Organic Structure Information". *J. Chem. Doc.* **1962**, *2*, 7-11.
(26) Gelberg, Alan. "Chemical Notations"; Marcel Dekker: New York, 1970; Encyclopedia of Library and Information Science 4, pp 510-528.
(27) Bonnett, H. T.; Calhoun, D. W. "Application of a Line Formula Notation in an Index of Chemical Structures". *J. Chem. Doc.* **1962**, *2*, 2-6.
(28) Tomea, Albert V.; Sorter, Peter F. "On-line Substructure Searching Utilizing Wiswesser Line Notations". *J. Chem. Inf. Comput. Sci.* **1976**, *16*, 223-227.
(29) Landee, Franc A. "Computer Programs for Handling Chemical Structures Expressed in the Wiswesser Notation"; "Abstracts of Papers", 147th National Meeting of the American Chemical Society, Philadelphia, PA, 1964; American Chemical Society: Washington, DC, 1964; p F7.
(30) Thomson, L. H.; Hyde, E.; Matthews, F. W. "Organic Search and Display using a Connectivity Matrix Derived from Wiswesser Notation". *J. Chem. Doc.* **1967**, *7*, 204-209.
(31) Granito, Charles E.; Gelberg, Alan. "Symposium on Notation Systems: Introduction". *J. Chem. Doc.* **1968**, *8*, 127.
(32) "Index Chemicus Registry System (ICRS)"; Institute for Scientific Information, Philadelphia, PA, 1968 (February announcement).
(33) Granito, Charles E.; Rosenberg, Murray D. "The Chemical Substructure Index (CSI), a New Research Tool". *J. Chem. Doc.* **1971**, *11*, 251-256.
(34) Granito, C. E.; Becker, G. T.; Roberts, S.; Wiswesser, W. J.; Windlinx, K. J. "Computer-Generated Substructure Codes (Bit Screens)". *J. Chem. Doc.* **1971**, *11*, 106-110. See: Wiswesser, W. J.; Windlinx, K. J. "Computer-Generated Substructure Codes (Bit Screens)"; U.S. Army, Fort Detrick: Frederick, MD, 1970.
(35) "Wiswesser Line Notations Crresponding to Ring Index Structures" (Report to the National Science Foundation under Contract NSF-C521); American Chemical Society: Columbus, OH, 1968.
(36) "Parent Compound Handbook"; American Chemical Society: Columbus, OH, 1976-1982.
(37) Farrell, Carl D.; Chauvenet, Alan R.; Koniver, Deena A. "Computer Generation of Wiswesser Line Notation". *J. Chem. Doc.* **1971**, *11*, 52-59.
(38) Koniver, Deena; Wiswesser, W. J.; Usdin, Earl; "Wiswesser Line Notation: Simplified Techniques for Converting Chemical Structures to WLN". *Science (Washington, D.C.)* **1972**, *176*, 1437-1439.
(39) Usdin, Earl; Efron, D. H. "Psychotropic Drugs and Related Compounds", 2nd ed.; NIMH: Rockville, MD, 1972 (WYLBUR-processed book with WLN index and 5-page introduction to WLN).
(40) Feldmann, R. J.; Koniver, D. A. "Interactive Searching of Chemical Files and Structural Diagrams Generation from Wiswesser Line Notation". *J. Chem. Doc.* **1971**, *11*, 154-159.
(41) Miller, George A. "Encoding and Decoding WLN's. *J. Chem. Doc.* **1972**, *12*, 60-66.
(42) "Bibliography of Reports Relating to Chemical Notations", CWIK List News, Oct-Dec 1979, pp 2-12.
(43) Wiswesser, W. J. "Fundamental Tetracyclic Ring Forms", Airlie House Conference, 1963.
(44) Smith, Elbert G. "The Wiswesser Line-Formula Chemical Notation"; McGraw-Hill: New York, 1968.
(45) Smith, Elbert G.; Baker, Peter A. "The Wiswesser Line-Formula Chemical Notation (WLN)", 3rd ed.; Chemical Information Management, Inc.: Cherry Hill, NJ, 1976.
(46) Johns, Trisha M.; Clare, Michael. "Getting the Most Out of the WLN. WLN as a structural Summary Medium".
(47) Bond, V. B.; Bowman, C. M.; Davison, L. C.; Roush, P. F.; Young, L. F. "Applications of the Wiswesser Line Notation at the Dow Chemical Company".
(48) (a) Walker, S. B. "Development of 'CAOCI'", presented at the "Symposium on Uses and Applications of the Wiswesser Line Notation (WLN) Today", Las Vegas, NV, Aug. 27, 1980. (b) Warr, Wendy A. "Diverse Uses and Future Prospects for Wiswesser Line-Formula Notation".
(49) Eakin, D. R. "Graphics Challenge WLN—Can WLN Hold Fast?".
(50) Slossen, Edwin E. "Creative Chemistry"; The Century Co.: New York, 1919; p 296.
(51) Berzelius, J. J. "On the Elemental Signs and the Method of Employing Them to Express Chemical Proportions". *Ann. Philos.* **1814**, *3*, 51-52. An extract of this paper is given in: Leicester, H. M.; Klickstein, H. S. "Source Book in Chemistry"; Harvard University Press: Cambridge, 1952; pp 263-265.

Chemical Supply Catalog Indexing: Now and the Future. An Ideal Place for Use of the Wiswesser Line Notation†

MURRAY D. ROSENBERG,* MARIAN Z. DEBARDELEBEN, and JOHN F. DEBARDELEBEN‡
Research and Development Department, Philip Morris Research Center, Richmond, Virginia 23261

Received February 11, 1982

This paper covers the traditional indexes found in single-source catalogs, complied catalog listings, and approaches to improve access to a chemical supplier. The approach by Fraser Williams Ltd. in their originally called "Index to Commercially Available Chemicals" (CAOCI) in compiling the listings of many chemical supply catalogs in one source, with frequent and accurate updates, and the use of Wiswesser line notation (WLN) as an additional index, now called "The Fine Chemicals Directory", are recommended.

INTRODUCTION

As the basic source for the starting materials or building blocks for practically all research and many development projects, chemicals (organic, inorganic, etc.) serve many masters in industry, government, and academia. The user of chemicals at these institutions and organizations must at one time or another resort to a chemical supply listing or catalog or, as in most cases, to many listings or catalogs to find a source or sources at a reasonable price. Significant savings are to be made by ensuring that a chemist uses a readily available starting material to prepare a compound by the most practical synthetic route. To reach this goal, chemists should have access to files of internally prepared compounds as well as commercial sources.

For those compounds prepared within a company, manual or automated systems are usually available to provide information. On the other hand, access to commercially available compounds is somewhat limited. Until recently there have been only two types of sources. We have had the individual supplier catalogs, either of a specialist nature (e.g., Alfa Products'[1] phosphonic acids listing or Pyrazine Specialties[2] catalog) or a general nature (e.g., Eastman's[3] organic chemicals catalog, Aldrich's[4] handbook of fine chemicals, and catalogs by companies like Takeda Chemical Industries,[5] Hopkin & Williams,[6] and Fluka[7]). We have also had such compilations as Chem Sources-USA,[8] Chem Sources-Europe,[9] Chemical Week's[10] Buyer's Guide, Chemical Marketing Reporter's OPD[11] (Oil, Paint, and Drug) Chemical Buyers Directory, Stanford Research Institute's[12] Chemical Economics Handbook, etc.

In February 1979, Fraser Williams (Scientific Systems) Ltd.[14] offered to make a set of the Index of Commercially Available Chemicals (CAOCI) available to the Technical

† Presented before the Division of Chemical Information, 180th National Meeting of the American Chemical Society (Second Chemical Congress of the North American Continent), Las Vegas, NV, August 27, 1980.
‡ Richmond Organics, Inc., Ashland, VA 23005.

0095-2338/82/1622-0093$01.25/0 © 1982 American Chemical Society

Information Section of the Philip Morris Research and Development Department. To evaluate the new service, those usually involved in identifying sources of chemicals at the Philip Morris Research Center such as organic chemists and the Center's librarian experimented with CAOCI and were then interviewed about its usability and possible improvements.

Traditional Searches for a Source(s) in Company-Supplied Catalogs. Traditionally the chemist has looked to a variety of supply catalogs for sources of research chemicals. The success of such a search relies upon the chemist's knowledge of the individual suppliers. Compilation of chemicals from various suppliers arose with the increasing specialization of manufacturers and with an overall lack of adequate indexing available in most individual catalogs. Although automation has become commonplace in and around the laboratory, sources of raw materials has not been an area of prime importance. Presently only Aldrich Chemical Co. offers an automated search of its own list of chemicals.

A review of supply catalogs shows that it is the exception to have multiple index points to the chemicals. Most catalogs simply list the compounds alphabetically by name. As there are a multitude of potential names for a single compound, a name index may not be a satisfactory way to search for a given chemical.

Because of nomenclature problems, some suppliers have wisely included molecular formula indexes as an alternate access point. This type of index helps in the search for a specific compound, but it is useless when a class of compounds is desired. Other suppliers provide structural drawings as user aids. But beyond the levels of indexing noted, the catalogs with sufficient access points to their contents are countable on a single hand.

Attempts to Aid in Searches for a Source(s) of Supply. (1) Older Compilations of Chemical Supply Catalogs. It was evident to many that a better way had to be found to locate a supplier of a research and/or bulk amount of a given chemical. From these realizations, the most well-known compilation evolved, Chem Sources, in 1958. Today this is considered *the* standard source.

Chem Sources is issued in January of each year by its producers.[8,9] They claim each edition is compiled anew each year, on the basis of changes supplied by the participating suppliers. The number of suppliers in the 1980 edition is 792 with "90 500 organic, inorganic, pharmaceutical, and biochemical" items listed.

The "Chemical Section" lists chemicals in alphabetic order by chemical name. "The nomenclature used in Chem Sources is taken directly from lists and product catalogs that are supplied by producers and distributors". The publishers themselves advise a check of each synonym of a compound "because of different systems of nomenclature".

Under each listed chemical, there are alphabetic–character codes (in alphabetic order) which represent the firms offering the material. After finding the desired material, the user turns to the "Companies by Code Section" to identify the vendor. If an asterisk appears beside a vendor's code, it is an indication that the vendor is a producer of the material. This information is derived from U.S. government reports and from the vendors. Also, a pair of brackets surrounding a company code under a material's name indicates a high-purity supply.

There is a "use index", e.g., aroma chemicals, antioxidants, etc. Also, an alphabetical listing by company, containing the complete name, address, code, manufacturing and sales office locations, and telephone number, is provided.

Overall, directories such as Chem Sources–U.S.A., its companion Chem Sources–Europe, Chemical Week's Buyer's Guide, Chemical Marketing Reporter's OPD (Oil, Paint, and Drug) Chemical Buyers Directory, etc. have served as useful tools to the chemical community. Also available is the Stanford Research Institute's Directory of Chemical Producers, which should be consulted for producers of commercially available chemicals for the chemical process industry.

But all these compilations have inherent problems due to limited indexing, inadequate indexing, nomenclature, outdated information, etc.

(2) CAOCI: Commercially Available Organic Chemicals Index. The next evolutionary step in this area occurred in 1973. It was a major step since it not only was a voluntary effort but also involved a new approach to the problem of chemical "sourcing". The United Kindgom (UK) chapter of the Chemical Notation Association (CNA) formed a committee to explore the feasibility of producing a comprehensive index of commercially available organic chemicals to allow information to be found at the total structure and substructure levels.[13] The committee identified companies with an interest in such a project. The companies which became involved on a continuing basis were Wellcome Foundation, Pfizer, Boots, Beecham Pharmaceuticals, Glaxo, Allen & Hanburys, Roussel-Uclaf, ICI (Plant Protection Division), and ICI (Pharmaceuticals Division).

Other companies maintained an interest in the project, although they were unable to commit any resources when the initial work began in 1975.

The catalogs chosen initially for inclusion in the indexing effort were to be limited because of the scope of the project and the effort involved. It was felt by the committee [CNA (UK)] that it would be better to use the initial stages to index a controlled number of catalogs, while gaining experience which would be useful when the catalog coverage was being enhanced. Analyses of available catalogs showed overlapping coverage. This led to the hypothesis that the best way to proceed was to include a small number of the larger, general catalogs, e.g., Aldrich, Eastman Kodak, and Pfaltz & Bauer, plus a larger number of the smaller, more specialized catalogs, e.g., PCR, Pierce, and Fluorochem.

As each catalog was added for input, significant savings in time (35–55% on each catalog) were achieved through an automated matching process based on the chemical names. Only the nonmatching names had to be processed manually.

The initial 46 catalogs provided 100 000 chemical names representing some 40 000 unique chemicals. Note the interesting "numbers game" between Chem Sources and CAOCI:Chem Sources with 90 500 chemicals from 792 sources and CAOCI with 100 000 from 46 sources. This is an interesting example of Bradford's application of Zipf's law. Somewhere between these two the middle ground lies. The file thus generated contains all the organic chemicals of known chemical structure from the catalogs noted in Table I.

The following types of compounds were excluded in the first indexing effort: inorganic compounds, polymers, NMR shift reagents, deuterated/tritiated compounds, complexes, dyes (unless the structure is well-known), proteins, vitamins, ambiguously named compounds.

The file contains at least one reference to a grade of purity of each compound, although complete details must be obtained from the original source catalog. Prices are not included.

(A) Data Contents of the File. Each entry (or compound) has the following data provided: (1) catalog name provided as a "coden", e.g., KOCH LT for Koch-Light; (2) catalog number from the vendor's catalog as given; (3) compound name as given in the vendor's catalog (foreign names are translated into English); (4) molecular formula using the Hill convention, i.e., carbon, hydrogen, and then the remaining elements in alphabetical order; two-letter atomic symbols are preceded by a hyphen (for example C_5H_5FO or C_9H_{11}-NaO_2); (5) Wiswesser line notation (WLN) as the unique and un-

Table I. Chemical Vendor Catalogs and Numbers of Compounds from Each (CAOCI)

Air Products (297)	Imhausen (213)
Aldrich (9400)	Jakem (103)
Bader (20 300)	K&K (19 083)
BASF (438)	Koch-Light (4782)
BDH Chemicals (2547)	Lancaster (1953)
BNM Labs (651)	Lonza (61)
Bristol Organics (734)	Merck
Cambrian (4983)	Mersey (211)
Chemical Samples Co. (2838)	Nipa (2088)
Cilag (168)	National Physical Labs (83)
Croda (132)	Parish (1391)
Eastman Kodak (4264)	Pfaltz & Bauer (24 143)
Emkachem (189)	PCR Inc. (2124)
Evans Chem (71)	Petrarch (515)
Fairfield (3871)	Pierce (362)
Fisons (945)	Pyraspec (25)
Fluorochem (941)	Raschig (116)
Fluka (8058)	ROC/RIC (1133)
Frinton (1076)	Schuchar (3071)
Fulmer (134)	Swiss Explosive Works (333)
Graddlink (2620)	Transworld (3728)
Hetchem (695)	Waukesha (74)
Hoechst (5909)	Willowbrook (107)
Hopkin & Williams (1193)	

ambiguous code for any one compound, which serves to gather all the synonymous names based on a given structure, and as a link between the various catalog entries.

(B) Types of Printed Indexes. There are four indexes provided.

(1) Molecular Formula Index (Figure 1). For each unique molecular formula, the molecular formula index lists the chemical names and the catalog references for all compounds represented by that molecular formula. From the information supplied, the user can directly select compounds in which he is interested without reference to further indexes. The molecular formula index is invaluable for persons who are not experienced in chemical nomenclature.

(2) WLN Index (Figure 2). For each unique compound, as defined by a unique WLN, the WLN index gives the chemical names and the catalog references for all compounds represented by that WLN. The WLN brings together compounds with the same basic nucleus and so can be used to identify salts of the same compound, derivatives of the same ring system, and so on.

(3) Source Vendor Index = Supplier Datasheets (Figure 3). This lists the names, addresses, and telephone numbers for all suppliers included in the directory and their agents throughout the world. It briefly describes the coverage of the catalog and any restrictions on purchase of the chemical. The supplier datasheets help the chemist or purchasing agent to maintain with minimum effort copies of the latest catalog available.

MOLECULAR FORMULA INDEX

```
C9H22N2
    ZY1&1N3&3
       N1,N1-DI-N-PROPYL-1,2-PROPANEDIAMINE          FAIRFLD  D-7970

    ZY1&3N2&2
       2-AMINO-5-DIETHYLAMINOPENTANE, 99%            ALDRICH  A04880
       4-AMINO-1-DIETHYLAMINOPENTANE                 BASF     10.4.1
       1-DIETHYLAMINO-R-AMINOPENTANE                 FAIRFLD  D-4340
       N*,N*-DIETHYL-1,4-PENTANEDIAMINE              FLUKA    U07641
       2-AMINO-5-DIETHYLAMINOPENTANE  85%            K & K    K 9964
       2-AMINO-5-DIETHYLAMINOPENTANE                 P & B    A20430
```

Figure 1. CAOCI Molecular Formula Index sample.

WLN INDEX

```
T56 BMJ DVH GF (C9H6FNO)
    5-FLUOROINDOLE-3-CARBOXALDEHYDE                  BADER    S36206-9

T56 BMJ DVH GO1 (C10H9NO2)
    5-METHOXYINDOLE-3-CARBOXALDEHYDE                 ALDRICH  M01494-3
    5-METHOXY-3-INDOLE CARBOXALDEHYDE                K & K    K10526
    5-METHOXYINDOLYL-3-ALDEHYDE,PURE                 KOCK LT  3652H
    5-METHOXYINDOLE-3-CARBOXALDEHYDE                 LANCSTER LAN-1117
```

Figure 2. CAOCI WLN Index.

FLUKA

CATALOGUE:	Fluka catalogue 12, 1980/81
Coden used in Directory	FLUKA
Name and address of supplier:	Fluka AG, Chemische Fabrik, CH-9470 Buchs, Switzerland
UK/European agent:	Fluorochem Ltd., Dinting Vale Trading Estate, Dinting Lane, Glossop, Derbyshire, SK13 9NU.
US agent:	Tridom Chemical Inc. 255 Oser Avenue Hauppauge, N.Y. 11787
Japanese agent:	None known
Approximate number of compounds:	10,000
Type of compounds:	General research chemicals
Ordering information:	
File Notes:	The information from the catalog has been updated with information from later supplements.

Figure 3. CAOCI Source Vendor Index/Supplier Datasheet.

Table II

chemical compilation used	compd sought: proline	chloro-acetyl chloride	furyl-acrylic acid	citrone-liol	1,1-dimethyl-hydrazine	4-ethoxy-benz-aldehyde	yohim-bine	ethyl 2-fluoro-propan-oate	thujone
Chem Sources-U.S.A.									
total no. sources found	49	27	11	20	12	21	5	3	3
unique sources			6	9	9	13	4	3	2
Chem Sources-Europe									
total no. sources found	22	18	7	11	7	8	4	1	3
unique sources			4	8	4	6	4	1	2
OPD Chemical Buyers Directory									
total no. sources found	16	5	1	6	*a*	2	1	*a*	*a*
unique sources			0	1		0	1		
Chemical Week Buyers Guide									
total no. sources found	7	6	*a*	6	2	*a*	*a*	*a*	*a*
unique sources				1	2				
CAOCI									
total no. sources found	12	14	10	6	7	9	7	2	2
unique sources			4	0	3	2	6	2	1

[a] Not indexed.

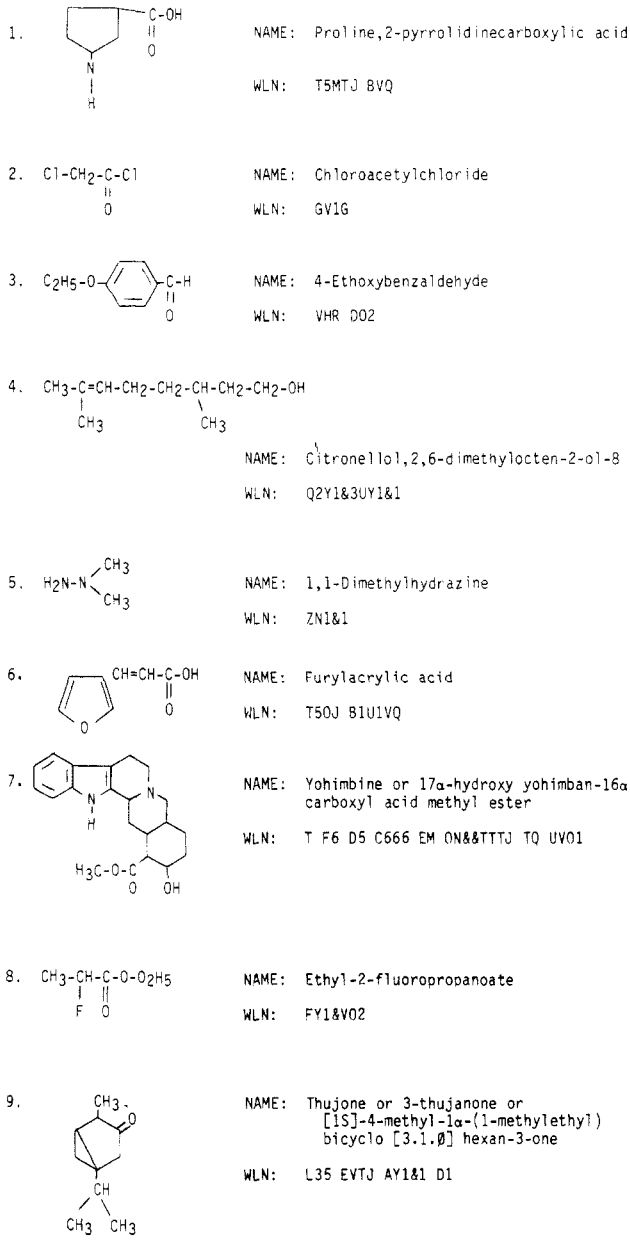

Figure 4. Compunds in search.

(C) Types of Computerized Indexes. In addition to the printed indexes in microfiche format, magnetic tapes of the three files used to generate the computer output microfiche (COM) are available.

(1) Registration File. This contains one entry for each catalog compound indexed. The file contains the following: (a) catalog coden, (b) catalog number, (c) WLN, (d) suffix portion of chemical name, (e) molecular formula, (f) chemical name.

(2) WLN File. This contains one notation for each compound in the Registration File and includes the following: (a) registration file number, (b) WLN, (c) molecular formula.

(3) Ancillary File. This contains all the additional information for each Registration File entry, except the WLN and molecular formula. The file contains the following: (a) suffix of chemical name, (b) chemical name, (c) catalog coden, (d) catalog reference number.

(D) Use of CAOCI. The CNA (UK) committee which developed CAOCI turned the further development of the files and their marketing rights over to Fraser Williams (Scientific Systems) Ltd.[14] in 1979. Fraser Williams offered the product on a commerical basis beyond those currently using the files in 1980. Originally, only the participating companies and their subsidiaries had access to the information with one exception. In late 1979, a trial set of the microfiche was offered to the Technical Information Section of the Research and Development Center of Philip Morris in Richmond, VA, for testing and evaluation.

Table III. Search Results for a Source of Furylacrylic Acid [$(C_4H_3O)CH{=}CHCOOH$][a]

Chem Sources-U.S.A.	Chem Sources-Europe	OPD Chemical Buyers Directory	CAOCI
Aldrich (ALH)	Aldrich-Europe (ALD)	ROC/RIC	Aldrich
Chemical Dynamics (CDC)	Chemie KG (EGA)		Apin Chemsmco
Chemical Procurement (CPL)	ICN Life Sciences (ICH)		Eastman
Pfaltz & Bauer (PNB)	R. N. Emanuel (RNE)		Fluka
Sigma (SIG)	Ferak Berlin (FRK)		Koch Lt ROC/RIC
International Enzymes (INZ)	Koch-Light (KLL)		Fairfield
ICN/K&K (KNK)	Nipa Labs (NIP)		K&K P&B
ICN Life Sciences (NBC)			
Spectrum Chemicals (SPE)			
Tridom Chemicals (TRI)			
Trans World Chemicals (TWO)			

[a] Not indexed in "Chemical Week Buyer's Guide".

(1) Comparison of Compound Searches Using CAOCI and Other Compilations. For comparison, the five available chemical compound or catalog compilations used were as follows: Chem Sources–U.S.A., Chem Sources–Europe, OPD Chemical Buyer's Directory, Chemical Week Buyer's Guide, CAOCI.

Nine compounds were randomly selected in discussions with chemists to provide a cross section of common organic chemical "building blocks" used in everyday syntheses. The nine compounds are listed in Figure 4.

The results of the searches through the five sources present a mixed lot. Table II provides an approximation of what one would see by using the compilations. It is evident until the middle ground is reached that there is no one compilation which can be universally recommended. We have tried to reduce the number of sources found in each to "unique" sources. This is based on eliminating common supplier names and/or subsidiary firms, where they are known. Of course, no claim for infallibility is being made. It must be suspected, also, that some of the so-called suppliers may only list a compound in its "catalog" and then purchase it elsewhere when it is ordered. Further, in comparing compilations, CAOCI is based on somewhat older catalogs than Chem Sources as it is a "testing-stage product". But, overall, it is clearly seen in the examples shown that at this time the prudent seeker of a compound should use all compilations available.

As an example of a search for a source of a compound, we have chosen furylacrylic acid. Table III presents the overall results of the search.

Where does this take us? It should be obvious, as was stated previously, that the supplier catalogs leave much to be desired in general, and the present commercially available catalog compilations still are not the best answer. It seems obvious that a novel approach to the problem was necessary. The solution was CAOCI. But CAOCI is in its infancy, and many improvements are necessary before it can be a successful commerical product.

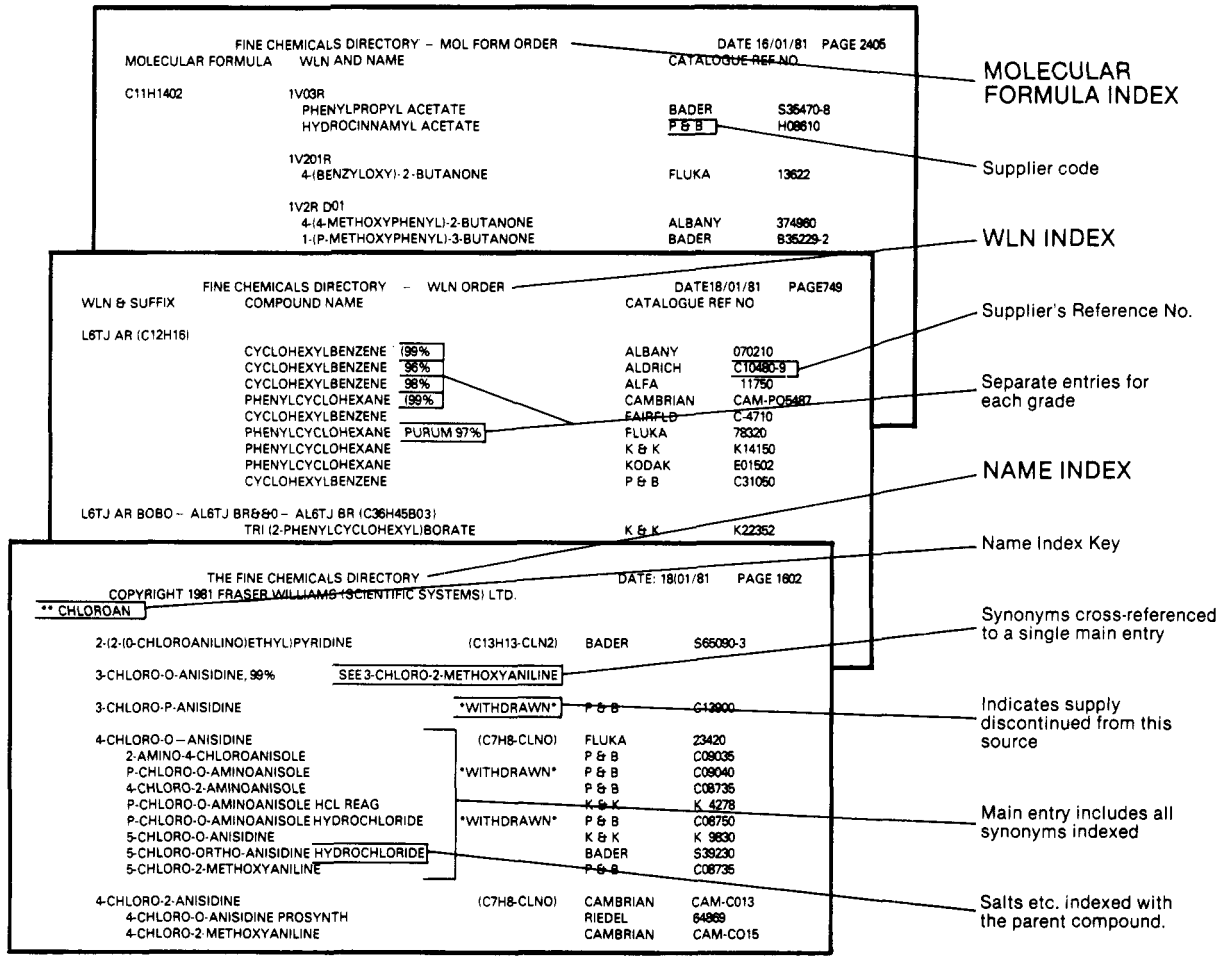

Figure 5. Indexes available (in microform) to the "Fine Chemicals Directory". These sample pages illustrate the format of the three different indexes.

(2) Improvements Needed in CAOCI. Any discussion here is directed at the microfiche and, potentially, a printed version of this service.

(a) An additional field should be added to the file for each unique compound, the Chemical Abstracts Service registry number.

(b) An additional "printed" index should be generated, again through the WLN serving as the basic reference point, and that is a chemical name index which collects all synonymous names under one entry point with in-depth cross-indexing.

(c) In the microfiche format of CAOCI, each fiche should contain unmagnified contents of each fiche; i.e., the Molecular Formula Index should indicate the range of formulas contained on a particular fiche, for example, C_5H_5Cl–$C_8H_4Br_2$ as a heading.

(d) Access to the computerized version should be obtainable via one of the universally available database vendors such as Lockheed Information Services' DIALOG or Systems Development Corporation's (SDC) ORBIT, etc. for a reasonable fee.

(e) A much larger number of vendors' catalogs should be included as our examples have shown too many gaps in each of the services.

(f) Frequent updating of the service should be carried out to reflect changes in a compound's availability and pricing.

(g) All available materials should be indexed, even if sold only by "name".

EPILOG

In the intervening months between the presentation of this paper at the 180th American Chemical Society Meeting in Aug 1980 and Nov 1981, there have been no announced plans for modifications in their printed formats or automation hints by the traditional chemical supply compilations previously cited. The American Chemical Society has announced a new service, but it will include only the bulk commercial chemicals.

The Fraser Williams (Scientific Systems) Ltd.[14] organization offered its own major revision of the original Index of Commercially Available Chemicals (CAOCI) which is called the Fine Chemicals Directory in January 1981. The improvements made over their initial offering were quite pronounced. For example: (1) addition of a chemical name index with cross references to the "main name" and the other indexes are shown in Figure 5; (2) all materials in each supplier's catalog are being indexed in some manner, i.e., by WLN and/or chemical name; (3) additional supplier catalogs are being added to the original listing of 47; (4) the annual subscription includes periodic updates; (5) the microform version includes easily read titles and contents on each fiche.

Two other areas listed in the paper of Aug 1980 which would be valuable improvements in the service are being considered but were not available at this date. They are (1) the Chemical Abstracts Service registry number index to supplier catalog entries and (2) access to the computer version of the Fine Chemicals Directory via one or more of the widely used database vendors.

The only improvement which has not been considered, but would make the service with its periodic updates almost a necessity for all purchases of chemicals, is inclusion of prices.

Overall, these actual improvements and those under consideration along with the original CAOCI concept make the Fine Chemicals Directory a "must" in a chemically oriented information center.

ACKNOWLEDGMENT

We thank the organic chemists who took the time to use the CAOCI fiche. Additionally, appreciation is due to Mrs. Frances Sanderford for her innovative efforts in preparing this manuscript.

REFERENCES AND NOTES

(1) Alfa Products, Thiokol/Ventron Division, P.O. Box 299, Danvers, MA 01923.
(2) Pyrazine Specialties, P. O. Box 6933, Atlanta, GA 30315.
(3) Eastman Kodak Co., Eastman Organic Chemicals, Rochester, NY 14650.
(4) Aldrich Chemical Company Inc., 940 W. Saint Paul Ave., Milwaukee, WI 53233.
(5) Takeda Chemical Industries Ltd., 7 Edobashi Nichome, Nihinbashi, Chuo Ku, Tokyo, Japan
(6) Hopkin & Williams, P.O. Box 1, Romford, Essex RM1 1HA, England.
(7) Fluka AG, Chemische Fabrik, CH-9470 Buchs, Switzerland.
(8) "Chem Sources-U.S.A."; Directories Publishing Company, Inc.: Ormond Beach, FL 32074.
(9) "Chem Sources-Europe"; Directories Publishing Co., Inc.: Flemington, NJ 08822.
(10) "Chemical Week Buyer's Guide"; McGraw-Hill, Inc.: New York, NY 10020.
(11) "OPD Chemical Buyers Directory"; Schnell Publishing Co., Inc.: New York, NY 10007.
(12) "Chemical Economics Handbook"; SRI International Chemical Information Services: Menlo Park, CA 94025.
(13) Walker, S. B. "Development of CAOCI and Its Use in the ICI Plant Protection Division". *J. Chem. Inf. Comput. Sci.*, paper included in this issue.
(14) The Fraser Williams Group, 140 Route 17 North, Paramus, NJ 07652, or Fraser Williams (Scientific Systems) Ltd., Flendower House, London Road South, Poynton, Cheshire SK12-1NJ, England.

Diverse Uses and Future Prospects for Wiswesser Line-Formula Notation

WENDY A. WARR

ICI Pharmaceuticals Division, Alderley Park, Macclesfield, Cheshire SK10 4TG, England

Received August 24, 1981

The many and varied uses found for Wiswesser Line Notation (WLN) in the 1970s are reviewed. The advantages of WLN-based systems for storage and retrieval of chemical structures have been proved over several years of increasingly wide usage. Future prospects for the notation are briefly considered in the light of these advantages and certain shortcomings.

Wiswesser Line Notation (WLN)[1] is a compact way of uniquely and unambiguously representing the complete topology of a chemical molecule by a linear sequence of characters. It is used by around 150 organizations worldwide in industrial, governmental, and educational fields. In addition, between 30 and 40 educational institutions teach WLN even though they do not maintain files of notations. The Chemical Notation Association (CNA) mails information to Canada, U.S., United Kingdom, France, Italy, Sweden, Japan, Denmark, Netherlands, Israel, Switzerland, India, Australia, West Germany, Romania, Austria, Belgium, Norway, and Yugoslavia. The rules of WLN are available in English, French, German, and Japanese translations.[2]

Numerous programs have been written to handle WLN, some part of in-house systems and some, such as the CROSSBOW[3] programs, commercially available. The CROSSBOW suite comprises programs for registration with WLN/molecular formula check, fragment generation, bit and string search, connection table generation, atom-by-atom search, structure display, and KWIC listing. Sixteen different organizations, to date, are using the CROSSBOW system.

Most organizations handling WLN need a checker program which generally validates notation syntax and then generates a molecular formula from the notation for comparison with an input molecular formula. ICI and other companies have used such programs for years. The CROSSBOW one is commercially available.[3] Other well-known checker programs are the ones used by the Institute for Scientific Information in connection with Index Chemicus Registry System (ICRS)[4] and the program WISCT (commercially available[5]) written at Pomona College to verify WLNs and generate connection tables. WISCT is also used to generate connection tables for the MACCS system.[6]

There are two WLN-string-search programs on the market, RADIICAL[7] from ISI (*R*etrieval and *A*utomatic *D*issemination of *I*nformation from the *I*ndex *C*hemicus and *L*ine Notation) and CROSSBOW. Reckitt and Colman in England have purchased certain CROSSBOW programs but use the IBM package INQUIRE to search notations, molecular formulas, and CROSSBOW fragments.

Many American companies, and Pfizer in England, have in-house software for checking and searching WLNs and for generating connection tables (from WLN) to be used in atom-by-atom search and structure display. On the connection table front, CROSSBOW,[3] WISCT,[5] and MACCS[6] are for sale. Two software houses, CIMI[8] and Fraser-Williams (Scientific Systems) Ltd.,[3] specialize in writing WLN systems for customers. Granito of CIMI has written a program to generate ring codes from WLN[9,10] (i.e., the code of the Pharma Documentation Ring).

Finally, this brief survey exemplifying (but not comprehensively covering) WLN software would not be complete without mentioning the PATHFINDER program originally written by the Dow Chemical Company but now available from Chemical Abstracts Service[11] (or in England via the United Kingdom Chemical Information Service—UKCIS). The program is used to check locant paths in ring systems with bridges and/or crossed bonds. It could, of course, be used for any ring system but it is of most use for "tricky" notations where the encoder is not sure whether he has picked the optimum path for lettering ring atoms.

From software we move to examples of the various data bases that have been created by using WLN. The biggest is ICRS.[4] ISI supply very up-to-date tapes, on a monthly basis, of novel compounds appearing in the literature, together with bibliographic data and various flags and subject terms. they

0095-2338/82/1622-0098$01.25/0 © 1982 American Chemical Society