Searching the *Chemical Abstracts Condensates* Data Base via Two On-Line Systems†

BARBARA G. PREWITT

Rohm and Haas Company, Research Division, Spring House, Pennsylvania 19477.

Received April 7, 1975

A comparison of the most used features of System Development Corporation's ORBIT and Lockheed's DIALOG systems is made, especially in reference to conducting searches of *Chemical Abstracts Condensates*. Many of the operations are similar in nature. However the capabilities are sufficiently different that an experienced searcher can select the system which gives the best results.

The Rohm and Haas Company's Research Division Library has been involved in on-line searching for over 2 years. Our initial experiences were with Lehigh University's LEADERMART System and System Development Corporation's ORBIT system. Shortly thereafter we signed up with Lockheed's DIALOG system. Our searching has been primarily with the *Chemical Abstracts Condensates* (Condensates), although we have made fairly extensive use of other bibliographic files such as NTIS, CAIN, Medline, Toxline, Compendex, Predicast's Chemical Market Abstracts, etc. This discussion will be limited to searching Condensates on-line using either ORBIT or DIALOG. It is intended for persons with some experience on one or both systems. The short forms and symbols which normally are used by an experienced searcher are shown in the examples. ORBIT provides three versions of standard program messages for searching, and I use the most abbreviated one, the symbolic form. Tables I and II list the symbols and abbreviations found throughout this paper.

Before beginning the comparison of these two software packages, I would like to stress the differences between the volume Subject Indexes to *Chemical Abstracts* and the Condensates keyword index. The keyword index appears in the back of the weekly issues of *Chemical Abstracts,* uses uncontrolled vocabulary, and must be searched using synonyms, abbreviations, various word endings, etc. Chemical nomenclature in the Condensates keyword index does not contain positional indicators. Thus, 3,4-dichlorobenzene would appear in Condensates as dichlorobenzene or possibly chlorobenzene. Condensates does not adhere to systematic IUPAC nomenclature. The volume Subject Index uses a controlled vocabulary for its main headings and systematic nomenclature for chemical compounds.

One other point to be made is the necessity to check the printed volume Subject Indexes of *Chemical Abstracts* in addition to the results of an on-line Condensates search if a complete search is to be made. Often references located in the printed Subject Index are not found in the on-line search of Condensates. This is due primarily to the deeper and more complete indexing done for the Subject Index or to the availability of the empirical formula index. However, owing to the ability to search on fields not available in the printed indexes, the on-line search may turn up references not located in a manual search of the Subject Index.

SEARCHABLE FIELDS

The first comparison is of the fields which are searchable by DIALOG and ORBIT. It is apparent from Table III that in both DIALOG and ORBIT the most frequently searched fields, such as title words, keywords, author, patent assignee, language, etc., are accessible directly. ORBIT's lack of directly searchable fields for the abbreviated journal title, country of granted patents, and publisher has caused us no hardship, and we have seldom, if ever, searched on these fields. The ability to search the publication date field in ORBIT for the year of publication has been used frequently to limit our searches.

Condensates from 1972 to date is available during the entire working day at Lockheed. At SDC the corresponding file is available in windows 1 (8:00 AM to 11:00 AM), 2 (11:10 AM to 2:00 PM), and 4 (5:10 PM to 8:00 PM). Times are Eastern Standard Time. During window 2 the 1970–71 Condensates file is available, too. Thus searches must be scheduled at SDC during the appropriate window, while they can be run anytime during the working day at Lockheed. We have found the morning hours to be best for searching. Response time appears to be better then. The 1970–71 file is used infrequently, since it is contained in the Eighth Collective Index. Our searches of this file have been restricted to those which cannot be performed manually or which require a large amount of time to carry out manually.

SUBJECT SEARCHING

Next I would like to compare subject searching via the Lockheed and SDC systems. In DIALOG single words from the title or keyword fields may be searched using the SELECT command. Word proximity searches are available in the title field only, since the keyword field is not full text indexed. Multiple occurrences of a word in the keyword field are not stored by DIALOG. However, it is possible to specify the occurrence of two or more terms in the same title or keyword field by typing SELECT SULFUR(F)DIOXIDE. This will ensure that both terms are present in the keyword or title field, but not the word order. Thus, this input might retrieve a reference containing keyword phrases such as the following:

SULFUR AND NITROGEN PLANT METABOLISM
CARBON DIOXIDE PLANT RESPIRATION

When searching titles via DIALOG, word order can be specified by SELECTing as follows:

SELECT INSECTICIDE(W)PREPARATION/TI

The references retrieved will have the word insecticide immediately preceding the word preparation in the title. However, an article entitled, "The Preparation of an Insecticide for Cockroaches", will not be retrieved because the

† Presented in the Symposium on "User Reactions to CAS Data and Bibliographic Services," 169th National Meeting of the American Chemical Society, Philadelphia, Pa., April 7, 1975.

Table I. DIALOG Commands and Abbreviations

Commands		Abbreviations	
#	SELECT	AU	Author
$	COMBINE	CS	Corporate source
*	AND	LA	Language
+	OR	NPT	Nonpatent
—	AND NOT	PA	Patent assignee
'	TYPE	SC	*Chemical Abstracts* section number
)	LIMIT		
"	EXPAND		
?	TRUNCATION		
=	END		

Table II. Orbit Abbreviations

AU	Author
CA	*Chemical Abstracts* section number
CONT?	Continue?
CP	*Chemical Abstracts* reference
CS	Corporate source
D	Down
GEN	General
IT	Index term
KW	Keyword
LA	Language
MM	Multimeaning
N	No
NBR	Neighbor
NP	No postings
OVFLW	Overflow
PA	Patent assignee
PN	Patent number
PROG	Program
PRT	Print
PY	Publication year
QUAL	Qualify
SCHD	Searched
SO	Source
SS 1/C	Search Statement 1 or Command?
SS 1 PSTG	Search Statement 1—Number of postings
STRS	Stringsearch
TI	Title
UC	Publication
WL	Work location
Y	Yes

Table III. Fields Searchable by DIALOG or ORBIT[a]

Field	DIALOG	ORBIT
Title	X	X
Keyword phrase	X	X
Author	X	X
Abbreviated journal title	X	—
Coden	X	X
Publication date	—	X
Patent assignee	X	X
Patent number	X	X
Country of granted patents	X	—
Patent classification	X	X
Country of patent application or priority	X	X
Language	A	X
CA publication section	X	X
CA publication citation	X	X
Corporate author	X	X
Work address	X	X
Work location	X	B
Publisher	X	—
Publication class	C	D

[a] A—Directly searchable only for specific foreign languages; can be limited to foreign or English. B—Company name only; not division and city. C—Can be limited to patents or nonpatents. D—Directly searchable only for patents.

words are not present in the order specified nor are they adjacent. Therefore, all possibilities must be considered before searching. A strategy would be to search as follows: SELECT INSECTICIDE(3W)PREPARATION; SELECT PREPARATION(3W)INSECTICIDE. Using this strategy, all titles or keyword phrases would be found that have the words insecticide and preparation within three words of each other. This, of course, does not consider the plural, insecticides, nor synonyms and other forms of the word preparation. For a thorough search all possible word combinations must be utilized in the strategy.

```
                                        Set  Items  Description

? #TLC;#THIN;#LAYER;#CHROMATOG?;#PESTICIDE?

          1      75   TLC
          2    6731   THIN
          3    6865   LAYER
          4   16121   CHROMATOG?
          5    3373   PESTICIDE?

? $5*(1+(2*3*4))

          6      98   5*(1+(2*3*4))
```

Figure 1. DIALOG sample search for thin layer chromatography of pesticides.

```
SS  1 /C?
USER:
THIN AND LAYER AND ALL CHROMATOG: OR TLC

PROG:
SS  1 PSTG (2586)

SS  2 /C?
USER:
1 AND ALL PESTICIDE#

PROG:
SS  2 PSTG (93)
```

Figure 2. ORBIT sample search for thin layer chromatography of pesticides.

```
SS  3 /C?
USER:
SULFUR AND DIOXIDE OR CARBON AND DIOXIDE

PROG:
SS  3 PSTG (8220)

SS  4 /C?
USER:
SULFUR OR CARBON

PROG:
SS  4 PSTG (38012)

SS  5 /C?
USER:
4 AND DIOXIDE

PROG:
SS  5 PSTG (8220)
```

Figure 3. ORBIT search logic for sulfur dioxide or carbon dioxide.

In DIALOG after having SELECTed the appropriate single words or word phrases, the COMBINE command is used to connect them with the Boolean operators AND, OR, and AND NOT. Complex logic statements can be built with the use of parentheses (Figure 1).

When using ORBIT the search terms and the appropriate Boolean operators can be entered in one search statement (Figure 2). It is not necessary to first select each term and then combine them. However, complex logic statements cannot be built in one search statement because the operator AND takes precedence over OR and parentheses are not permitted. If one is searching via ORBIT for sulfur dioxide or carbon dioxide, the search should be typed as shown in Figure 3. If the search statement were typed SULFUR OR CARBON AND DIOXIDE, the computer program would interpret this to mean all references on carbon dioxide or sulfur. ORBIT does not consider word order or word proximity in the normal searching mode. ORBIT's STRINGSEARCH, which is described below, does permit searching for words in a given order in specific fields.

In addition to the capability of searching directly for words or numbers, ORBIT has the ability to search for character strings. This is called STRINGSEARCH or

```
USER:
ALL INSECTICIDE# OR ALL ACARICIDE# OR ALL MITICIDE#

PROG:
SS  6 PSTG (6247)
SS  7 /C?
USER:
STRS (TI) :PHOSPHO: OR :PHOSPHO: (KW)
PROG:
(130) SCHD (25) QUAL; CONT? (Y/N)
USER:
Y
PROG:
(278) SCHD (56) QUAL; CONT? (Y/N)
USER:
Y
PROG:
(442) SCHD (78) QUAL; CONT? (Y/N)
USER:
N
PROG:
SS  7 PSTG (78)
SS  8 /C?
USER:
"PRT 2, TI,KW,CP,INDENTED"
PROG:

TITLE OF DOCUMENT              METABOLISM OF PHOSPHORIC ACID ESTER AMIDES
KEYWORD PHRASES                AMINOPHOSPHATE INSECTICIDE
KEYWORD PHRASES                DICHLOROVINYL DIMETHYL  AMINO PHOSPHATE
                               INSECTICIDE
KEYWORD PHRASES                PHOSPHOROAMIDATE INSECTICIDE
CA PUBLICATION CITATION        CA076150821767

TITLE OF DOCUMENT              CHOLINESTERASE INHIBITION AND TOXICITY OF
                               VARIOUS ORGANOPHOSPHORUS INSECTICIDES TO
                               HIBERNATING CHILO SUPPRESSALIS  LARVAE
KEYWORD PHRASES                ORGANOPHOSPHATE INSECTICIDE CHILO
KEYWORD PHRASES                INSECTICIDE RICE BORER CHOLINESTERASE
KEYWORD PHRASES                CHOLINESTERASE CHILO INSECTICIDE
KEYWORD PHRASES                ANTICHOLINESTERASE INSECTICIDE RICE BORER
CA PUBLICATION CITATION        CA076150821467
```

Figure 4. Sample of ORBIT STRINGSEARCH for phosphorus-containing insecticides.

STRS for short. Once a group of references have been found, for instance by a keyword or title word search, it is possible to sequentially search any field in these records for character strings. For example, in a search on phosphorus-containing insecticides, one would search for references with the words insecticide, insecticides, acaricide, acaricides, miticide, etc., in the title or keyword fields. Then these references would be stringsearched for the characters :PHOSPHO: embedded in the associated keywords or title words (Figure 4).

There is no equivalent to STRINGSEARCH in DIALOG. STRINGSEARCH has made possible substructure searching on a set of references, such as illustrated in the preceding paragraph. Other applications of STRINGSEARCH have been found, such as enabling a search to be completed that would have created an overflow condition in the computer storage area if attempted by direct searching, e.g., all CA volume 81 references. STRINGSEARCH is a valuable and powerful tool, but it is slower than direct searching and requires many yes responses to the continuation message if the set being processed contains over 500 references.

It should be mentioned that STRINGSEARCH can be used to search for character strings in more than one word. Thus, it is possible to search for :PHENYL ETHER: or for :INSECTICIDE:PREPARATION:. The embedded colon here indicates the presence of from 0 to 7100 characters. Also available is the capability to search for several strings in more than one field and to use Boolean operators in the same searching operation (Figure 5).

Both DIALOG and ORBIT are capable of displaying alphabetically adjacent terms from their index. DIALOG uses the EXPAND command which displays up to 51 terms, 20 lines at a time (Figure 6). Additional terms beyond the 51st may be displayed, but one must SELECT appropriate terms before going beyond term E51 since DIALOG then begins renumbering with E1. ORBIT's

```
SS 1 /C?
USER:
FUNGICIDE OR FUNGICIDES OR FUNGAL
PROG:
SS 1 PSTG (4334)
SS 2 /C?
USER:
STRS (KW) :CARBAMODITHIO: OR :CARBAMODITHIO: (TI) OR :DITHIOCARBAM: OR
PROG:
CNT 2
USER:
:DITHIOCARBAM: (TI)
PROG:
(90) SCHD (3) QUAL; CONT? (Y/N)
USER:
Y
PROG:
(182) SCHD (7) QUAL; CONT? (Y/N)
USER:
N
PROG:
SS 2 PSTG (7)
SS 3 /C?
USER:
"PRT  TI,AU,SO,PY,PN,PA,KW,CP"
PROG:

TI - GERMICIDE COMPOSITION FOR SILK WORMS
AU - IMANISHI, KOSAKU; YAMANO, TOGO
SO - JAPAN. 4 PP.
PY - 18 AUG 71
PN - 71 28429
PA - TAKEDA CHEMICAL INDUSTRIES, LTD.
KW - ETHYLENEBISDITHIOCARBAMATE FUNGICIDE SILKWORM
KW - SALICYLATE SILKWORM FUNGICIDE
CP - CA076050219597D
```

Figure 5. ORBIT substructure search utilizing STRINGSEARCH for dithiocarbamate fungicides.

```
? "ACRYL

REF     INDEX-TERM                    TYPE ITEMS    RT

E1      ACROZONASE-------------------        1     0
E2      ACRYAMIDE--------------------        2     0
E3      ACRYATE----------------------        1     0
E4      ACRYDANO---------------------        1     0
E5      ACRYIC-----------------------        2     0
E6      ACRYL------------------------       29     0
E7      ACRYLACETATE-----------------        2     0
E8      ACRYLALDEHYDE----------------        8     0
E9      ACRYLAMAIDE------------------        1     0
E10     ACRYLAMIDASE-----------------        2     0
E11     ACRYLAMIDE-------------------     1207     0
E12     ACRYLAMIDES------------------       31     0
E13     ACRYLAMIDESS-----------------        1     0
E14     ACRYLAMIDINE-----------------        2     0
E15     ACRYLAMIDINES----------------        2     0
E16     ACRYLAMIDIZATION-------------        1     0
E17     ACRYLAMIDO-------------------       23     0
E18     ACRYLAMIDOANTHRAQUINONE------        1     0
E19     ACRYLAMIDOBUTANE-------------        1     0
E20     ACRYLAMIDOCARBAMATE----------        1     0
                                                -MORE-

? #E9,E11-E13

          1    1225   E9,E11-E13
          E6:  ACRYL
```

Figure 6. Sample of DIALOG alphabetical term display of acryl.

NEIGHBOR command will display up to 10 terms initially and will permit the user (Figure 7) to browse additional terms at either end of the display. Both systems list the term and the number of times it occurs in the file. DIALOG allows the searcher to SELECT terms directly from the display. If more than one is chosen, the terms are ORed together in one set. All 51 terms from an EXPAND display may be ORed together, if desired. ORBIT requires the searcher to retype the appropriate terms in a search statement.

TRUNCATION FEATURES

Another feature common to both systems is truncation. DIALOG performs term truncation by using a ? after the last character of the stem. In DIALOG a truncated term will include up to the first 100 terms beginning with that

```
SS  8 /C?
USER:
"NBR ACRYL 10"

PROG:

POSTINGS      TERM
    17      ACRSAJ (JC)
     2      ACRYAMIDE  /  (IT)
     1      ACRYATE  /  (IT)
     1      ACRYDANO  /  (IT)
     2      ACRYIC  /  (IT)
    29      ACRYL  /  (IT)
     2      ACRYLACETATE  /  (IT)
     8      ACRYLALDEHYDE  /  (IT)
     1      ACRYLAMAIDE  /  (IT)
     2      ACRYLAMIDASE  /  (IT)
UP  N  OR  DOWN  N?

USER:
D 5

PROG:

POSTINGS      TERM
  1216      ACRYLAMIDE  /  (IT)
    31      ACRYLAMIDES  /  (IT)
     1      ACRYLAMIDESS  /  (IT)
     2      ACRYLAMIDINE  /  (IT)
     2      ACRYLAMIDINES  /  (IT)
UP  N  OR  DOWN  N?
0

SS 8/C?
USER:
ACRYLAMAIDE OR ALL ACRYLAMIDE:

PROG:
SS 8 PSTG (1225)
```

Figure 7. Sample of ORBIT alphabetical term display of acryl.

```
SS  3 /C?
USER:
ALL BENZISOTHIAZOL:

PROG:
SS 3 PSTG (139)

PROG:

SS 4 /C?
USER:
ALL POLY:

PROG:
GEN TERM OVFLW

PROG:

SS 4 /C?
USER:
BENZISOTHIAZOLIN:

PROG:
MM (BENZISOTHIAZOLIN:) (9)
          1    BENZISOTHIAZOLIN (IT)
          2    BENZISOTHIAZOLINE (IT)
          3    BENZISOTHIAZOLINES (IT)
          4    BENZISOTHIAZOLINONE (IT)
          5    BENZISOTHIAZOLINONECARBOXYLATE (IT)
          6    BENZISOTHIAZOLINONES (IT)
          7    BENZISOTHIAZOLINYL (IT)
          8    BENZISOTHIAZOLINYLIDENE (IT)
          9    BENZISOTHIAZOLINYLIDENECYANOACETA (IT)
SPECIFY NUMBERS, ALL, OR, NONE-

USER:
ALL

PROG:
SS 4 PSTG (139)
```

Figure 8. ORBIT truncation samples.

character stem. To ensure that no terms are missed, the DIALOG searcher is advised not to use truncation, but to EXPAND on the stem and select the appropriate terms from the display.

Truncation in ORBIT is performed using a colon after the last character of the stem. All terms beginning with the character string may be retrieved if there are not more than 400 terms (Figure 8). If the truncated term is not preceded by the word ALL, up to ten terms will be displayed auto-

```
SS 4 /C?
USER:
ENGLISH AND CLAY OR ENG AND CLAY

PROG:
MM (ENG) (2)
          1  ENG (IT)
          2  ENG (LA)
SPECIFY NUMBERS, ALL, OR, NONE-

USER:
1

PROG:
SS 4 PSTG (1)
```

Figure 9. ORBIT multimeaning message.

ORBIT

```
SS 5 /C?
USER:
TOXICITY OR ALL CA004: (CA)

PROG:
SS 5 PSTG (20670)

SS 6 /C?
USER:
ENG (LA) AND ALL HERBICIDE# AND REVIEW

PROG:
TIME OVFLW: CONT? (Y/N)

USER:
Y

PROG:
SS 6 PSTG (136)

SS 7 /C?
USER:
PREWITT BARBARA G.

PROG:
SS 7 PSTG (1)
```

DIALOG

```
? #LA=ENG
          1    619   LA=ENG
? #SC=CA004
          2  20721   SC=CA004
? #LA=GER
          3  45761   LA=GER
? #AU=PREWITT, BARBARA G.
          4      1   AU=PREWITT, BARBARA G.
```

Figure 10. Specification of search fields.

matically. The appropriate terms are selected by number from the display. If more than ten terms exist, the user is given only the option to select all or none of the terms. ORBIT also uses the hash mark (#) to represent a single variable character or space. POLYMER# will retrieve POLYMER or POLYMERS, but not POLYMERIZE.

Truncation should be used with care since unintended terms may be retrieved. The stem POLYMER will include the word POLYMERCAPTAN, a term probably not desired. For this reason, it is often preferable to display an alphabetical list of terms and select only those that are appropriate. However, when one is reasonably sure of the results, it is quicker and easier to use the truncation feature. This can be achieved by consulting the SDC microfiche of the Chemcon index to Condensates prior to going on-line.

The truncation feature in ORBIT can be used to avoid having to retype the search terms obtained in a NEIGHBOR display as mentioned above (Figure 7). Typing BENZISOTHIAZOLIN: will cause up to ten terms with this stem to be displayed automatically via the MULTIMEANING feature. At present more than ten terms cannot be displayed, an obvious shortcoming. Another limitation of MULTIMEANING is that it does not list the number of

```
? #ROHM(F)HAAS(F)PA/CS

       16      101      ROHM(F)HAAS(F)PA/CS
? ' 16/5/1-2

1
CA082040213751
   DECOLORIZING DYE WASTES
   AUTHOR: KENNEDY, DAVID C., STEVENS, BRUCE, KERNER, JEFFREY W.
   LOCATION: ROHM AND HAAS CO., PHILADELPHIA, PA.
   SECTION: CA060002, CA040000 PUBL.-CLASS: J   COVERAGE:  1
   JOURNAL: AM. DYEST. REP.   CODEN: ADREAI   PUBL: 74   SERIES: 63
ISSUE: 8 PAGES: 11-14
   DESCRIPTORS:  DYE WASTEWATER DECOLORIZATION
2
CA081251643450
   METABOLISM OF THE HERBICIDE PRONAMIDE IN SOIL
   AUTHOR:  FISHER, JAMES D.
   LOCATION:  RES. LAB., ROHM AND HAAS CO., SPRING HOUSE, PA.
   SECTION:  CA004004   PUBL.-CLASS:  J   COVERAGE:  1
   JOURNAL: J. AGR. FOOD CHEM.   CODEN:  JAFCAU   PUBL: 74  SERIES:
22    ISSUE: 4  PAGES: 606-8
   DESCRIPTORS:  PRONAMIDE METABOLISM SOIL, HERBICIDE PRONAMIDE
METABOLISM SILO
```

Figure 11. DIALOG corporate source search for Rohm and Haas Company Pennsylvania locations.

```
SS 8/C?
USER:
(WL)ALL PENNA# OR ALL PA# OR PENNSYLVANIA

PROG:
NP (PENNA# (WL))
NP (PENNSYLVANIA (WL))
SS 8 PSTG (14588)

SS 9/C?
USER:
8 AND ROHM/CS AND HAAS/CS

PROG:
SS 9 PSTG (114)
```

Figure 12. ORBIT corporate source search for Rohm and Haas Company Pennsylvania locations.

times a term occurs in the Condensates index file. This number can be helpful in deciding the significance of a particular term to the search at hand.

It should be mentioned that ORBIT's MULTIMEANING feature has another use. This is to distinguish between terms as index terms and as, for example, a language (Figure 9). Here the subject is ENGLISH CLAY not CLAY and ENGLISH language. DIALOG requires the use of a two-letter field code in a SELECT command if one is searching on a field other than title words and keywords. The searcher also can specify the field to be searched in ORBIT by typing in the two letter field code and avoiding a MULTIMEANING message (Figure 10).

CORPORATE SOURCE SEARCHES

A searchable corporate source field exists in both DIALOG and ORBIT. In DIALOG the corporate source field, which is full text indexed, contains information on the corporate author, work address (company name, division, city), and work location (country, possession, province, or state). These terms are searched with a /CS following the term (Figure 11). In ORBIT the corporate source field contains corporate author, patent assignee, and company name from the work address field. The state and country elements are searchable in the separate work location field. City is not searchable in ORBIT. The terms in the corporate source are entered for searching with a /CS following the term (Figure 12).

When searching for a patent assignee in DIALOG, the patent assignee field should be EXPANDed for the company name (Figure 13). This is necessary to ensure that the correct spelling and punctuation are used. If the search terms as entered do not match the index terms exactly, a zero hits message will result. It is then possible to EXPAND this term by simply entering a quote sign and display alphabetically adjacent terms. In ORBIT a patent assignee search can be performed by entering one or more significant words from the company name and intersecting this with patents (Figure 14). Thus, it is not necessary to know the exact form of the company name to do a patent assignee search in ORBIT.

```
? "PA=ROHM AND HAAS

REF    INDEX-TERM          TYPE TERMS   RT
E1     PA=ROFAMEL A.-G.---------   1    0
E2     PA=ROGERS BROTHERS CO.---   1    0
E3     PA=ROGERS CORP.----------   1    0
E4     PA=ROGOX A.-G.-----------   1    0
E5     PA=ROHE SCIENTIFIC CORP.-   2    0
E6    -PA=ROHM AND HAAS---------   0    0
E7     PA=ROHM AND HAAS CO.------ 189    0
E8     PA=ROHM G.M.B.H.---------   2    0
E9     PA=ROHM AND HAAS
          G.M.B.H.--------------   1    0
E10    PA=ROHNE-PROGIL----------   1    0
?  #E7;'17/5/1-2

         17  189  PA= ROHM AND HAAS CO.
1
CA082040216831
   WATER TREATMENT
   AUTHOR:  KUNIN, ROBERT
   SECTION: CA061005, CA037000  PUBL.-CLASS: P  COVERAGE: 1
   JOURNAL: GER.  OFFEN.   CODEN: GWXXBX  PUBL: 740919  PAGES: 16
PP.
   DESCRIPTORS: RESIN  ION  EXCHANGE  WATER,  METHACRYLIC ACID EXCHANGE
RESIN, DIVINYLBENZENE EXCHANGE RESIN
   PATENT-NO: 2409261  APPLIC-NO: 340,503  DATE: 730312  CLASS: C
02B  COUNTRY:  US
   ASSIGNEE:  ROHM AND HAAS CO.

2
CA082040186481
   NITROCELLULOSE-MODIFIED URETHANE COATING COMPOSITIONS AND THEIR USE
IN FINISHING LEATHER
   AUTHOR:  LEWIS, SHELDON N., YUNASKA, MATTHEW R.
   SECTION: CA041004  PUBL.-CLASS: P  COVERAGE: 1
   JOURNAL: U.S.   CODEN: USXXAM  PUBL: 740611  PAGES: 5 PP.
DIVISION OF U.S. 3,763,061 (CA 80;16473U).
   DESCRIPTORS:  NITROCELLULOSE POLYURETHANE COATING LEATHER
   PATENT-NO: 3816168  APPLIC-NO: 161,987  DATE: 710712  CLASS:
117/142, B 44D
   ASSIGNEE:  ROHM AND HAAS CO.
```

Figure 13. DIALOG patent assignee search for Rohm and Haas Company.

```
SS 10 /C?
USER:
ROHM/CS AND HAAS/CS AND P (UC)

PROG:
SS 10 PSTG (192)

SS 11 /C?
USER:
"PRT 2,PATENT,INDENTED"

PROG:

TITLE OF DOCUMENT           STABLE, POURABLE ALKALI METAL DITHIONITE
                            DISPERSION
AUTHOR NAME (PERSONAL)      OWEN, ROBERT M.; RAMM, HENRY L.
PATENT ASSIGNEES            ROHM AND HAAS CO.
CA PUBLICATION CITATION     CA082060328751
SOURCE REFERENCE            U.S. 5 PP.
PATENT NUMBER               3839217
PUBLICATION DATE            01 OCT 74
PATENT DATE                 090170
PATENT CLASSIFICATION       252/188; C 01B, D 21C
PATENT PRIORITY NUMBER      1,835

TITLE OF DOCUMENT           ALLYL METHACRYLATE POLYMERS
AUTHOR NAME (PERSONAL)      LEWIS, SHELDON NOAH; MILLER, JOHN JOSEPH
PATENT ASSIGNEES            ROHM AND HAAS CO.
CA PUBLICATION CITATION     CA082060317365
SOURCE REFERENCE            GER. OFFEN. 22 PP.
PATENT LOCATION             US
PATENT NUMBER               2357615
PUBLICATION DATE            12 JUN 74
PATENT DATE                 061272
PATENT CLASSIFICATION       C 08F
PATENT PRIORITY NUMBER      312,433
```

Figure 14. ORBIT patent assignee search for Rohm and Haas Company.

LIMITING OUTPUT

The results of a search in *Chemical Abstracts Condensates* may be limited in both DIALOG and ORBIT to lan-

```
? #ISOTHIAZOL?
          4    131 ISOTHIAZOL?

? )4/NPT
          5     81 4/NPT

? )4/ENG
          6     87 4/ENG

? )4/80000001-82000222
          7     48 4/80000001-82000222
```

Figure 15. Use of DIALOG limit command for nonpatents, language, and accession number range.

```
SS 3 /C?                    SS 6 /C?
USER:                       USER:
ALL ISOTHIAZOL:             3 AND GREATER THAN 73

PROG:                       PROG:
SS 3 PSTG (130)             SS 6 PSTG (30)

SS 4 /C?                    SS 7 /C?
USER:                       USER:
3 AND NOT P (UC)            3 AND LESS THAN 71

PROG:                       PROG:
SS 4 PSTG (80)              SS 7 PSTG (2)

SS 5 /C?                    SS 8 /C?
USER:                       USER:
3 AND GER (LA)              3 AND FROM 74 THRU 75

PROG:                       PROG:
SS 5 PSTG (7)               SS 8 PSTG (30)
```

Figure 16. Restricting ORBIT results to patents, language or year.

guage, patents, nonpatents, and time period. DIALOG uses a LIMIT command for these operations (Figure 15). ORBIT performs these operations by direct searching (Figure 16).

ERASING STORED SETS

DIALOG can accommodate up to 98 sets if the number of items stored does not exceed the capacity of about 400,000 records. ORBIT can store up to 25 search statements, each of which can contain a number of search terms and the associated logic. ORBIT's storage capacity is 80,000 to 200,000 records depending on what the system is doing at the time. The ORBIT searcher can selectively erase previous search statements; the DIALOG searcher cannot. Thus, it is possible to circumvent storage overflow conditions at ORBIT with the RESTACK or ERASEBAK commands. Both DIALOG and ORBIT permit the searcher to erase everything and begin anew. When switching data bases, for example, Condensates to Compendex, DIALOG retains previous sets unless a BEGIN or BEGIN B (for bypass) command is entered. ORBIT automatically erases all previous search statements when switching data bases. At times it is convenient to be able to return in DIALOG to a previously used data base and revise the search strategy with new terms obtained in another data base without reentering the earlier search terms. However, this can be a problem if many of the search terms are heavily posted and cause an overflow condition to occur subsequently. This does not happen frequently with DIALOG.

SEARCH INTERRUPTION AND ERROR CORRECTION

DIALOG can be interrupted by pressing the break key. ORBIT cannot be stopped once a command or search has been transmitted. Often it is convenient to interrupt a display after examining only the first few items. For example, it can be frustrating to wait while the printer continues to list nonpertinent references or to display author names once the one you seek is printed. Also, the break key is useful for interrupting the introductory message from DIALOG.

Typographical errors can be corrected on both DIALOG and ORBIT. Single characters or whole lines can be erased before they are transmitted. The methods of doing this vary with the type of terminal being used.

STORING SEARCHES

DIALOG also has the ability to store a profile or a list of synonyms which are used frequently and require considerable typing. This is possible by using the SEARCH SAVE feature described in the manual. SEARCH SAVE is available in all files at Lockheed. It should be mentioned that recalling a stored search has a lower priority than the other searching commands and is executed more slowly than other operations. Also the EXPAND command should not be used in SEARCH SAVE. This capability is useful, too, for updating searches at longer intervals than the normal weekly or biweekly intervals for Condensates SDI profiles.

PRINTING

To print references obtained via DIALOG, one enters either a print on-line command (TYPE or DISPLAY) or a print off-line command (PRINT). The set number, the number of the desired format, and the range of references are entered. The user cannot format the output to suit his needs, but must use one of the standard formats provided for Condensates. The address and labeling information if required, is entered during the BEGIN operation. In any case, all off-line prints are mailed to the appointed contact person whether they are labeled or not. The pricing of off-line prints varies according to the format, and lower prices are charged for partial output.

The ORBIT user may print on-line or off-line and may use either fixed or user-formatted printouts. There is a limit of 1000 references on off-line prints. The capability of printing only the desired fields is very convenient, especially when browsing results on-line to determine pertinency. ORBIT's ability to store a mailing address using the STORAD command is also convenient. Only title and requester information need be entered for an off-line print and the printout is sent to the address stored. If the STORAD option is not used, the mailing address must be entered on-line. Off-line prints are priced the same regardless of the format used.

SEARCHING TECHNIQUES

The following are some searching techniques that we have found useful for *Chemical Abstracts Condensates*. Checking the SDC microfiche of index terms before going on-line gives the searcher a good idea of appropriate search terms and their frequency and often suggests additional terms. It is very helpful for determining how chemical names are entered. For example, polyvinyl chloride is also found as polyvinylchloride and PVC. Another related technique is to enter the *Chemical Abstracts* reference number to determine how a known pertinent article was indexed. Still another invaluable tool for search term generation is the Index Guide, both the current and 8th collective versions. Recently Chemical Abstracts Service has made available a User Aid Package to assist searches of *CA Condensates*.

We have been experimenting with the use of patent classification numbers, both U.S. and international classes. These provide an additional handle for retrieval. However, Condensates does not include the group and subgroups of the International Patent Classification (IPC). As a result hundreds or thousands of references are retrieved under the searchable portion of the IPC, making it a less precise entry point than desired. If the whole IPC number were

searchable, more precision would be possible and this mode of searching would be more valuable. The complete U.S. classification number is given for U.S. patents. However, since U.S. patents issue more slowly than their foreign equivalents, relatively few U.S. patents are abstracted in *Chemical Abstracts*. Therefore, a search of U.S. class numbers normally results in the retrieval of only a small fraction of the pertinent patents covered in *Chemical Abstracts*. It should be mentioned that the Condensates record also includes the patent classification used on Japanese Kokai.

The availability of the *Chemical Abstracts Patent Concordance* in an on-line system would provide quicker and better access to equivalent patents. Currently a manual search must be conducted through the volume and weekly patent concordances. Often an on-line search will not retrieve a previously known patent, and the only way to determine if an equivalent patent was indexed is by a manual search.

Another technique that I have found useful is to save a copy of the news and update information that is printed on-line. This allows me to update searches more easily and to answer questions about search coverage from requesters at a later date.

Often it is advantageous to search on the *Chemical Abstracts'* section number, as well as the appropriate index terms. I have found many useful references in this manner which were not subject-indexed to the desired concept. An example would be a search on Section 5 for articles on structure–activity studies for a question about pesticidal structure–activity relationships.

We are continuing to use DIALOG and ORBIT to search Condensates and the other available data bases. Our experience shows that both systems provide good service and are reliable. It is my opinion that ORBIT and DIALOG are valuable services and that all organizations providing computerized bibliographic retrieval service should utilize both.

Computerized Chemical Information Retrieval Techniques†

DANIEL U. WILDE

New England Research Application Center, University of Connecticut, Storrs, Connecticut

Received April 3, 1975

The introduction of powerful, low-cost computers now make it possible for most organizations to perform SDI and retrospective searching. The use of such a machine, located at the New England Research Application Center (NERAC) was reported earlier. New improvements to NERAC's SDI Chemical Search System are now discussed.

INTRODUCTION

For the past eight years, the New England Research Application Center (NERAC) has operated as a NASA Industrial Application Center at the University of Connecticut. The Center aids and promotes technology transfer by helping business and industry, colleges and universities, and state and local governments locate appropriate scientific and technical information. During this period, NERAC has performed some 8000 retrospective searches, while its data base has grown to over 2,250,000 documents. This data base now includes retrospective searching of nine files: National Aeronautics and Space Administration (NASA), Engineering Index (COMPENDEX), Department of Defense (DDC), Government Reports Abstracts (NTIS), Education Resources Information Center (ERIC), American Society for Metals (METADEX), Alloys Index (AI), World Aluminum Abstracts (WAA), and Abstracted Business Information (INFORM). Inspec and *CA Condensates* are also available for current awareness or Selective Dissemination of Information (SDI) bringing to eleven the total number of files available for SDI service.

When NERAC was established there were two choices for computational power. Time could be purchased on a large, extensive computer system, or a small machine could be rented for the exclusive use of the center. As was described earlier, NERAC decided to rent its own machine.[1,2] As a result, computer costs were fixed and increased computer usage decreased unit search costs. NERAC's present computer system consists of an IBM 370/115 with a 128K byte memory. Search questions are read via a 2501 card reader, and results are printed on a 5203 line printer. Data files are read from and intermediate results are saved on two 3410 tape drives with transmission rates of 80,000 characters per second at 1600 bits per inch. User programs and temporary data sets are saved on two 3340 disk drives with 140 million bytes of storage.

Random access searching of inverted files is not practical with such a limited machine. This is particularly true for large data bases, such as the NASA file with over 1,000,000 documents or Engineering Index with 400,000. Consequently, NERAC's data files are organized serially. This requires that all documents be examined individually in sequence during each search. Needless to say, search times can be expected to be especially long for character-by-character comparisons required for successful searching of *CA Condensates*.

Linear search methods are well known,[3] and the retrieval capabilities of NERAC's Chemical Search System have already been published.[4] This paper describes additional techniques for reducing linear search times and how they have been implemented on NERAC's small computer system without sacrificing retrieval capabilities.

† Presented before the Division of Chemical Literature in the Symposium on "User Reactions to CAS Data and Bibliographic Services," 169th National Meeting of the American Chemical Society, Philadelphia, Pa., April 7, 1975. This research was sponsored in part by the National Aeronautics and Space Administration, Contract NASW-2516.