

KGS
OF
68-4

Status: Open File
Date: October 24, 1968
By: Owen T. Spitz
Chief, Operations
Research

AN INFORMATION RETRIEVAL SYSTEM
For Sub-Surface Brine Analysis
Within the U. S. Bureau of Mines

by
Donald I. Good
and
Floyd W. Preston

The University of Kansas

in cooperation with
U. S. Bureau of Mines
Bartlesville, Oklahoma

April 1966

CONTENTS

	Page
Description of Data File.	1
Program MAINTAIN.	9
Program DUMP.	17
Program DICTIONARY.	19
Program FETCH	22

ACKNOWLEDGMENTS

The authors would like to acknowledge the assistance of the University of Kansas Computation Center for their use of the computer and their help in expediting the processing of the long data runs. The help of Mr. Marvin Donaldson is particularly appreciated. The encouragement of Dr. W. W. Hambleton is also appreciated. It was his foresightedness that led to the establishment of this cooperative project.

PURPOSES AND GENERAL DESCRIPTION OF PROJECT

PURPOSE

The purpose of this report is to describe the use of an information retrieval system for brine data developed for the U. S. Bureau of Mines. The computer program which constitutes the information retrieval system performs three functions.

- a. Creation of a magnetic tape file of all available brine records.
- b. Preparation of a printed paper output file of all records in the tape file, this paper output being in a very readable tabular form, one sheet per brine sample or a compressed form listing of data cards.
- c. Search and retrieval of any brine record or records conforming to certain categories desired and specified by the file user.

INTRODUCTION

Over the last 15 years the U. S. Bureau of Mines has been systematically gathering and analyzing subsurface brine from oil wells. The number of analyzed samples now exceeds 3000, with every indication that this number will continue to rise. The file was rapidly reaching the point at which it was difficult for one individual to search and assimilate the information. It was therefore suggested that it would be appropriate to store the information on a magnetic tape file and build a flexible retrieval system for the data. This report describes such a system and gives instructions for using it.

Objectives are accomplished by means of a package of four computer programs: MAINTAIN, DUMP, DICTIONARY, and FETCH. MAINTAIN is a general program for maintaining the magnetic tape file. It provides for a) original loading of the tape, b) adding new records to the tape, c) replacing old records with new records, and d) deleting records from the tape. DUMP allows for printing on paper selected records that are on the magnetic tape file. DICTIONARY is a program needed in the search routines. It creates dictionaries of six of the data entries (state, basin, county, formation, system and series) of each record. FETCH is the retrieval program for searching the file and writing prescribed information in tabular form.

The complete tape system is composed of two files. USBM-FILE is the file on which the brine data are stored and may, if necessary, involve more than one reel of magnetic tape. USMB-DICT is a single reel file or card deck which contains the indexed dictionary information generated by DICTIONARY and also certain other information that is necessary to transfer from one program to another.

PREPARATION OF THE BRINE DATA INPUT CARDS

In the initial phases of this project, the brine data were given to the keypunch operator on several different source documents. This was done because the data were already prepared on them at the start of the project and this move would avoid transcription of large amounts of data to

80 column coding forms. It is recommended that, in the future, the original source data be recorded on either coding forms or some other standard forms that are consistent with the punched card format.

A set of such forms is described in a later part of this report.

Throughout this report the word "record," when used with reference to the brine file, denotes the complete description of a single brine analysis of a specific sample. In other words, all of the information for one sample constitutes a single record. The word "field" denotes any one of the terms of information recorded in the brine analysis. In terms of the punched cards a field denotes a continuous sequence of columns containing a single item of information, as for example, state. Because of the large amount of information recorded for each sample, each record is composed of nine 80 column punched cards. The sample description must be located on these cards precisely as given below. The following general rules apply to preparation of data.

- a. If information for any field does not exist, it should be left blank or contain some other special symbol(s) such as a negative integer denoting that the information is not available or why it is not available.
- b. All entries should be right justified (right-most letter or number in right-most column of field) in their appropriate fields unless specifically stated otherwise.

Card	Column	Item
Card 1	1-22	State. Left justified.
	23-44	Basin. Left justified.
	45-66	Field. Left justified.
	68-72	The numeric part of the Sample No.
	73	The letter of the Sample No.
	75-79	File Sequence No.
		This quantity was devised as a reference number, purely for this filing system, and is not found on any of the original source documents. This number is a positive five digit integer that is assigned uniquely to each record being entered into the file. The file sequence number of any record in the file must be larger than the file sequence number of the preceding record. The file sequence number provides a means of uniquely "naming" each specific record.
	80	The digit "1" (Column 80 on each of the nine cards contains the card number within the specific record).
Card 2	1-22	County. Left justified.
	23-66	Location or Legal Description. Left justified. It should be noted in preparing data that the retrieval program, FETCH, prints only the first 22 columns of this field and that the program DUMP when using the open

file format writes the legal description on two lines breaking the description between the 18th and 19th characters.

75-79 File Sequence No.*
 80 The digit "2".

Card 3 1-22 Formation. Left justified.
 24-29 Top of Formation Depth (Ft.).
 31-36 Bottom of Formation Depth (Ft.).
 38-43 Elevation (ground level) (Ft.).
 45-46 Day Sample Obtained.
 47 / or blank
 48-49 Month Sample Obtained.
 50 / or blank.
 51-52 Last two digits of year sample obtained.
 54-59 Resistivity (Ohm-meters). Decimal must be punched in column 56.
 61-62 Day well completed, expressed as a number from 01 to 12.
 66 / or blank.
 67-68 Last two digits of year well completed.
 75-79 File Sequence No.*
 80 The digit "3".

Card 4 1-22 System. Left justified.
 23-44 Series. Left justified.
 46-51 Specific gravity. Decimal must be punched in column 48.

*The file sequence number must be the same for all nine cards.

53-58 Total dissolved solids, as a whole number.
60-65 Depth of well (Ft.), as a whole number.
67-72 pH. Decimal must be punched in column 70.
75-79 File Sequence No.*
80 The digit "4".

Card 5 1- 6 Lithium (mg/l)
8-13 Sodium (mg/l)
15-20 Potassium (mg/l)
22-27 Rubidium (mg/l) Decimal must be punched in
column 26.
29-34 Cesium (mg/l) Decimal must be punched in
column 33.
36-41 Calcium (mg/l)
43-48 Magnesium (mg/l)
50-55 Strontium (mg/l)
57-62 Barium (mg/l)
64-69 Strontium-Barium (mg/l)
75-79 File Sequence No.*
80 The digit "5".

Card 6 1- 6 Beryllium (mg/l)
8-13 Boron (mg/l)
15-20 Germanium (mg/l)
22-27 Copper (mg/l)
29-34 Chromium (mg/l)
36-41 Aluminum (mg/l)
43-48 Lead (mg/l)

*The file sequence number must be the same for all nine cards.

	50-55	Tin (mg/l)
	57-62	Silver (mg/l)
	64-69	Cobalt (mg/l)
	75-79	File Sequence No.*
	80	The digit "6".
Card 7	1- 6	Nickel (mg/l)
	8-13	Bismuth (mg/l)
	15-20	Vanadium (mg/l)
	22-27	Iron (mg/l)
	29-34	Manganese (mg/l)
	36-41	Chloride (mg/l)
	43-48	Bromide (mg/l)
	50-55	Iodide (mg/l)
	57-62	Bicarbonate (mg/l)
	64-69	Carbonate (mg/l)
	75-79	File Sequence No.*.
	80	The digit "7".
Card 8	1- 6	Sulfate (mg/l)
	8-13	Sulfide (mg/l)
	15-20	Nitrate (mg/l)
	22-27	Phosphate (mg/l)
	29-34	Organic Acid as Acetic (mg/l)
	36-41	Organic Carbon (mg/l)
	43-48	Organic Nitrogen (mg/l)
	50-55	Vanable Organic Nitrogen (mg/l)
	57-62	Stable Organic Nitrogen (mg/l)

*The file sequence number must be the same for all nine cards.

	64-69	Permanganate Oxidation as Ocygen (mg/l)
	75-79	File Sequence No.*
	80	The digit "8".
Card 9	1- 6	Iodate Oxidation as Oxygen (mg/l)
	8-13	Unused
	15-20	Unused
	22-27	Unused
	29-34	Unused
	36-41	Unused
	43-48	Unused
	50-55	Unused
	57-72	Unused
	64-69	Unused

These nine unused fields are read from card 9 and recorded on the file but are never used or printed in the current program.

These fields are included to provide for the addition of new data entries in the future.

	75-79	File Sequence No.*
	80	The digit "9".

*The file sequence number must be the same for all nine cards.

MAINTAIN

General Summary

The program MAINTAIN is a generalized maintenance program for the magnetic tape file system which performs any one of four major functions. Since MAINTAIN, in three of its four functions, copies certain information from an old USBM-FILE onto a new USBM-FILE, it is necessary to have a way of labeling these different versions. It is suggested that a number simply be added, USMB-FILE-n. Thus the original version of USBM-FILE will be called "USBM-FILE-1", the next version "USBM-FILE-2", etc.

Functions

- a. MAINTAIN is used initially to load the magnetic tape with data. The "initial loading" run refers to the first production run during which data is first placed on the tape file. Thus, the original loading run will always produce USBM-FILE-1. During the initial loading run the brine records are input in card form and then written on the magnetic tape USBM-FILE-1, so long as no obvious errors are encountered in the card data.
- b. The second major function of MAINTAIN is to add new records to USBM-FILE-n. These new records may be added only at the end of USBM-FILE-n and not inserted within it. Because of this, the file sequence number of the first new record added must be larger than largest file sequence number

currently on the file. To add the new records, the old USBM-FILE-n is first copied onto a new magnetic tape (USBM-FILE-n+1) and then the new records are read in card form and written onto USBM-FILE-n+1 after the old records. Here again each new record is checked for certain obvious errors.

- c. The third function of MAINTAIN is updating. This allows the replacement of an old record by a new one, hence records already on the file can be corrected or modified. The records to be replaced are referred to by their file sequence number. For instance, let us suppose that we wish to correct the two records whose file sequence numbers are 17 and 120. First, the complete new nine card record for each of these two records is punched on cards in the normal data format. The program during execution will then proceed as follows. First it copies all records with file sequence numbers less than 17 from USBM-FILE-n onto a new tape, USBM-FILE-n+1; then it reads new card record 17 and writes it on USBM-FILE-n+1. Next, it copies the records whose file sequence number are greater than 17 but less than 120 from USBM-FILE-n onto USBM-FILE-n+1; and then reads new card record 120 and writes it on USBM-FILE-n+1. Finally the remainder of USBM-FILE-n is copied onto USBM-FILE-n+1.

- d. The fourth major function of MAINTAIN is the deletion of records from USBM-FILE-n. For this function the card data is composed primarily of a list of file sequence numbers. USBM-FILE-n is copied onto a new tape, USBM-FILE-n+1, except for those records whose file sequence number appears in the card data.

Within each of these four functions the program records, and at the end of each run writes the total number of records on the file after the termination of the run and writes out, on the line printer, each of the records as it is loaded, added, replaced, or deleted.

Preparation of Card Input

Card 1. Only the first two columns of this card are used by the program.

Column 1. This column contains the function indicator.

1 indicates that an initial loading run is to be performed (function a).

2 indicates an addition run (function b).

3 indicates a replacement run (function c).

4 indicates a deletion run (function d).

Column 2. This column contains an indicator for the output format of each record loaded, added, replaced, or deleted.

0 indicates that the records are not to be written on the line printer.

1 indicates that the records are to be written in the format of the nine cards used as input.

2 indicates the records are to be written in the open file format.

3 indicates that only the file sequence number of the records is to be written.

After card one comes a deck of cards containing the brine records (in the previously specified format) which are to be read by the program. After this deck is placed a blank card which serves as an end-of-data indicator. The preparation of these remaining cards depends on the function specified on column one of card one.

Original loading run-

All records that are to be placed on the file are placed in order of increasing file sequence number. It is extremely important that the records be ordered in this way and that each of the nine cards of each given record also be in order. After all of the records are placed in their proper sequence, a blank card must be placed after them. This blank card denotes the end of the input data.

Add run-

All records that are to be added to the file must be placed in order of increasing file sequence number. The file sequence number of the first record to be added must be larger than the largest file sequence number previously on file. A blank card must follow these cards.

Replacement run-

The new records that are to replace the old records are arranged in order of increasing file sequence number. The additional restriction in this case is that each new record have a file sequence number less than or equal to the largest file sequence number currently on file. A blank card must follow these cards.

Deletion run-

For each record to be deleted only one card is prepared, instead of nine, as in the three previous functions. This card contains the file sequence number of the record to be deleted in columns 75-79 (right justified). Also column 80 must contain the digit 1 (one). These cards are then arranged in ascending order according to file sequence number. And finally, a blank card must follow these cards.

Description of Output

The output from the program MAINTAIN is of two forms, tape and printer. The tape output has meaning only for the three programs of this package and hence is not described in detail. The following discussion pertains only to the output printed on the line printer. The first message tells which type of function MAINTAIN has performed. This message may be any one of the following four: "ORIGINAL LOADING RUN", "ADD RUN", "REPLACEMENT RUN", or "DELETION RUN".

The next section of the output is the printing of the records processed. The format of this section is controlled by the indicator in Column 2 of Card 1 of the input cards. In a loading run, each record that is written on the tape file is also written on the line printer according to its specified format, unless the format indicator is zero. In an add run only those records that are added to the file are written on the printer in the specified format. If in a replacement run the format indicator is 1 or 2, both the old record that is being replaced and the new record that replaces it are written on the printer. First the new record is printed, next the word "REPLACES", and finally the old record is printed. If the format indicator is 3, only the file sequence number of the record being modified is printed. In a deletion run each record that is deleted is printed in the appropriate format.

The last section of the output contains three messages. First, is a message telling the total number of records now on the tape file. Second, is a message telling the largest file sequence number now on the file. Third, are the words "NORMAL RUN". This completes the normal output for MAINTAIN.

Error Messages

In addition to the normal output, there are several programmed error messages in MAINTAIN.

IMPROPER PARAMETERS

This message occurs if the value in column one of card one is anything but 1, 2, 3, or 4; or if the value in column two of card one is anything but 0, 1, 2, or 3. If either of these errors are encountered, the program is terminated and nothing is written on any of the output tapes.

CARD SEQUENCING WRONG AT xxxxx PROCESSING TERMINATED

This message is printed if any of the nine cards of a brine record are out of order, or if the file sequence numbers on the nine cards are not the same. In a deletion run, this message signifies that the digit in column 80 is not a one. The number xxxxx is the file sequence number of the first card read for the bad record. If this problem is encountered, the bad record is treated as an end-of-data indicator. The bad record is not written on tape or printer and the program transfers control to its normal routines for closing out the tapes and finishing the run. Thus, processing is completed correctly up to the bad record.

FILE SEQUENCING WRONG AT xxxxx PROCESSING TERMINATED

The number xxxxx denotes a card record whose file sequence number is not greater than the file sequence number just previously written on the tape file. This message can also occur in a replacement or deletion run if a record with a file sequence number larger than the largest existing value is to be replaced or deleted. Here again,

if this message is encountered, the normal finishing routines are used and the bad record is treated as an end-of-data indicator.

FILE SEQUENCE NO. xxxxx NOT ON FILE PROCESSING TERMINATED

This message may occur only in the replacement or deletion runs. It means that a record was specified to be replaced or deleted but that no such record can be found on the file. This condition is also treated as an end-of-data indicator.

DUMP

General Summary

The program DUMP writes USBM-FILE in one of several allowable forms. DUMP writes each record of the tape file on either the printer or card punch in either the card input format or the open file format. DUMP may use both the printer and the card punch in any given run.

Data Preparation

The program DUMP has one and only one input data card. This card contains the option specification for writing the records. This card is prepared as follows:

Column 1. This column contains the line printer indicator.

0 (zero) indicates that the records are not to be written on the line printer.

1 indicates that the records are to be written on the printer in the same format in which they were read from cards in MAINTAIN.

2 indicates that the records are to be written on the line printer in the open file format.

Column 2. This column contains the card punch indicator.

0 (zero) indicates that the records are not to be written on the card punch.

1 indicates that the records are to be written on the line printer in the card input

format. This gives the same results as a one in column one.

2 indicates that the records are to be punched onto cards in the same format in which they were read in MAINTAIN.

Columns 3-7 Smallest File sequence number to be written.

Columns 8-12 Largest File sequence number to be written.

After a successful run of DUMP, USBM-FILE-n must be saved by dismounting the tape.

Description of Output

The first message printed is "DUMP RUN." This message is printed only on the line printer. Following this preliminary message, the entire USBM-FILE is printed and/or punched as per specifications on the input card for DUMP.

Error Messages

There is only one possible error message that may be printed by this program. This is the message "IMPROPER PARAMETERS". This message indicates that one of the values on the input card for DUMP is not valid or that both of the values are zero specifying that nothing is to be written. If this message is written, the program is terminated and no processing occurs.

DICTIONARY

General Summary

The program, DICTIONARY, provides a means of indexing USMB-FILE. There are two different functions that DICTIONARY may perform. The first function is indexing six different items of the brine records; state, basin, county, formation, system, and series. DICTIONARY reads each record of USMB-FILE and enters into a list the name of the state, basin, county, formation, system, and series found on that record. In this way DICTIONARY records each different state name that is entered in the file and also the number of times this name is used. The same is done for each of the five other items, basin, county etc. After the entire USMB-FILE is read and indexed, these six lists are written onto USMB-DICT for future reference. Also the six lists are written on the line printer in "alphabetical" order.

Once USMB-FILE has been indexed, the six lists are written onto USMB-DICT. The second function of DICTIONARY simply reads these lists from USMB-DICT and writes them on the line printer.

Preparation of Input Cards

Normally there is only two cards of input for DICTIONARY. The first card contains three numbers as follows:

Column 1. This column indicates the type of function DICTIONARY is to perform. It must contain a 1 or a 2.

1 indicates that the index lists are to be compiled from USBM-FILE-n and written on the printer or cards.

2 indicates that the index lists have already been compiled and are available on USBM-DICT and that these lists are to be written.

Columns 2, 3. These columns contain the two digit file code of the output unit for the index lists.

Columns 79-80 These columns contain the file code of USBM-DICT. The remainder of this input card must be blank.

The second card contains integer numbers in columns one to five and six to ten. These numbers are to be right hand justified in their field. The first number is the maximum number of records to be searched. The second number is the largest file sequence number to be processed. If USBM-FILE has no sequence errors and has not been sorted these numbers will be the same.

Description of Output

The first message is "DICTIONARY RUN". This message is written on the line printer only. Following this are the index lists, either on the printer or on cards as specified. Each of the six lists are written in the same general format. Hence, we will consider only the format for state. The first message is of the form xxxxx DIFFERENT STATES. The number xxxxx gives the number of different state names that occur on the file. Following this message, the index list

for the states is printed in four columns. Each column contains a state name such as, OKLAHOMA and a number to the right of the name. This number is the number of times that this particular state name has been found in the file. For instance, KANSAS 10 means that Kansas has been found on ten different brine records.

After all six index lists are written the following messages are written on the line printer: The total number of records on the file, the largest file sequence number on the file, and the words "NORMAL RUN".

Error Messages

There are no error messages as such in DICTIONARY. If any of the indicators on the input data card are invalid the program simply stops and execution is not attempted. The only output in this case is the DICTIONARY RUN message.

It is also possible that the program will terminate before the entire USBM-FILE has been examined. This occurs because of the fact that certain limits are necessary as to the number of items in any index list. This limit refers to the number of different items and are the following:

States	60
Basins	76
Counties	500
Formations	550
Systems	40
Series	122

As soon as any one of these limits is reached, the program writes the lists and finishes.

FETCH

General Summary

FETCH is the data retrieval program of this data processing package. Its primary function is to search USBM-FILE for records satisfying certain conditions and then print certain information in either tabular form or in the open file format from these records. It is primarily for use in FETCH that the index lists are generated by DICTIONARY. These lists are used to insure that a search parameter requested by FETCH does indeed exist somewhere on USBM-FILE.

There are two parameter sets contained in the card input data for FETCH. These are the search parameters and the write parameters. The search parameters tell FETCH what type of records are to be extracted from the file. The write parameters tell FETCH what parts of the extracted record to write either on the printer or on cards, or both. The information written on the printer may be in either a tabular format or in the open file format. The information written on cards may be in tabular form only.

In order to facilitate searching and writing, the items of the brine record are given a code number and the item itself is considered as either a numeric or alphabetic item. These code numbers are given below.

Code Number	Item Type	Description
1	Numeric	Top of formation depth (ft)
2	Numeric	Bottom of formation depth (ft)
3	Numeric	Elevation (ground level) (ft)

4	Numeric	Resistivity
5	Numeric	Specific gravity at 80 F
6	Numeric	Total dissolved solids
7	Numeric	Lithium
8	Numeric	Sodium
9	Numeric	Potassium
10	Numeric	Rubidium
11	Numeric	Cesium
12	Numeric	Calcium
13	Numeric	Magnesium
14	Numeric	Strontium
15	Numeric	Barium
16	Numeric	Strontium-Barium
17	Numeric	Beryllium
18	Numeric	Boron
19	Numeric	Germanium
20	Numeric	Copper
21	Numeric	Chromium
22	Numeric	Aluminum
23	Numeric	Lead
24	Numeric	Tin
25	Numeric	Silver
26	Numeric	Cobalt
27	Numeric	Nickel
28	Numeric	Bismuth
29	Numeric	Vanadium
30	Numeric	Iron

31	Numeric	Manganese
32	Numeric	Chloride
33	Numeric	Bromide
34	Numeric	Iodide
35	Numeric	Bicarbonate
36	Numeric	Carbonate
37	Numeric	Sulfate
38	Numeric	Sulfide
39	Numeric	Nitrate
40	Numeric	Phosphate
41	Numeric	Organic acid as acetic
42	Numeric	Organic carbon
43	Numeric	Organic nitrogen
44	Numeric	Variable organic nitrogen
45	Numeric	Stable organic nitrogen
46	Numeric	Permanganate oxidation as oxygen
47	Numeric	Iodate oxidation as oxygen
48	Numeric	Unassigned
49	Numeric	Unassigned
50	Numeric	Unassigned
51	Numeric	Unassigned
52	Numeric	Unassigned
53	Numeric	Unassigned
54	Numeric	Unassigned
55	Numeric	Unassigned
56	Numeric	Unassigned
57	Numeric	Depth of well (ft)

58	Numeric	pH
59	Alphabetic	State
60	Alphabetic	Basin
61	Alphabetic	Field
62	Alphabetic	County
63	Alphabetic	Formation
64	Alphabetic	System
65	Alphabetic	Series
66	Alphabetic	Location
67	Numeric	Date sample obtained
68	Numeric	Well completion date
69	Numeric	Lab number without letter
70	Alphabetic	Letter of lab number

The following limits for the tabular format are established arbitrarily for the printer and card output. On the printer there are always eleven columns printed plus a column for the file sequence number; however, any of these columns may be blank. After the file sequence number, the first three columns may contain any three of the alphabetic items listed above. Certain special considerations are given to the Location (code 66). If this code is specified, only the first 24 characters of the location given on the brine record are printed in the column. To the right of the alphabetic columns are eight columns that may contain any eight of the numeric items listed above. Here special consideration must be given to the date items (codes 67 and 68). If either of these items are to be written, they are written in the form xxyyzz where xx = day, yy = month, and zz = year.

If punched card output is specified, there are ten columns that may be used in addition to the file sequence number. These ten columns may contain only numeric items with the same conventions applying to codes 67 and 68 as on the printer.

Preparation of Card Input

The first input card for FETCH immediately following the "DICTIONARY" data deck is a specification card that controls the checking of the index lists (data deck) generated by DICTIONARY. The second is a title card for the alphabetic tabular columns that are to be printed on the printer, the third, a title card for the numeric columns on the printer. The fourth and fifth cards contain the write parameters for the printer and punch respectively. Next are the cards containing the search parameters, and finally comes a blank card indicating the end of the search parameters.

Card 1. The only column on this card in which information may be punched is column 1. If column 1 contains the digit 1 (one), the index lists generated by DICTIONARY are checked for each state, basin, county, formation, system, and series search parameter that is specified on the search parameter cards. If the specified parameter is not found in its appropriate index list, an error message is printed and the search of the tape file is not made. This prevents searching the tape for an item that does not exist. If column 1 contains a 0 (zero), the check against the index lists are not made. Columns 2-80 of Card 1 must be blank.

Card 2. This card provides a title for each of the 3 alphabetic columns of tabular data to be printed on the printer. If the open file format is to be used on the printer this card should be left blank. This card should be prepared after card 4.

Columns 2-23 Title for alphabetic column one.

Columns 25-46 Title for alphabetic column two.

Columns 48-69 Title for alphabetic column three.

The titles given on card 2 are printed character for character as the title for the three alphabetic columns on the printer.

Card 3. This card provides a title for each of the 8 numeric columns of tabular data to be written on the printer. If the open file format is to be used on the printer, this card should be left blank. This card should be prepared after card 4.

Columns 2- 7 Title for numeric column one.

Columns 9-14 Title for column two.

Columns 16-21 Title for column three.

Columns 23-28 Title for column four.

Columns 30-35 Title for column five.

Columns 37-42 Title for column six.

Columns 44-49 Title for column seven.

Columns 51-56 Title for column eight.

Card 4. This card contains the write parameters for the printer. If no output is desired from the printer, this card should be left entirely blank. If the

open file format is desired for each extracted record, columns 2-3 should contain 99. If these columns do contain 99, the remainder of card 4 should be left blank.

Columns 2- 3 Item code for first alphabetic column.
Columns 5- 6 Item code for second alphabetic column.
Columns 8- 9 Item code for third alphabetic column.
Columns 11-12 Item code for first numeric column.
Columns 14-15 Item code for second numeric column.
Columns 17-18 Item code for third numeric column.
Columns 20-21 Item code for fourth numeric column.
Columns 23-24 Item code for fifth numeric column.
Columns 26-27 Item code for sixth numeric column.
Columns 29-30 Item code for seventh numeric column.
Columns 32-33 Item code for eighth numeric column.

If the item code for any of the 11 columns to be written on the printer is left blank, the program assumes that this column is to be left blank in the tabular output.

Card 5. This card contains the write parameters for the card punch. If no punched output is desired, this card should be left entirely blank.

Columns 2- 3 Item code for numeric column one.
Columns 5- 6 Item code for numeric column two.
Columns 8- 9 Item code for numeric column three.
Columns 11-12 Item code for numeric column four.
Columns 14-15 Item code for numeric column five.

Columns 17-18 Item code for numeric column six.
Columns 20-21 Item code for numeric column seven.
Columns 23-24 Item code for numeric column eight.
Columns 26-27 Item code for numeric column nine.
Columns 29-30 Item code for numeric column ten.

If the item code for any of the 10 columns to be written on the cards is left blank, the program assumes that this column is to be left blank in the tabular output.

Cards 6, 7,....These are the cards that contain the search parameters by which records are to be extracted from the file. Each search card is divided into four major fields. Field one is in column 1-2; field two, in columns 4-25; field three, in columns 26-47; and field four, in columns 48-69. Field one always contains the item code of the data item to which the remainder of the card pertains. In general, fields two, three, and four specify the conditions that a brine record must satisfy in order for it to be extracted from the file. These three fields are treated in an inclusive either/or sense so that if one of the conditions specified in any of the three fields is satisfied, the condition on the item referred to by the item code in field one is considered to be satisfied. The method of placing information in fields two, three, and four depends upon whether or not the item referred to by field one is considered as numeric^{or} alphabetic.

Alphabetic Item-

For alphabetic items there is only one way in which information may be placed in fields 2, 3, and 4. Let us consider only field 2, since 3 and 4 are treated in the same way. If, for instance, we wish the extraction condition for the item state to be that the state name is KANSAS, we would write KANSAS in field 2 (left justified). This means that any brine record from the state of Kansas will be extracted from the file, provided the conditions on each other search item as specified on the remaining cards are also satisfied. If we wish the condition for extraction for the item State to be that the state name is either OKLAHOMA or TEXAS or COLORADO, we would place OKLAHOMA in field two, TEXAS in field three, and COLORADO in field four. This means that a brine record will be extracted if it is from either Oklahoma, Texas, or Colorado, provided all other extraction conditions as specified on the other search parameter cards are also satisfied. The fields 2, 3, and 4 must be used in that order; however, the sequence of names is irrelevant. For instance, KANSAS in field two and TEXAS in field three would give the same condition as TEXAS in field two and KANSAS in field three.

Numeric Item-

For numeric items the placing of information in fields 2, 3, and 4 is more involved, because there are three ways in which any one of these three fields may be

handled. These three ways correspond the three conditions:

$$a < x, \quad b \leq x \leq c, \quad x < d.$$

In order to handle these three different conditions, fields 2, 3, and 4 are subdivided into subfields - 2.1, 2.2; 3.1, 3.2; 4.1, 4.2. These subfields correspond to the following cards columns.

2.1 Columns 4- 9

2.2 Columns 10-15

3.1 Columns 26-31

3.2 Columns 32-37

4.1 Columns 48-53

4.2 Columns 54-59

If a number "a" is in subfield n.1 and subfield n.2 is blank, this specifies the condition $a < x$ where x is the value found on a brine record for the item specified in field one. If a number "b" is in n.1 and another number "c" is in n.2, this specifies the condition $b \leq x \leq c$. If n.1 is left blank and n.2 contains the number "d", this specifies the condition $x < d$. In the case of these numeric subfields, all numbers should be right justified. It should be noted in passing that the condition $a = x$ is equivalent to the condition $a \leq x \leq a$.

To further confuse the issue, it is possible for these "numeric" subfields to contain alphabetic information, such as the letters, ND. The relative "numeric" value of these alphabetic characters may be determined for any machine. Thus any alphabetic character has a "numeric" value.

In this way a search card can be prepared for any of the 70 data items. It is not necessary, however, to prepare a search card for each of the 70 items. If no search card is prepared for a certain item, this data item is not considered in the extraction of records. As implied previously, the search conditions for all data items that are specified as search parameters are considered simultaneously.

To help clarify all of these ideas let us consider an example. Let us suppose that we wish to find all records that satisfy the following conditions:

1. The state is either KANSAS or NEBRASKA.
2. The top of formation depth is greater than 1000 ft.
3. The sodium content is less than 5 or the sodium content is greater than 300.
4. The magnesium content is recorded as ND.

These four conditions define the search parameters such that a brine record is extracted if and only if all four of these conditions are satisfied simultaneously. Let us also suppose that we wish to write on the printer in tabular form the following items:

state, system, formation, depth to top of formation, sodium content, magnesium content, resistivity, and well completion date.

The search parameters are first compared with the index lists compiled by DICTIONARY to assure availability of the information.