DEPARTMENT OF THE INTERIOR

U.S. GEOLOGICAL SURVEY

Accessing geochemical data from the Branch of Analytical
Chemistry computer database

by

Michael M. Kelly

Open-File Report
86-523

DISCLAIMER

ALTHOUGH THIS PROGRAM HAS BEEN EXTENSIVELY TESTED, THE U.S.
GEOlOGICAL SURVEY CANNOT GUARANTEE THAT IT WILL GIVE ACCURATE
RESULTS FOR ALL APPLICATIONS OR THAT IT WILL WORK ON ALL
COMPUTER SYSTEMS.

ANY USE OF TRADE NAMES IS FOR DESCRIPTIVE PURPOSES ONLY AND DOES
NOT IMPLY ENDORSEMENT BY THE AUTHOR OR THE U.S. GEOLOGICAL
SURVEY. THIS REPORT IS PRELIMINARY AND HAS NOT BEEN REVIEWED
FOR CONFORMITY WITH THE U.S. GEOLOGICAL SURVEY EDITORIAL
STANDARDS.

Menlo Park, California 1986

ABSTRACT

This report describes a general method for electronically transferring chemical data stored in the Branch of Analytical Chemistry (BAC) HP1000 database to micro computers. The included software supports porting the data to an IBM PC for use with SYMPHONY, MINITAB, EXCEL, or GPP. Files may then be sent to the ISD VAX 11/785 mainframe from the PC. This transfer method can move 50 samples (100 elements per sample) into a spreadsheet in under 10 minutes.

INTRODUCTION

As more geologists use spreadsheets and graphic computer programs to aid in chemical data analysis, the need to accurately and quickly enter large amounts of data into electronic databases has become apparent. It is now possible to access and copy the analytical results that are stored on the BAC HP1000 computer database by linking with a micro-computer over a telephone line. An example file transfer, using an IBM-PC is documented in this report. Hardware and software requirements are outlined, and use of the included software for formatting data for different spreadsheets is explained.

HARDWARE AND SOFTWARE

The access method is designed for the IBM-PC computer family, although the requirements are similar for other operating systems. The minimum required system is a PC with 256K RAM, a serial port, a modem and matched communication software. Many combinations of software and modems are possible, and some guidelines are mentioned below.

The modem should "communicate" at 1200 or 300 bits per second. It should have a cable that connects it to the serial port on the PC, and a plug for a standard phone jack. Acoustic modems that have cups for the telephone handset instead of the plug for the telephone cable are also adequate.

The communication software must be capable of working with the selected modem. Newer modems conform to a standard known as the Hayes command set. Since most software is written with this standard in mind, using a Hayes compatible modem almost guarantees that the communications software will work correctly. The software must be capable of receiving and sending ASCII data files.

Once the system is configured, with the communications settings at 300 or 1200 baud (1200/300 bits per second, Parity = E, 7 data bits, 1 stop bit), call the HP1000 using the procedure descibed under the Method section. Each group of chemistry samples is catalogued in the HP1000 database by its **job** number. For instance, my samples from 1985 had the same job number : LN26. The program on the HP1000 that collates the data for

transfer will request that number. In this example the program
will put all the **completed** parts of the job into a new file
named : DLN26. The "D" signifies a download file. The job
number can be found on the BAC form: "Request For Analysis " ,
which should have been returned to the geologist when the
analysis process began. Jobs that are older than 1982 are
probably not available, and must be personally requested from
BAC.
      The following description of a terminal session is based
on the communications program PC-Talk. PC-Talk is "shareware",
and can be tried prior to purchase. Purchasing information is
given in the appendix. The sample terminal session description
below, outlines the commands given to the HP1000, which are
applicable to any communication setup you might be using. It is
recommended that the user become familiar with the operation of
the modem and software before attempting to access the HP1000.
      The modem used for this session was the XEROX model 300
acoustic. It has two switches on it, a power switch, and a
duplex switch, which is set for full duplex. There is a 25 pin
cable connecting the modem to the serial port on the IBM-PC.

Method

Accessing the HP1000 computer, and downloading your files

   Dial up the BAC Computer at:

        321-5273   For 300 bps modem (like the XEROX 300)

        321-5293   For 1200 bps modem

The HP1000 will answer the call and ask you to log-on. Below is
a typical transaction between the IBM-PC using PC-Talk and the
HP1000. Remember the following conventions:

1)   Comments are compressed letters.

2)   <CR> represents a carriage return.

3)   : The HP1000 prompt is a colon
         indicating that the machine is waiting for a command.

4)   **The bold text is the response to a question, or a
command to be given.**

PLEASE LOG-ON: **user**   type 'user'

Now you will get the login message of the HP1000

```
*********************************************************

<<<<<<<<<<            RTE 6 V/M              >>>>>>>>>>

<<<<<<<<<< Branch of Analytical Chemistry >>>>>>>>>>

<<<<<<<<<<      Menlo Park California       >>>>>>>>>>

*********************************************************

  the  HP1000 cursor appears


 : bac01 <CR>  This command runs the program that makes your datafile


     File NAMR for data. [EX:DATA:0:14]  DLN26 <CR>   create a file for
 the data . Use the job number preceded by a "D". Note CAPITAL letters !!


Creating DLN26
 Do you want the samples identified by:
1-Branch lab. number.
2-Submitters field number. 2        Always use your field numbers !

Do you want a checksum at the beginning of each
 out-put line? N <CR>                           No checksum !

Getting all SAM projects.
Opening AUK data base
Closing the data base!
Opening SAM data base

Job number (EX:AA01) [END to end] LN26 <CR>      The job number again
                                                 but notice there is no
JOB -LN26MX exists.                                'D' preceding it !
JOB -LN26MP exists.
Closing the SAM data base!
Getting lab. numbers.
Getting answers.
Opening answer file LN26MX:30:40
Job LN26MX:30:40 is not finished - skipped!      Some part of this job is
Opening answer file LN26MP:30:40                 unfinished

Changing all "MT'S" to 0!
Writing data to DLN26
Opening SAM data base

Job number (EX:AA01) [END to end] END <CR>       Enter END to stop
Closing the SAM data base!
Closing DLN26
normal stop
```

You have just created your datafile, now tell PC-TALK to capture it on the IBM-PC

: **ALT-R**   You are now "communicating " with PC-Talk

specification: **A:DLN26 <CR>**   PC-Talk is saving whatever appears on the screen in a file called DLN26 on drive A, and now you are communicating with the HP1000

: **du,DLN26,1 <CR>**   This will list the file named DLN26 on the HP1000

```
LN26
$ 85MP 91MP 92MP 97MP                         Here is your data !
&MK85-13
0  625-20     58-20     90-20     10-2
&MK85-14
0  791-20    123-20    164-20     10-2
&MK85-15
0  305-20    222-20    148-20     12-2
&MK85-16
0  602-20     82-20    165-20     13-2
&MK85-17
0  512-20     77-20    138-20      4-2
:
```


PC-Talk has captured the file that was just listed. Now terminate the capture by typing ALT-R

: **ALT-R**

: **ex,rp <CR>**   This will log you off of the HP1000

Hang up the phone. Exit from PC-TALK by typing
ALT-X.

ALT-X

FORMATTING

        The coded chemistry data must now be deciphered using the software included in this report. If a 5.25 inch diskette was not included with this report, contact the Open File Services through the U.S.G.S. office in Menlo Park using the number listed in the appendix. AFORM is an interactive, self – explanatory TURBO PASCAL program that runs on the IBM-PC and very compatible clones. The program is in executable form so it does not require a compiler. Once the operating system is booted, AFORM may be run directly. AFORM formats the data so

that it can be imported into either Symphony, Excel, Minitab or
GPP. When the program is run, it prompts for the brand of
spreadsheet in use. The program decodes the data, into rows and
columns, with chemical species across, and samples down. To
prepare the BAC datafile for use with AFORM, a text editor must
be used to remove any extra text that was captured with the file
when it was transferred. In the sample file transfer listed
above, the HP1000 command DU,DLN26,1 was captured along with the
data, and this information must be deleted. The file should
have the job number on the first line, followed by line(s) of
coded species, followed in turn by line(s) of sample data. Here
is an example of an edited file:

```
@LN26                        First line is title
$ 85MP 91MP 92MP 97MP        This line has coded elements
&MK85-8                      First sample #
0  782-20  109-20  82-20   6-2   First sample data
&MK85-13
0  625-20   58-20  90-20  10-2
&MK85-14
0  791-20  123-20  164-20  10-2
&MK85-15
0  305-20  222-20  148-20  12-2
&MK85-16
0  602-20   82-20  165-20  13-2
&MK85-17
0  512-20   77-20  138-20   4-2   Last line
```

     Once the data is edited as above, AFORM can be used. To
run this program, type AFORM from the drive where it is located.
The program will prompt the user for information in a fairly
unobnoxious manner. AFORM allows the coded data to be formatted
for different spreadsheets. When specifying the BAC data file
to be formatted make sure that the filename is preceded by the
drive\directory specification. If an attempt is made to format
a file without specifying its location the program will crash if
the file is not on the same disk as the AFORM program.
Similarly,in order to send the new formatted file to another
disk,the drive letter and directory must be specified. AFORM
takes up 65K of disk space, and creates new files of about one
to two times the size of the original BAC file.
     AFORM can process up to 50 samples with up to 100 species
per sample at one time. If the datafile has more than fifty
samples, the file can be edited into two parts, and processed
seperately.In this case the title line and the coded element
lines must be copied to the split file, followed by the data
lines that exceed the fifty sample limit.
     It is also a good idea to name the new file with a suffix
that describes its contents. If the original file is named
KELLY.DAT, and AFORM is used to format a new file for Symphony,
then the new file would be called KELLY.SYM.

Sending the data files to VAX

5

If MINITAB on the ISD VAX (Menlo Park) is the favored spreadsheet, then the data can be transferred to that computer using a communication program that allows ASCII file transfers. The following is a sample transaction between the PC and VAX. First, log-on to VAX in the normal manner.

$ cr filename.dat <CR>.   Vax will record whatever is typed.

ALT-T   Instruct PC-Talk to dump the
                                contents of the formatted
                                datafile
                                PC-Talk will prompt for the file
                                to send. Indicate the drive and
                                filename ie A:filename.dat.
                                PC-Talk will now dump the file.

Cntrl-Z   Closes the new Vax file

The new file may now be opened by MINITAB.

<u>Appendix</u>

         PC-TALK (35.00$)
         The Headlands Press Inc.
         Box 862
         Tiburon,CA 94920
         415/435-9775

         USGS
         (415) 323-8111

         Cable Configuration for XEROX 300/1200 acoustic modem to IBM
         serial port:

         IBM   XEROX 300/1200
          DB25S   DB25P
            1----------------3
            2----------------2
            3----------------1