DOCUMENT RESUME

ED 344 592 IR 053 964

AUTHOR Cardullo, Pamela; Neuman, Michael
TITLE Catalogue of Projects in Electronic Text. User's Guide.
INSTITUTION Georgetown Univ., Washington, DC. Center for Text and Technology.
PUB DATE Nov 91
NOTE 33p.
AVAILABLE FROM Academic Computer Center, 238 Reiss Science Bldg., Georgetown University, Washington, DC 20057.
PUB TYPE Guides - General (050) -- Reference Materials - General (130)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Computational Linguistics; Content Analysis; Electronic Publishing; *Full Text Databases; Higher Education; *Humanities; Information Networks; Literary Criticism; Online Systems; *Primary Sources; *Search Strategies; Telecommunications
IDENTIFIERS *Electronic Text; Georgetown University DC

ABSTRACT

This guide to searching the Georgetown University Catalogue of Projects in Electronic Text database provides information on electronic text projects in the humanities throughout the world as well as a variety of information on the many collections of literary works, historical documents, and linguistic data that are available from commercial vendors and scholarly sources. It is noted that the electronic text projects documerted in the database are machine-readable files of primary materials from humanities disciplines which generally take the form of either large corpora for linguistic analysis or major works of major authors for analysis of style and content; these files are intended by their developers to be searched and manipulated directly by humanists. The guide provides instructions for moving around the computer screen and keyboard, obtaining a free communications package that is available for use with the catalogue, and accessing the catalogue from remote sites. An explanation of the structure of the database and instructions for searching the PRODUCT and WORK subsystems for complete product entries or information about individual texts are also provided. Sample screens are included. (MAB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

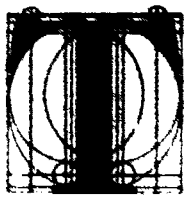
ED344592

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy



THE CENTER FOR TEXT & TECHNOLOGY of the Academic Computer Center

presents

The Georgetown University

Catalogue of Projects in Electronic Text

November, 1991

A project of the Center for Text & Technology.
Designed by Pamela Cardullo and Dr. Michael Neuman.
Documentation by Pamela Cardullo.

Michael Neuman, Ph.D., Director
CENTER FOR TEXT & TECHNOLOGY

Susan T. Stoler, Ph.D., Director
ACADEMIC COMPUTER CENTER

Georgetown University
Reiss Science Building, Room 238
37th and O Streets, NW
Washington, DC 20057

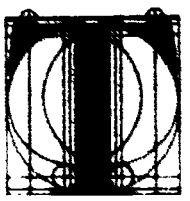
"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Michael Neuman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

12053964

BEST COPY AVAILABLE



THE CENTER FOR
TEXT & TECHNOLOGY
of the Academic
Computer Center

presents

The Georgetown University

Catalogue of Projects in Electronic Text

November, 1991

A project of the Center for Text & Technology.
Designed by Pamela Cardullo and Dr. Michael Neuman.
Documentation by Pamela Cardullo.

Michael Neuman, Ph.D., Director
CENTER FOR TEXT & TECHNOLOGY

Susan T. Stoler, Ph.D., Director
ACADEMIC COMPUTER CENTER

Georgetown University
Reiss Science Building, Room 238
37th and O Streets, NW
Washington, DC 20057

Table of Contents

Section **Page**

Introduction 1

Moving Around the Screen and Keyboard 4
- Menu Selections 4
- Key Mapping 4

Free Communications Package Available for Use With the Catalogue 5
- Obtaining a Copy of Kermit 5
  - Via Mail 5
  - Via FTP 5

Accessing the Catalogue from Remote Sites 8
- Via Telnet 8
- Via Modem 8
- The Initial Catalogue Screens 9

Structure of the Database 10

The PRODUCT Subsystem 11
- Simple Queries 11
- Multiple Queries 12
- Wildcards 14
- Query Returns 15
- Viewing a Complete Product Entry 16

The WORK Subsystem 22
- Simple Queries 23
- Multiple Queries 23
- Query Returns 25

APPENDICES

Appendix A Field Type and Length 29

Appendix B Electronic Documentation 31

The Georgetown University

Catalogue of Projects in Electronic Text

Introduction*

Since April of 1989, the Center for Text & Technology (CTT), under the aegis of the Academic Computer Center at Georgetown University, has been compiling a catalogue of projects that create and analyze electronic text in the humanities. The *Georgetown University Catalogue of Projects in Electronic Text* is a powerful database that includes information on electronic text projects throughout the world. The database includes a variety of information on the many collections of literary works, historical documents, and linguistic data which are available from commercial vendors and scholarly sources. The database is written in Ingres and resides on a VAX 8700 computer at Georgetown University. The database may be searched by off-campus users who can connect to the database using Telnet or a modem.

The electronic text projects documented in the database are machine-readable files of primary materials from humanities disciplines. Whether entered by keyboarding or by scanning with an optical character reader, these text files generally take the form either of large corpora for linguistic analysis (such as the new British National Corpus of one hundred million words currently being developed by Oxford University Press and others) or major works of major authors for analysis of style and content (such as the compact disc of the Thesaurus Linguae Graecae containing 1400 years of classical Greek texts). Beyond the scope of the project are electronic versions of encyclopedias, dictionaries, and secondary studies, concordances, databases, or computer-assisted instruction, unless such proejcts include full-text versions of primary works.

Unlike the databases that research libraries often make available, the electronic texts cataloged at Georgetown are intended by their developers to be searched and manipulated directly by humanists. Often, therefore, the text is encoded with markup language to facilitate integration with other files; occasionally, the texts are combined with a commercial text-analysis tool such as WordCruncher, Folio Views, or Micro-OCP.

With electronic text and integrated analysis software, the researcher not only has the equivalent of an interactive concordance for finding instances of key words but can also search for clusters of words, exact phrases, and co-occurrences of key words (sorted by boolean operators) in contexts of various sizes. Statistical programs show where the desired term or concept is concentrated in a work or series of works, and parsing programs can analyze parts of speech and syntactic structures.

In general, therefore, the combination of electronic text and searching software can be said to provide the researcher with both microscopic and macroscopic views of the text. The former provides access to small-scale features of a single work; for example, within seconds, a philosopher could locate the single occurrence of the phrase "consciousness of absolute being" from the nine-megabyte, three-volume translation of Hegel's *Lectures on the Philosophy of Religion*. By contrast, the macroscopic view of the text highlights the ways in which one work differs from other works by the same author or the author's contemporaries; for example, if one searches an eleven-megabyte file of

* A version of this section appeared in the *Western European Specialists Section Newsletter*, Vol.14, No. 2 (Spring 1991), pp. 1, 3-4.

Shakespeare's works for the word "time," one finds a greater concentration in *Macbeth* than in the other tragedies, and by exploring the contexts, one can see how the title character's over-reaching can be explained thematically in terms of his attempt to usurp the providential function that belongs to time.

Given these advantages, it is not surprising that the conversion of primary texts to electronic form is proliferating throughout the world. Nevertheless, because of the unpublicized academic nature of such projects, the process of locating them can be difficult. For this reason, we rely heavily on the discussion groups on BITNET and Internet, not only to identify new projects but also to request information about them and to disseminate the material we compile. Electronic mail provides access to the most recent developments and permits us to receive and transmit information throughout the world quickly and economically. Among the sixty discussion groups we monitor are those in language and literature (Ansax-L, C18-L, Chaucer, Contex-L, English, Ficino, Linguist, Litera-L, Literary, Reed-L, Rustex-L, Shakesper, and Wwp- L), culture and religion (Ccnet-L, Indology, Japan, Judaica, and Religion), libraries (Cdrom-L, Fisc-L, Libref-L, Pacs-L, and Tei-L), philosophy and history (History, Philos-L, ad Philosop), and the humanities in general (Erl-L, Gutnberg, Humanist, and Pmc-Talk).

In our search for news of projects, we also review a wide range of publications, including popular magazines and newspapers (such as the Chronicle of Higher Education), agency reports (such as the List of Awards of the National Endowment for the Humanities), trade publications (including InfoWorld and EDU Magazine), discipline specific journals (such as Computers and Philosophy and Computers and the Classics), the newsletters of numerous academic computing centers, and the journals central to humanities computing (Computers and the Humanities, Bits and Bytes Review, and the ICAME Journal).

Once we have identified a new project, we request ten categories of information:

0. Identifying acronym or short reference;
1. Name and affiliation of operation (including collaborators) with references to any published description;
2. Contact person and/or vendor with addresses;
3. Primary disciplinary focus (and secondary interests);
4. Focus: time period, geographical area, or individual;
5. Language(s) coded;
6. Intended use(s) and Size (number of works, or entries, or citations);
7. File format(s);
8. Form(s) of access (outline, tape, diskette, CD-ROM, etc.);
9. Source(s) of the archival holdings: encoded in-house, or obtained from elsewhere.

As of November, 1991, 340 projects are described im the Catalogue

Because the catalogue is constantly being updated, any printing would be almost immediately obsolete. Consequently, the CTT has converted the catalogue to an online database — searchable through Telnet and dial-in access — so that current information

can be made available to researchers. In addition, searches of the catalogue are performed on request, and updated lists of projects and addresses are posted regularly on the HUMANIST electronic bulletin board and distributed through surface and electronic mail.

For further information about the project, or to request a specific search, please contact:

James A. Wilderotter II, Project Assistant
The Center for Text and Technology
Academic Computer Center
238 Reiss Science Building
Georgetown University
Washington, DC 20057

(202) 687-6096
BITNET: wilder@guvax
Internet: wilder@guvax.georgetown.edu

Moving Around the Screen and Keyboard

Menu Selections

The Georgetown University Catalogue of Projects in Electronic Text is a largely menu-driven program. Menu selections are displayed at the bottom of the screen. Use the keypad to the right of your keyboard to select the various menu selections. You may need to turn off NUM LOCK to activate the keypad. The menu selections also require that you use PF keys. Normally PF1 corresponds to Function Key F1, etc.

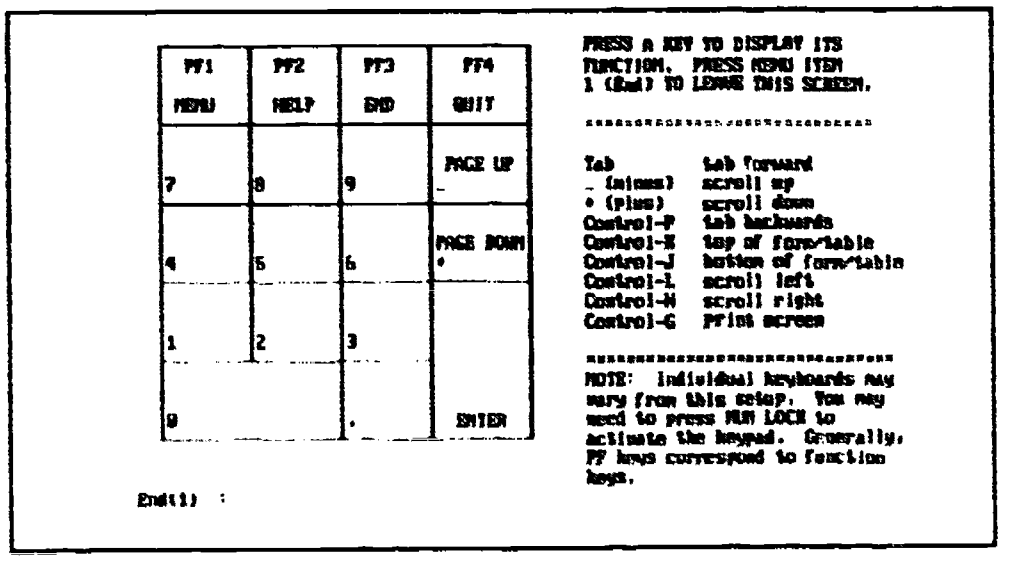
End The End key will always be PF3. End generally returns you to the normal sequence of events. In most cases, it will return you to the previous screen. When you are paging through a complete entry, however, it will take you to the next screen in sequence.

Quit The Quit key is always PF4. Pressing Quit will exit you from the database.

Occasionally there are more menu options available than can be displayed in one line at the bottom of the screen. You can tell if there are additional options available because a carat (>) will appear at the end of the menu items. To see the additional options, press PF1. Pressing PF1 toggles you through all the available options. Press PF1 again to see the beginning of the list. You can select any of the available options, even if you cannot see them.

Key Mapping

The various function keys, PF keys and menu item selection keys may vary, depending on the type of terminal you are emulating when you log onto the program. A basic keymap is available on-line from most screens. When you select the Key option from any screen, the following is displayed:



Note that in general the PF keys correspond to function keys. From this template you can press PF1, PF2, PF3 or PF4 for a description of what that key does. To leave this template and return to the previous screen, press End (1).

Free Communications Package Available for Use With the Catalogue

Users accessing the database from remote sites may connect to the Georgetown host computer via either Telnet or by dialing in from a modem. See page 8 below. If you do use a modem, or if you must connect via modem or a dataline connection to a larger computer running Telnet, you will inevitably make use of a communciations package to do so. VTerm, Proconm, MacTerminal, Kermit -- these are all examples of communications packages which you might use. Each of these communications packages configures the keyboard a little differently.

In order to facilitate your ability to use the Catalogue with success, The Center for Text & Technology has received permission to distribute to its users a customized version of Kermit, a terminal emulation program that adapts the user's keyboard to our on-line database. We can provide Kermit for use with either an IBM-compatible personal computer or a Macintosh. These versions of Kermit use the keyboard configuration outlined on page 4. We hope that by providing this service, we will make it easier to use both the database and the accompanying documentation While you do *not* have to use this particular package in order to connect to the database, we cannot provide information on keyboard mapping for other versions or software packages.

Obtaining a Copy of Kermit

Users can retrieve a copy of the Kermit program via FTP (File Transfer Protocol) from *guvax.georgetown.edu* (141.161.1.2). Specific instructions for retrieving the program are detailed below.

Users who do not have access to a machine running FTP may request to receive a copy of Kermit through the mail.

Via Mail

To have a copy of Kermit sent to you through the mail, send a note or mail message to the following address:

James A. Wildcrotter II, Project Assistant
The Center for Text & Technology
Academic Computer center
238 Reiss Science Building
Georgetown University
Washington, DC 20057

(202)687-6096
BITNET: wilder@guvax
Internet: wilder@guvax.georgetown.edu

Please provide your mail address and specify whether you use a Macintosh or IBM-compatible, and the appropriate diskette size (3-1/2" or 5-1/4").

Via FTP

File Transfer Protocol (FTP) is the fastest means of transferring a file from one computer to another. FTP allows users to log onto remote computers to send and retrieve files. To do so, one must have an account not only on one's home computer but also on the computer from which you

plan to retrieve the file. Fortunately, many computers (including the guvax.georgetown.edu site) allow one to log on as "Anonymous" and use one's real name as the password. If you have any questions about using FTP, please contact your system manager. Specific commands may vary a little, depending on your host computer. The following procedures, however, should work in most cases.

To retrieve a copy of Kermit for either the Mac or an IBM-compatible, perform the following procedure:

A. To initiate the FTP session, log onto a computer running FTP. Enter *ftp* followed by the name of Georgetown's host computer, guvax.georgetown.edu:

    ftp guvax.georgetown.edu

B. Once the connection has been successfully made, the message Connection Opened will be displayed along with some system information, followed by a login prompt (in this case Name:).

    Connection Opened
    Using 8-bit bytes
    220 guvax.georgetown.edu Wollongong FTP Server (Version 5.1)
    on 20-Sept-1991 at 13:09:21
    Name (guvax.georgetown.edu:pam):

    At the Name: login prompt enter 'anonymous'. You will then be prompted for a password. Enter your real life name as the password. Once you have successfully logged on, an asterisk prompt (*) will be displayed.

    Name (guvax.georgetown.edu:pam): ***anonymous***
    331 Anonymous user ok, send real identity as password.
    Password: ***<enter your name here>***
    230 ANONYMOUS logged in, directory DISK$ACC1:[ANONYMOUS],
        restrictions apply.
    RMS File Transfer mode in use.
    *

C. Now that you are logged onto Georgetown's host computer, GUVAX, you can enter FTP commands at the asterisk prompt (*) to access and retrieve the files you need.

    To retrieve the IBM-compatible version of Kermit (CPETKERM.EXE) and the accompanying text file (AAAREADME.TXT), enter the following series of commands at the asterisk (*). Commands that the user must enter are in ***bold italics***. Responses from the host computer are in this font:

    **** cd ms-dos***
    Working director changed to "Disk$ACC1:[ANONYMOUS.MS-DOS]"
    **** cd kermit***
    Working directory changed to
    "Disk$ACC1:[ANONYMOUS.MS-DOS.KERMIT]"
    **** cd cpet***
    Working directory changed to
    "Disk$ACC1:[ANONYMOUS.MS-DOS.KERMIT.CPET]"

```
* get aaareadme.txt
125 RMS transfer started for
"Disk$ACC1:(ANONYMOUS.MS-DOS.KERMIT.CPET)AAAREADME.TXT"
* binary
200 Port command OK
* get cpetkerm.exe
125 RMS transfer started for
"DISK$ACC1:(ANONYMOUS.MS-DOS.KERMIT.CPET)CPETKERM.EXE"
*
```

To retrieve the Macintosh version of Kermit (GTU-KERMIT-SIT.HQX) and the accompanying text file (GTU-KERMIT-SIT.TXT), enter the following series of commands at the asterisk (*). Commands that the user must enter are in ***bold italics***. Responses from the host computer are inthis font:'

```
* cd mac
Working directory changed to "Disk$ACC1:(ANONYMOUS.MAC)"
* cd kermit
Working directory changed to
"Disk$ACC1:(ANONYMOUS.MAC.KERMIT)"
* get gtu-kermit-sit.txt
125 RMS transfer started for
"Disk$ACC1:(ANONYMOUS.MAC.KERMIT)GTU-KERMIT-SIT.TXT"
* binary
200 Port command OK
* get gtu-kermit-sit.hqx
125 RMS transfer started for
"DISK$ACC1:(ANONYMOUS.MAC.KERMIT)GTU-KERMIT-SIT.HQX"
*
```

D. When you have successfully retrieved the appropriate files, log off of the Georgetown host computer and exit FTP using the following commands:

```
* close
* bye
```

Once you have closed your connection to GUVAX, you may have to transfer the files you just retrieved to your personal computer from your mainframe. Procedures for downloading will vary depending on the nature of your mainframe. Be sure that you SET FILE TYPE BINARY when you are transferring the executable file (CPETKERM.EXE or GTU-KERMIT-SIT.HQX, and that you SET FILE TYPE ASCII when you are transferring the accompanying textfile (AAAREADME.TXT or GTU-KERMIT-SIT.TXT).

Once the appropriate Kermit files have been downloaded or copied to your personal computer, place those files in a separate subdirectory or folder. Read the accompanying documentation for Kermit (AAAREADME.TXT or GTU-KERMIT-SIT.TXT) to help you configure the program to your particular system.

Accessing the Catalogue from Remote Sites

The Catalogue resides on a VAX computer owned and operated by Georgetown University. Users outside of Georgetown University may access the database through Telnet or via modem.

Via Telnet

To access the Catalogue via Telnet, log onto a local computer which runs Telnet and enter the following:

telnet guvax3.georgetown.edu

A series of status messages will be displayed, followed by the Georgetown login screen. At the Username prompt enter the name: *CPET*. Press Enter. No password is required. If the username entered is incorrect, the message 'User Authorization Failure' is displayed.

If you *are* prompted for a password, you have probably made one of two mistakes. Either you have misspelled the username (CPET) or you have tried to enter CPET via ftp rather than by telnet.

Via Modem

To access the Catalogue via modem, you must have the following:

- A personal computer configured with a Hayes-compatible modem; and
- A communications package, e.g. VTERM, Procomm, Kermit or the equivalent.

Modem access is available 24 hours a day by dialing:

- (202) 687-2616

Dialing this number connects the user to a modem pool; there are approximately 50 modems in this pool which provide access to Georgetown University's telecommunications network and subsequently to its computing systems. Users can communicate with the VAX computer running the Catalogue at speeds up to 2400 baud.

To establish a connection, perform the following procedure:

A. Using a Hayes-compatible modem, dial (202) 687-2616.

B. When you receive a CONNECT message, press Esc-+ followed by ENTER.

C. At the DIAL: prompt, type *GUVAX*. This will connect you to the terminal server.

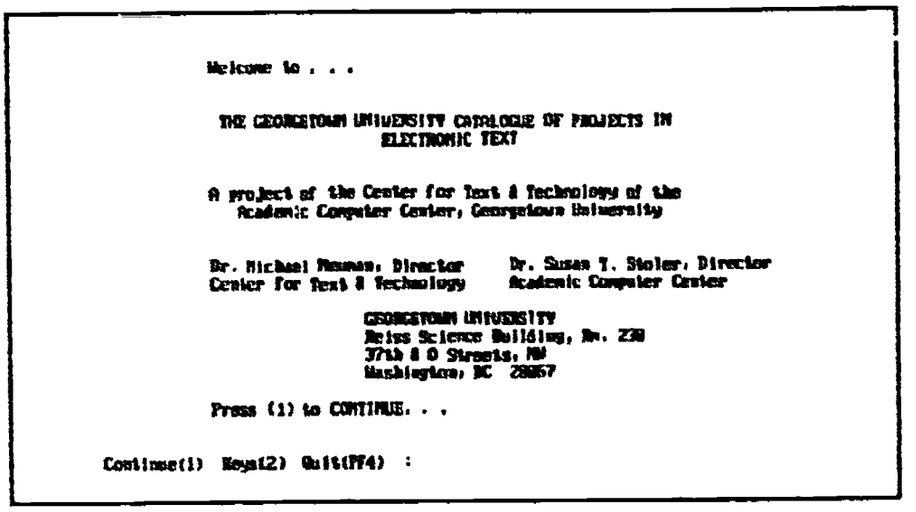
D. At the Local> prompt, type *CONNECT GUVAX3*.

Upon making the modem connection, the Georgetown login screen will then be displayed. At the Username prompt enter the name: *CPET*. No password is required. If the username entered is incorrect, the message 'User Authorization Failure' is displayed.

Once you have logged onto the CPET account, you will be prompted for your electronic mail address. The mail address is necessary if you use the Catalogue to generate reports based on your queries. When you request that a report or list be printed, the report or list is actually output to a file. That file is then sent to your electronic mail address when you log out of the program at the end of your session. Note that if you do not specify an electronic mail address, any reports you have generated will not be sent to you, and will instead be destroyed.

The Initial Catalogue Screens

Once you have logged onto the *CPET* account and specified your mail address, you will automatically be connected to the Catalogue. The initial Catalogue screen is displayed.



Press 1 to Continue.

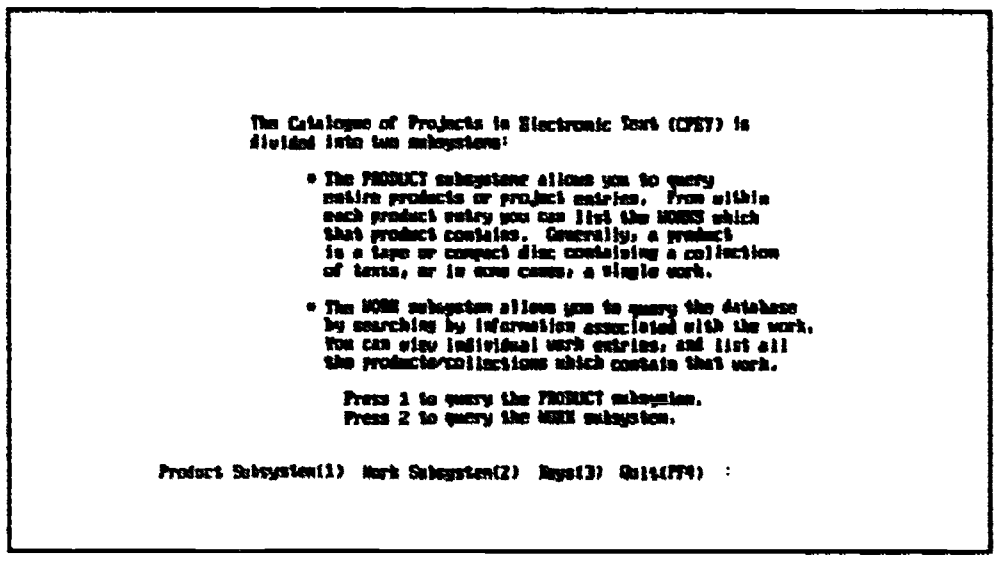
Structure of the Database

Information in the database is divided into two types: products and works. Products are actual academic projects or commercial products which you can purchase or acquire and which contain versions of electronic texts. Works are the actual literary, linguistic or scholarly texts themselves. A single product may include electronic versions of one or more works. For example, the *Thesaurus Linguae Graecae* is a product, and Plato's *Timaeus* is a work included as a part of that collection.

The database is divided into two subsystems which correspond to these two types of information.

- The PRODUCT subsystem allows you to query the database for information about various electronic text products and projects. Each entry is comprised of 8 screens. From within the PRODUCT subsystem you can also view information about WORKs that are associated with that product.

- The WORK subsystem allows you to query information about individual works, i.e. texts. All the information about a particular work appears on a single screen. This information is independent of any particular product entry and will appear as information associated with the work, wherever that work is referred to in the database.

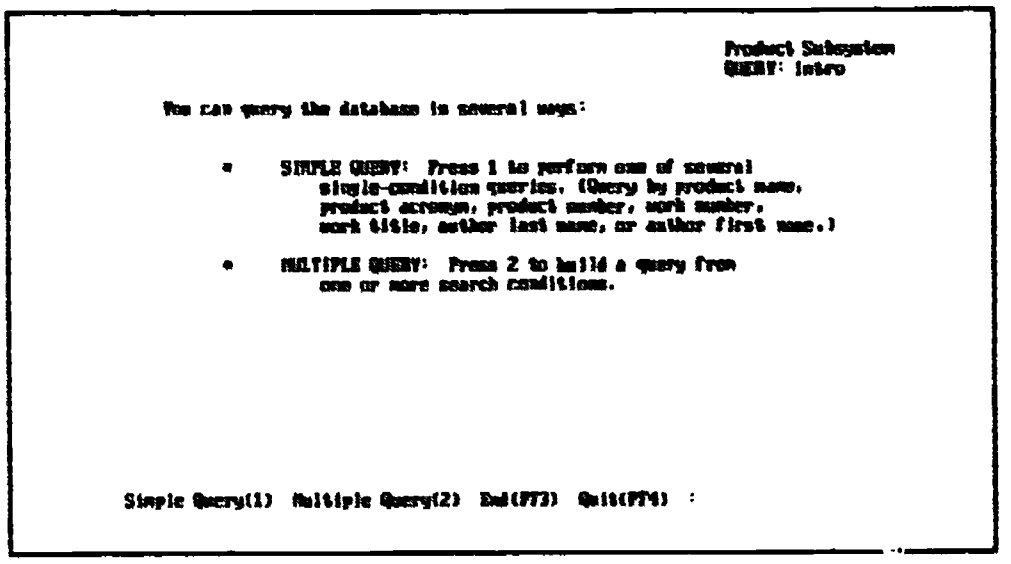


Select 1 to access the PRODUCT subystem. Select 2 to access the WORK subsystem.

The PRODUCT Subsystem

Use the PRODUCT subsystem to search the database for complete product entries which satisfy one or more search conditions. Once you have run a query and generated a list of returns, you can do one of several things:

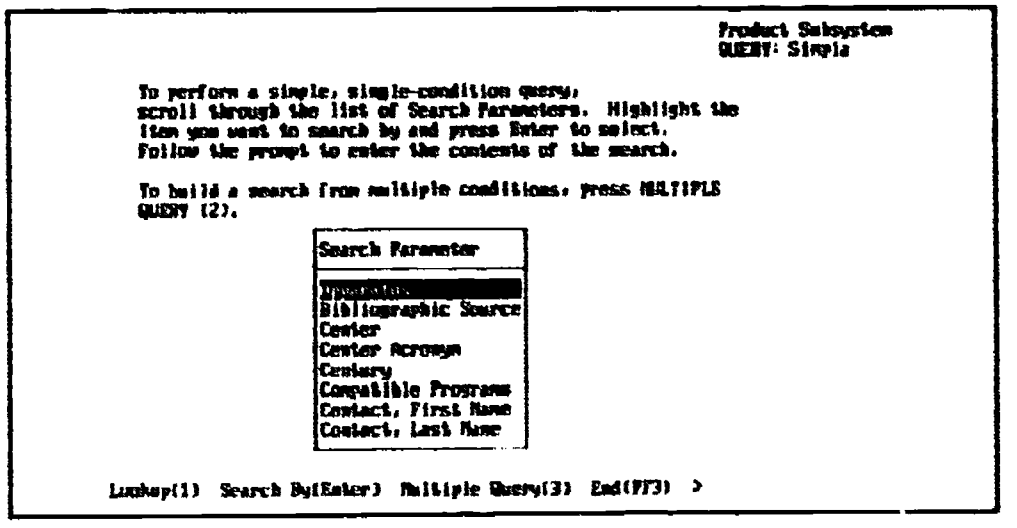
- View each of the entries returned by the query.
- Print the list of query returns.
- Print a single complete entry.



When you enter the PRODUCT subsystem, the **Query: Intro** screen is displayed. Select **Simple Query (1)** to run a search using only one search parameter. Select **Multiple Query (2)** to run a search using from 1 to 5 parameters. Using either a simple or multiple query you can search any of the fields comprising the database except those specifically relating to the work. You can search the WORK subsystem separately. However, you can search within the PRODUCT subsystem for products which include particular works by searching by work number.

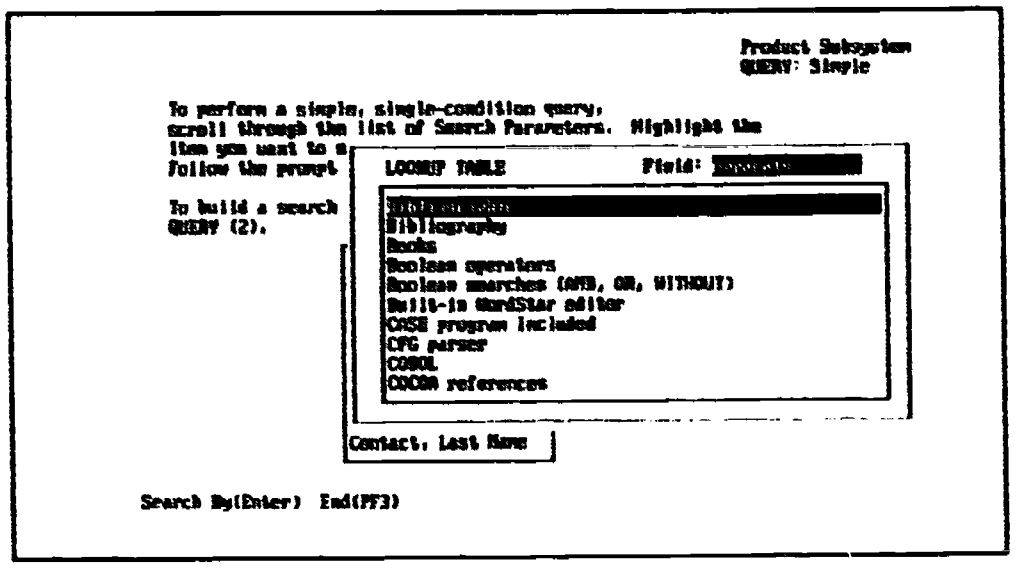
Simple Queries

Once you have selected **Simple Query (1)**, a screen is displayed listing the various fields which may be searched. Using the cursor control keys, or the + (for page up) and - (for page down) keys, scroll through the list of fields to highlight the one by which you want to search. Press **Search By (Enter)** to select that field as your search parameter.



At the prompt, enter the contents of that field for which you are searching. Once you have entered the contents, press Enter.

Note that you can also use the Lookup Feature to specify the search condition.



Highlight the field by which you want to search. Pres. Lookup (1). A list of possible values for that field is displayed. Use the cursor control keys to scroll through the list. Highlight the value you want to search for and press Search By (Enter). The program will perform the search. To leave the lookup list without performing a search, simply press End (PF3).

When you have performed a search, the results are displayed on the QUERY: Return screen. This screen includes a listing of the entries which satisfy your search.

Multiple Queries

Running a multiple query is similar to a simple query except that you must specify whether the various conditions are joined by an AND or an OR. The AND and the OR are called 'operands'. Each pair of conditions in a multiple-condition query is joined by an operand.

- An AND search is a *narrowing* search, i.e., the program will search for entries which satisfy both conditions.
- An OR search is *expansive*, i.e., the program will search for entries which satisfy either of the two conditions.

Think of a multiple s ch as a sentence with several phrases, each joined by either AND or OR. For example, a typical search might look like this:

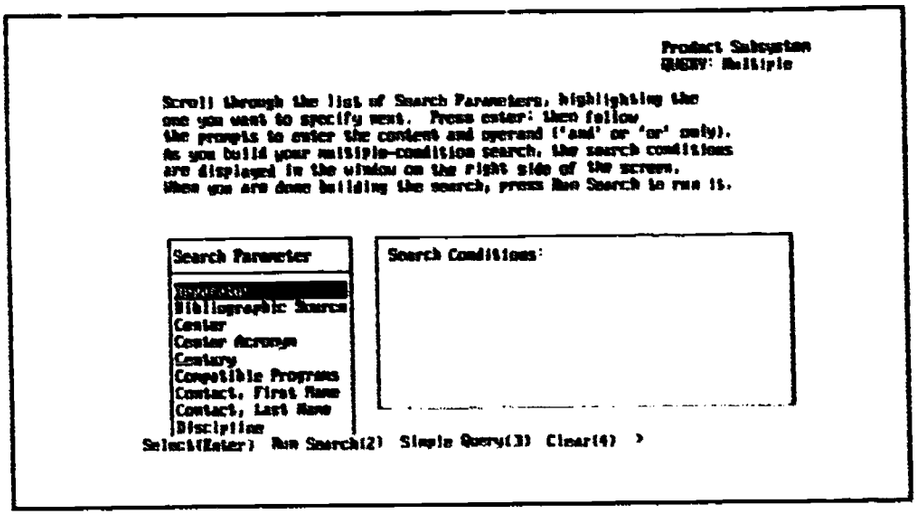
DISCIPLINE = History **AND** DISCIPLINE = Linguistcs **AND** LANGUAGE = French

The program would look for all entries which satisfied all three conditions. When you select the first parameter you will be prompted for only the contents of that parameter (in this case 'History'). When you select parameters two through five, however, you will first be prompted for the operand which precedes that parameter (enter AND or OR) and

then you will be prompted for the contents of that parameter. As you build the search the parameters and operands will be displayed in the Search Conditions field on the right-side of the screen.

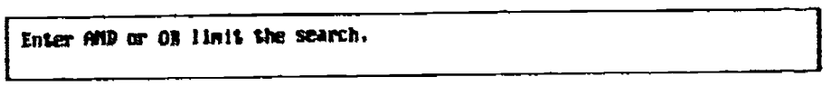
If you build a search with all ANDs, the program will search for entries which satisfy *all* conditions. If you build a search with all ORs, the program will search for entries which satisfy *any* of the conditions.

It is possible to combine ANDs and ORs in a single search. Because you cannot *group* conditions, however, it may be difficult to predict the outcome of such searches. The general rule is that the program gives AND searches higher precedence and OR searches lower precedence. Therefore it will seek to satisfy the AND conditions, and then will look for the OR conditions. The safest bet is, therefore, to run either pure AND or OR searches.



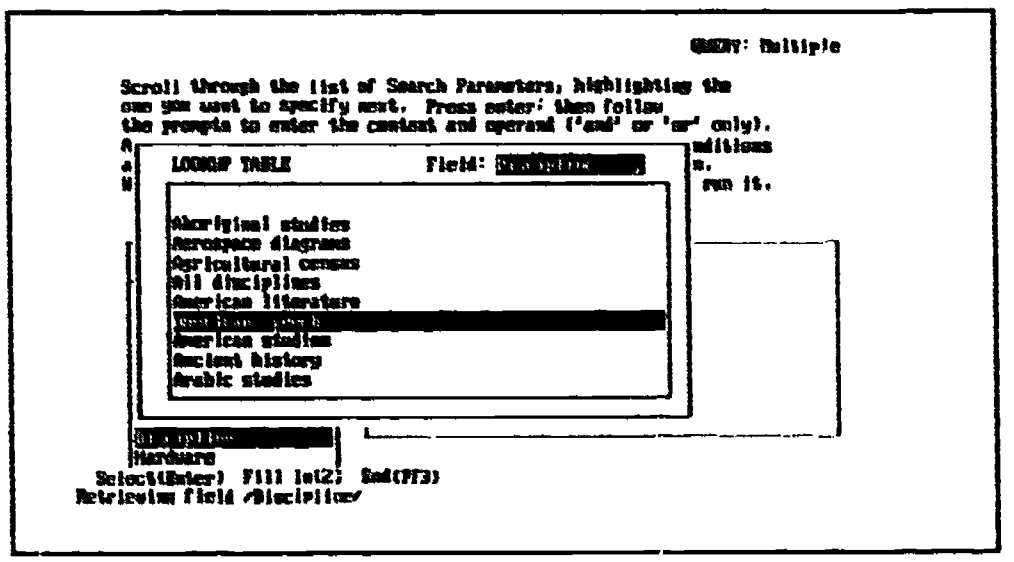
Note that you can build a search from one to five conditions. To select the first parameter, scroll through the list of fields using the arrow keys to highlight the desired field. Press Enter. At the prompt, enter the contents of that field for which you would like to search. Press Enter. Note that the display returns to the multiple query screen and the search parameter you have just specified is displayed in the box marked "Search Conditions:". This enables you to keep track of the various parameters you have selected.

To choose a second search parameter, scroll through the list of fields again, pressing Enter to select the desired field. Note that you can reselect search parameters several times. For example, you could search for entries where DISCIPLINE = History or DISCIPLINE = Biblical Studies. Before being prompted for the contents of the second parameter, however, the program will prompt you to enter either 'AND ' or 'OR', as the operand which will link parameters one and two.



Finally, you will be prompted to enter the contents of the second search parameter. Repeat this process until you have specified all the desired search parameters (up to five).

Note that when you select each parameter, the program will automatically display a lookup list.



Scroll through the list, highlight the value for which you want to search and press **Select (Enter)**. If there is no lookup list available for that field, or if you select **Fill In (2)** from the lookup list, then a prompt will be displayed, asking you to type in the value for which you want to search. If you want to use wild cards in your search, you should use the **Fill In** option.

When all search parameters have been specified, select **Run Search (2)**. The program will run the search and display the results in a list.

Wildcards

In both the PRODUCT and WORK subsystems, you can use wildcards to specify a single character or a string of characters within a text string. Use the following wildcards when specifying a query parameter:

%	for any string of 0 or more characters
_ (underscore)	for any single character

For example, a typical query parameter using wildcards might look like this:

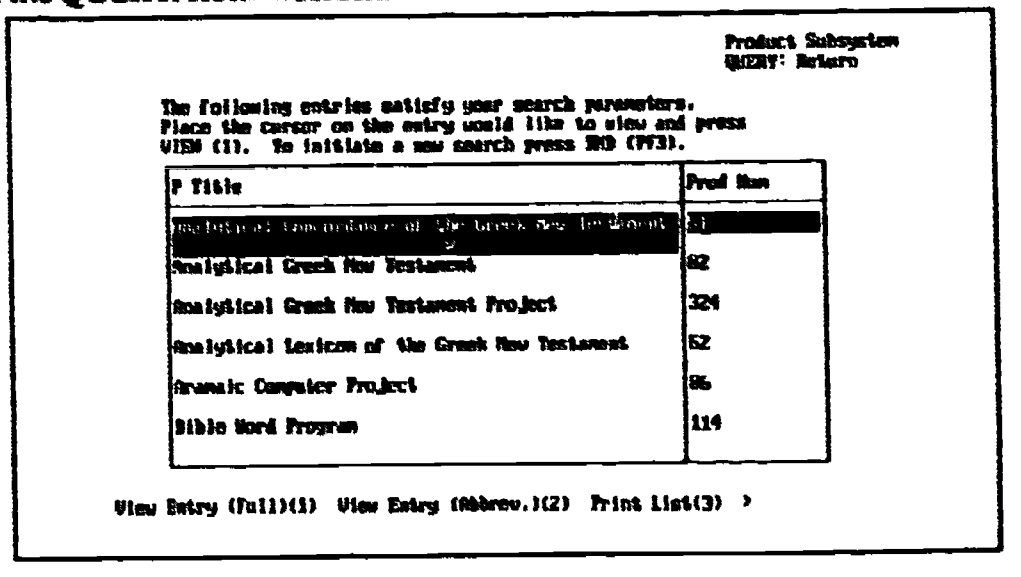
Title = English%.

Certain rules apply to the use of wildcards when specifying query parameters:

- You CANNOT use wildcards for INTEGER or DATE fields. These fields require exact matches. Use the Lookup feature to construct the query and ensure an exact match.
- A wildcard CAN NEVER be the first character in the search string.

Query Returns

Once you have run either a simple or a multiple query, the program displays the results on the **QUERY: Return** screen.



From this screen you can perform several functions:

- To VIEW a particular entry in the list, scroll through the list, highlighting the entry you want. Then select **View Entry Full (1)**. The program will display the first screen of the selected entry. Pressing **Continue** will enable you to walk through the entire entry. Once the entire entry has been viewed, the screen returns to the **QUERY: Return** listing.

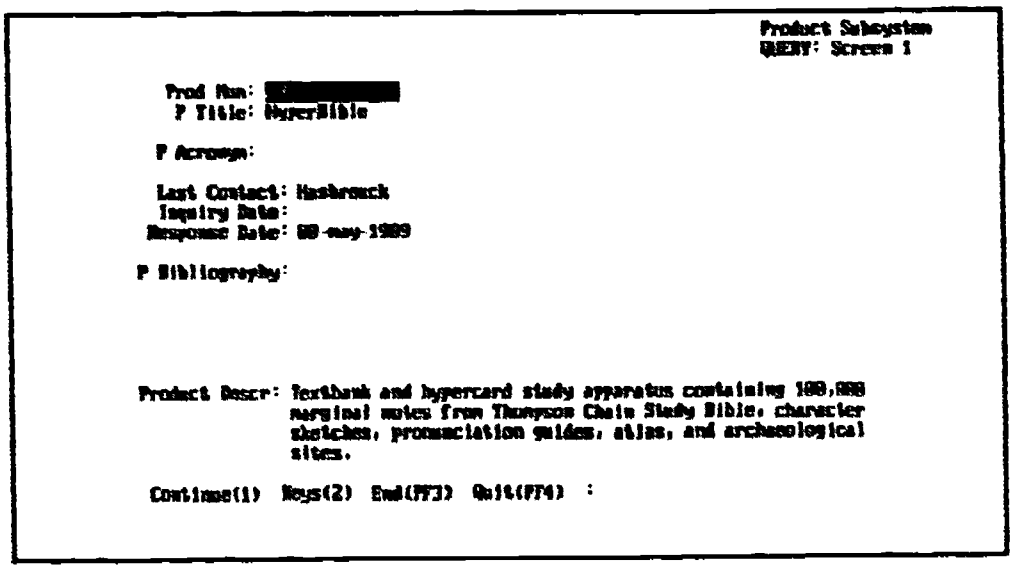
- To VIEW an abbreviated, one-screen version of the product entry, press **View Entry (Abbrev.) (2)**.

- To PRINT a LIST of returns, select **Print List (3)**. The program will automatically output to a file. This file is sent to your electronic mail address automatically when you log out of the program. Once the report has been generated, the screen returns to the **QUERY: Return** listing.

- To PRINT a FULL ENTRY, scroll through the list, highlighting the entry you want to print, and select **Print Full Entry (4)**. The program will then display a special report screen. You can cancel the report at this point by pressing **End (PF3)**. To go ahead and generate the report, select **Report (1)**. The report generator creates a file which is sent to your electronic mail address automatically when you log out of the program. Once the report has been generated, the screen returns to the **QUERY: Return** listing.

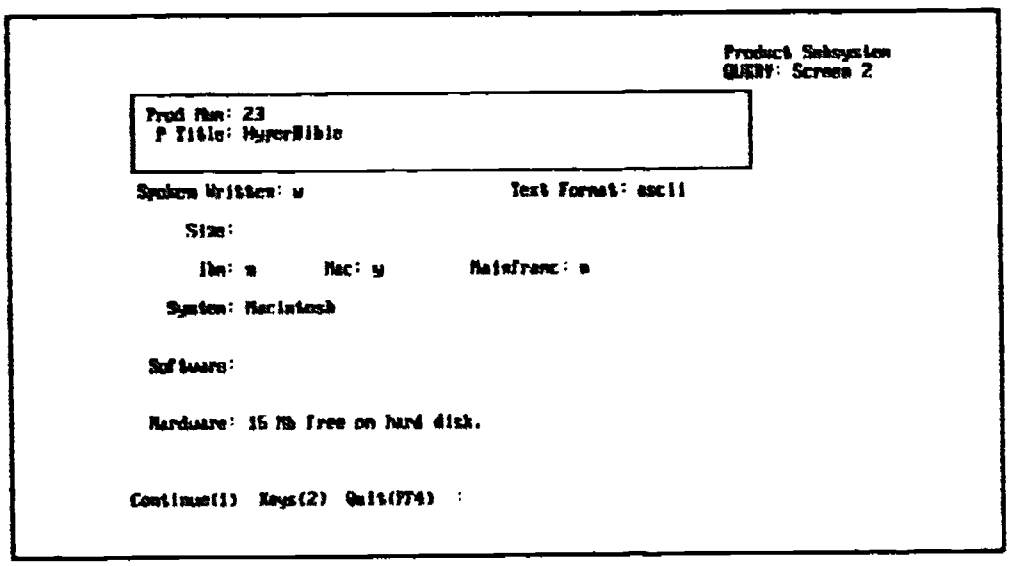
Viewing a Complete Product Entry

A complete product entry is comprised of eight or more screens. When you elect to view a complete entry, the first screen of the entry is displayed.

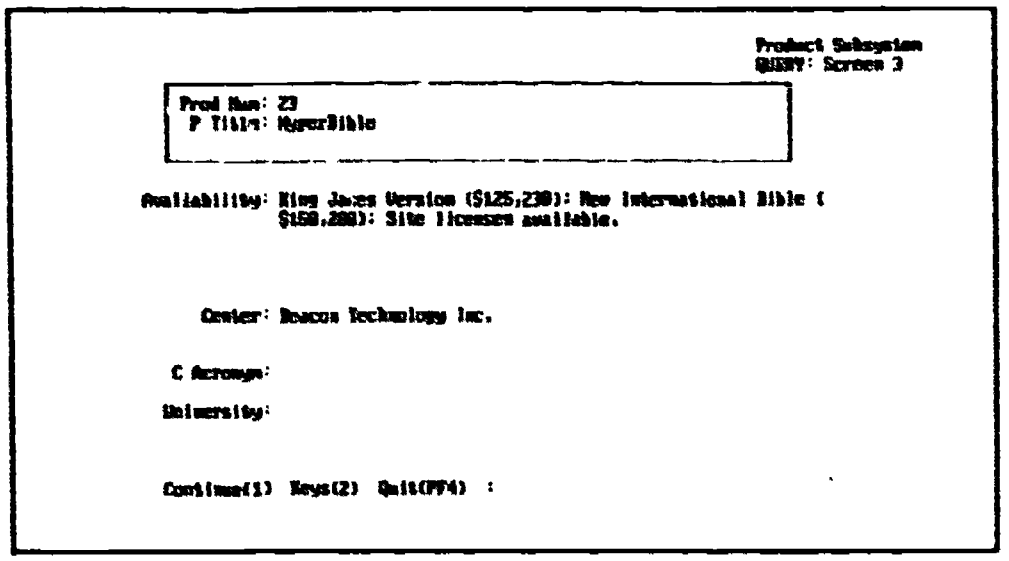


You can leave this screen, without viewing the rest of the product entry by pressing **End (PF3)**. If, however, you elect to view the next screen, you must scroll through the complete entry in its entirety.

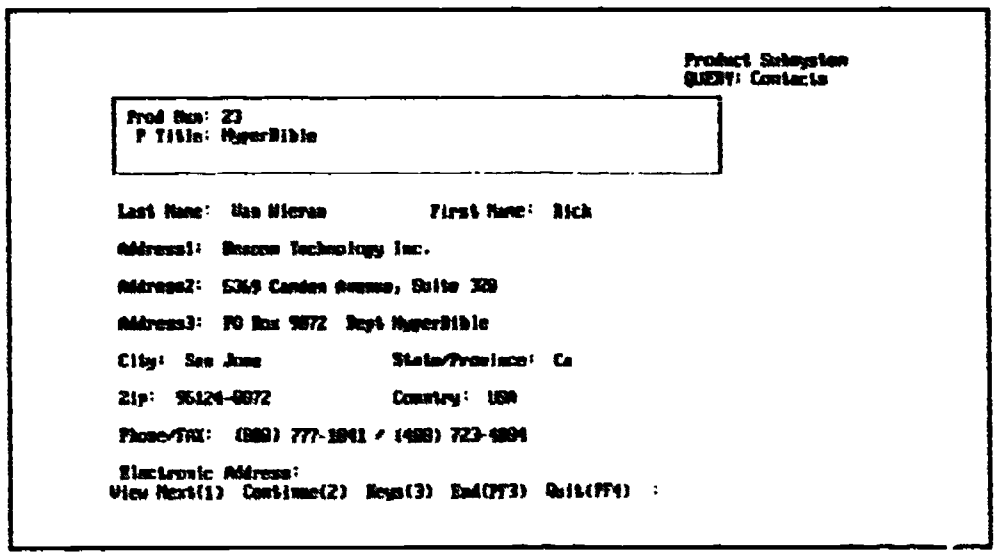
To continue on to the next screen, press **Continue (1)**.



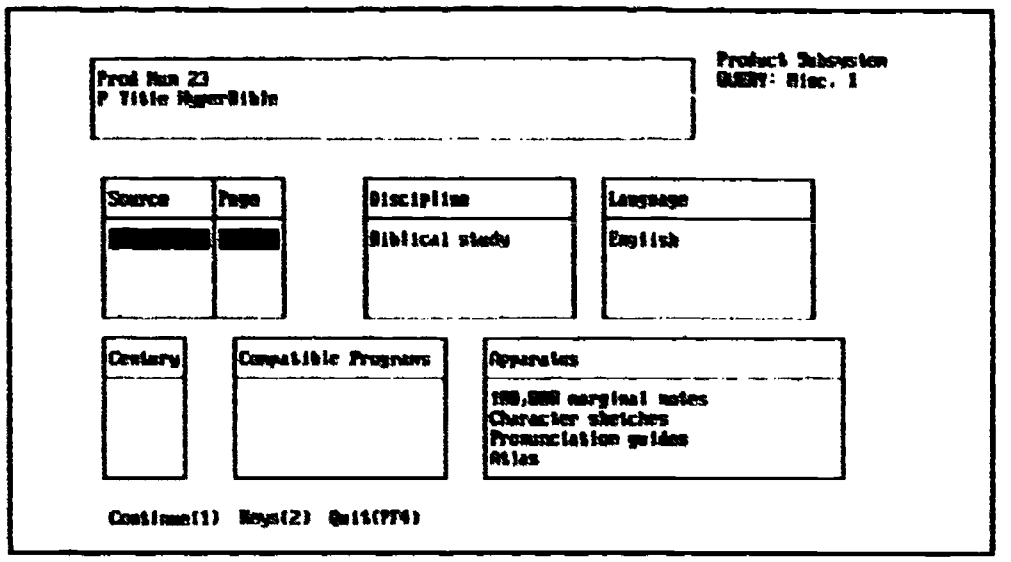
The second basic screen of information is displayed. Press **1** to continue.



The third basic screen of product information includes information on the center and/or university where the project was undertaken, or where the product was made. Press 1 to continue.



The program now displays information on one or more contact persons from whom the product or electronic texts may be obtained. To view information on other contact persons, if they exist, press View Next (1). To stop viewing contact names and go on to the next screen, press Continue (2).



The next screen displays a series of tablefields. Each tablefield lists specific information about the product/project.

- **Source:** This tablefield refers to the book or publication where information about this product was obtained. This is largely for internal Georgetown use.

- **Discipline:** This refers to the academic discipline associated with this product, or the discipline for which it was designed.

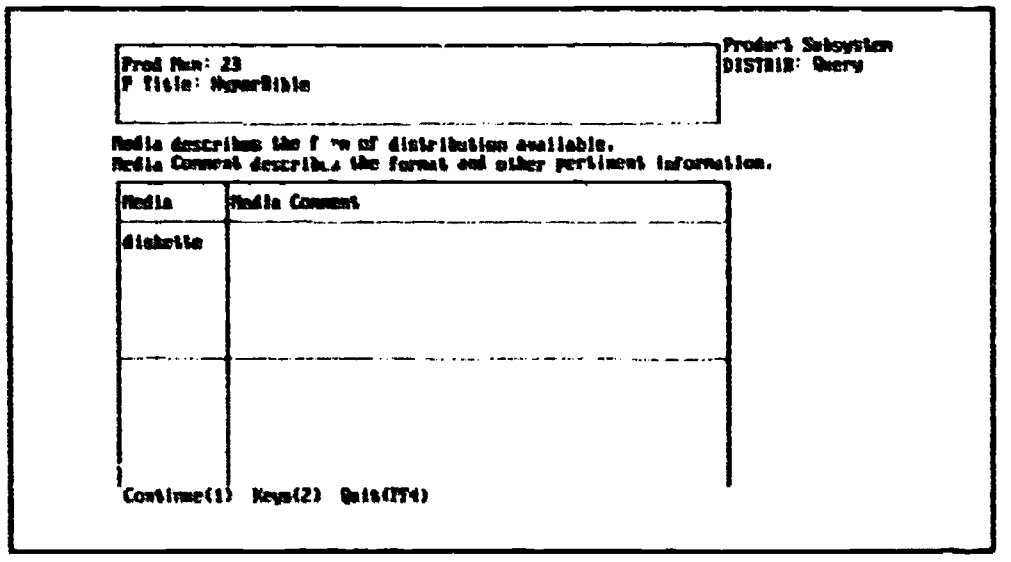
- **Century:** Century indicates the years represented by the texts in this collection. Centuries are indicated by integers. For example, the 20th Century is indicated by the number '20'. The 5th Century BCE, however, is indicated by the number '-5'.

- **Compatible Programs:** These refer to the various software packages and text analysis programs with which this product was designed to be used, or with which it is capable of being used.

- **Apparatus:** Apparatus refers to the various tools included with the product to facilitate the analysis of the texts included. For example, it might include morphological dictionaries, indices or other features which would make the product more accessible.

Tablefield lists may occasionally be longer than the list displayed. Use the tab key to place the cursor on the tablefield you want to view. Then use the cursor control keys or the (+) or (-) keypad keys to page up or down the list.

To leave this screen and move on to the next, press **Continue (1)**.

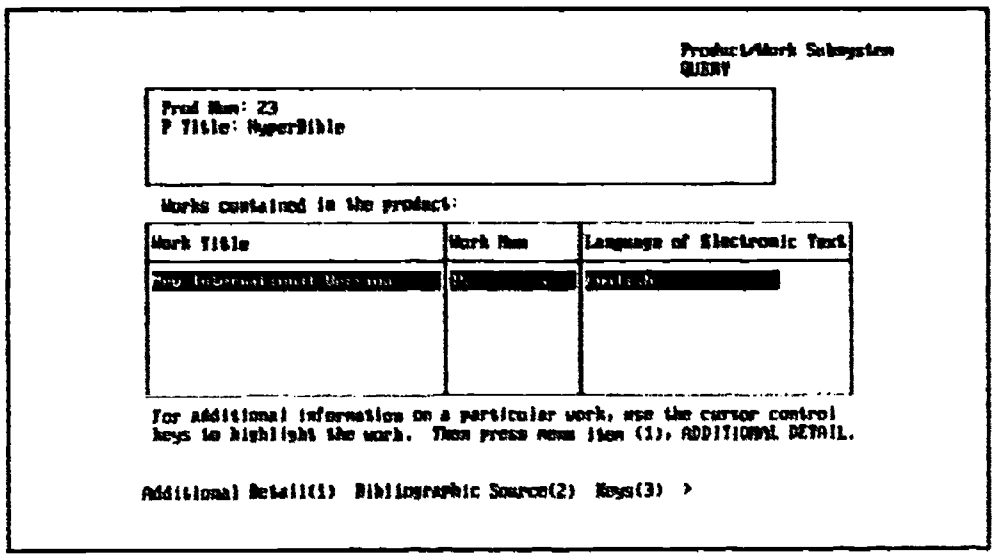


The program then displays information about how the product is distributed. Media refers to the format (CD, tape, diskette, etc.). Media Comment will often indicate how many CDs, diskettes or tapes are included, or information on tape format, etc. Press 1 to continue. If the list extends beyond the screen, use the tab key or cursor control keys to page through the list.

Upon exiting this screen, you will be prompted as to whether you would like to view information about the specific information about the WORKs which are included as texts in this collection at this time.

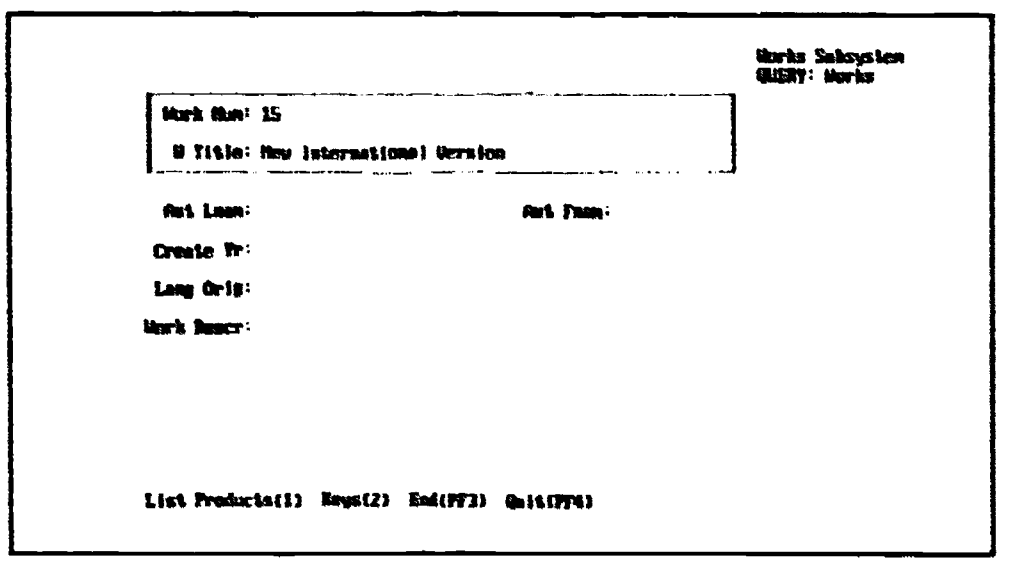
- To leave this entry and return to the list of query returns, enter N.
- To go on to view WORKs data, enter Y.

If you elect to view WORKs information, the program will then display a list of all the works (i.e. texts) which are part of the product collection.

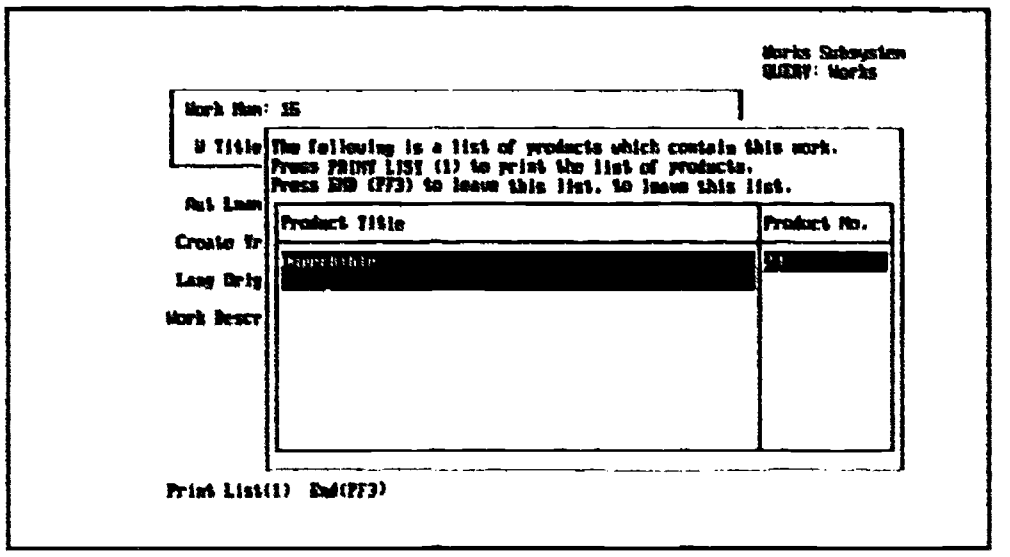


From this screen you can view detailed information about each work, and/or view bibliographic information about the versions used in creating the electronic version of that particular text.

To view additional detail about a particular work, scroll through the list to highlight the desired work and press 1 for **Additional Detail.**



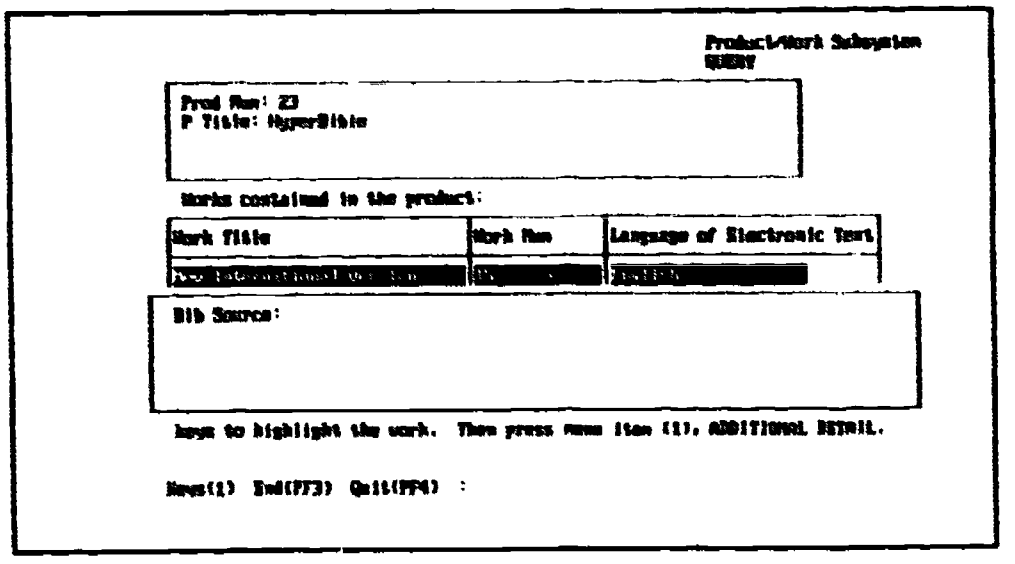
Note that you can list all the other products which might contain this work by pressing **List Products (1).**



The list of products may itself be used to generate a report. When you press **Print List (1),** a report is generated and output to a file. This file is then sent to your specified electronic mail address upon logout. To leave the listing of products, press **End (PF3).**

To leave the work detail screen and return to the list of works, press **End (PF3).**

To display bibliographic data, scroll through the list of works, highlight the desired work and press **Bibliographic Source (2).**



The bibliographic data specify the particular version of the text which was used to create the electronic version of the text which was included in this particular product. This information, therefore, is specific to the PRODUCT rather than the WORK. For example, one edition of Shakespeare's *Julius Caesar* may have been used in creating the electronic version of a particular product, but a different edition may have been used for a different product. To leave this screen, press **End (PF3).**

Once the complete entry has been viewed, and once information on the works included has been viewed, if desired, the program returns to the list of query returns. You can then go on to view another entry, print a report, or press **End (PF3)** to initiate a new query.

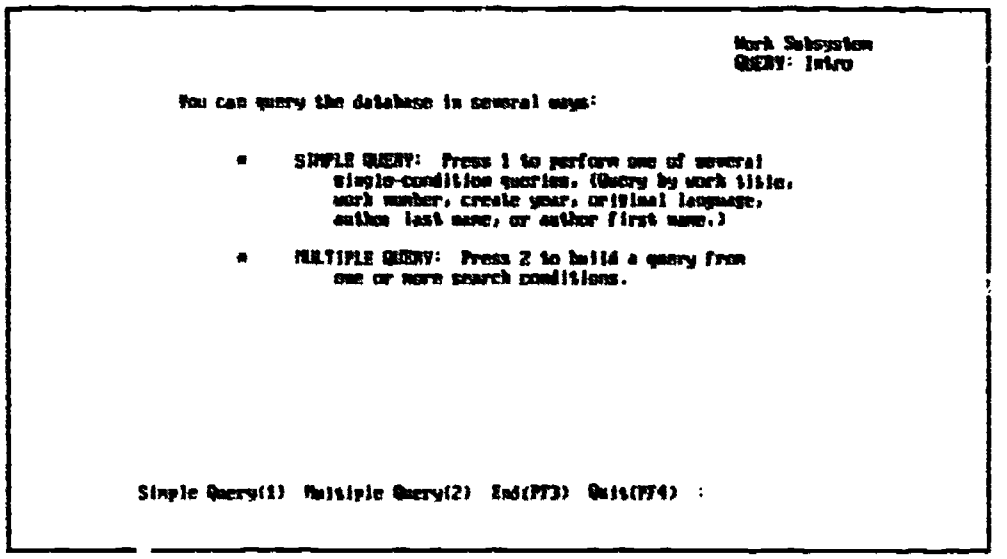
The WORK Subsystem

Use the WORK subsystem to search the database for information about individual texts. All the information about a particular work appears on a single screen. This information is independent of any particular product and will appear as information associated with that work or text, wherever that work is referred to in the database. Enter the WORK subsystem by selecting **Work Subsystem (2)** from the initial Catalogue screen.

The WORK subsystem can be used to search the database for works which satisfy one or more search conditions. Note that to search for all the products containing a particular work, search for the work, then press **List Products** to list all the products containing that work. Once you have run a work query and generated a list of returns, you can do one of several things:

- View each of the entries returned by the query.
- Print the list of query returns.
- Print a single complete entry for a particular work.

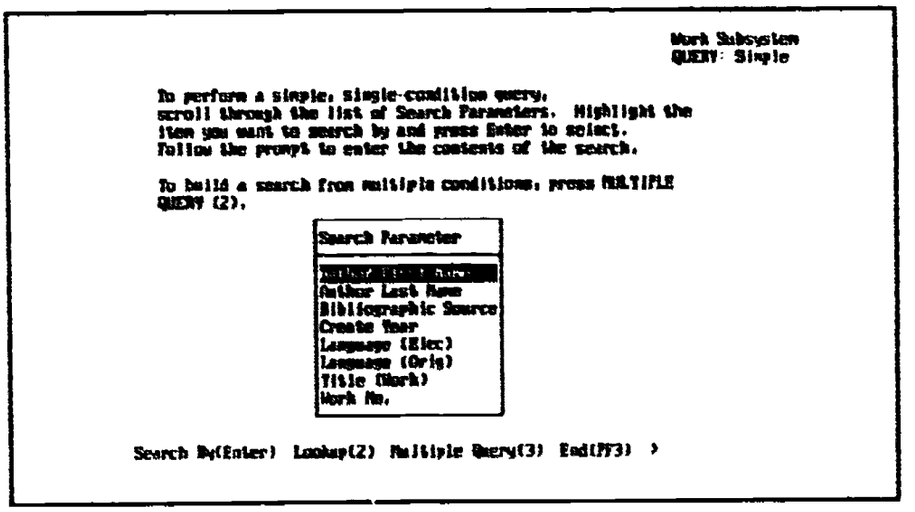
When you enter QUERY mode, the **Query: Intro** screen is displayed.



Select **Simple Query (1)** to run a search using only one search parameter. Select **Multiple Query (2)** to run a search using from 1 to 5 parameters. Using either a simple or multiple query you can search any of the fields comprising the database except those specifically relating to the product. You can search the PRODUCT subsystem separately. However, you can search within the WORK subsystem for products which include particular works by searching by work number.

Simple Queries

Once you have selected **Simple Query (1)**, a screen is displayed listing the various fields which may be searched. Scroll through the list of fields to highlight the one by which you want to search. Press **Search By (Enter)** to select that field as your search parameter.



At the prompt, enter the contents of that field for which you are searching. Once you have entered the contents, press **Enter**.

Note that you can also use the Lookup Feature to specify the search condition. Highlight the field by which you want to search. Press **Lookup (1)**. A list of possible values for that field is displayed. Use the cursor control keys to scroll through the list. Highlight the value you want to search for and press **Search By (Enter)**. The program will perform the search. To leave the lookup list without performing a search, simply press **End (PF3)**.

When you have performed a search, the results are displayed on the **QUERY: Return** screen. This screen includes a listing of the entries which satisfy your search.

Multiple Queries

Running a multiple query is similar to a simple query except that you must specify whether the various conditions are joined by an AND or an OR. The AND and the OR are called 'operands'. Each pair of conditions in a multiple-condition query is joined by an operand.

- An AND search is a *narrowing* search, i.e., the program will search for entries which satisfy both conditions.
- An OR search is *expansive*, i.e., the program will search for entries which satisfy either of the two conditions.

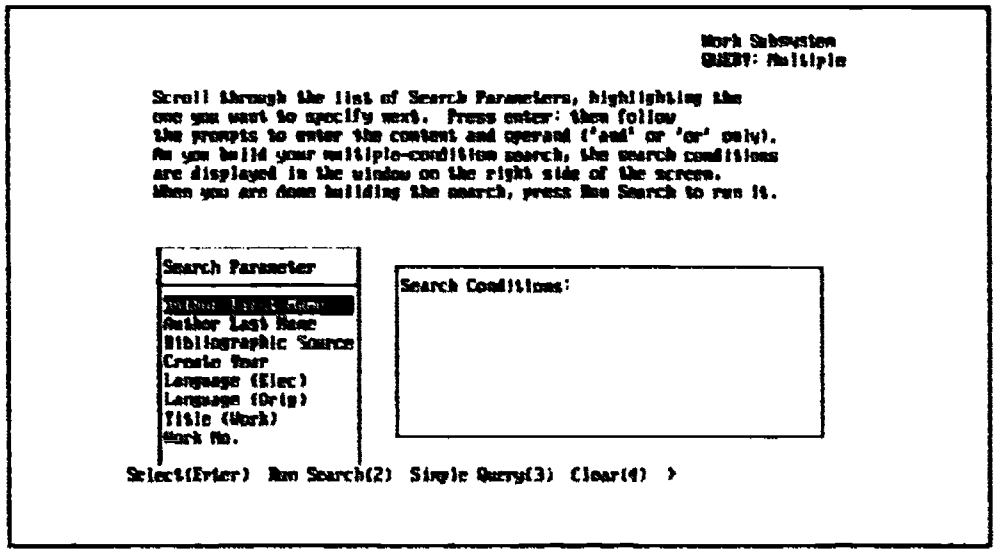
Think of a multiple search as a sentence with several phrases, each joined by either AND or OR. For example, a typical search might look like this:

LANG ORIG = German **AND** LANG ELEC = English **AND** CREATE YEAR = 1545

The program would look for all entries which satisfied all three conditions. When you select the first parameter you will be prompted for only the contents of that parameter (in this case 'German'). When you select parameters two through five, however, you will first be prompted for the operand which precedes that parameter (enter AND or OR) and then you will be prompted for the contents of that parameter. As you build the search the parameters and operands will be displayed in the Search Conditions field on the right-side of the screen.

If you build a search with all ANDs, the program will search for entries which satisfy *all* conditions. If you build a search with all ORs, the program will search for entries which satisfy *any* of the conditions.

It is possible to combine ANDs and ORs in a single search although it is not recommended. Because you cannot *group* conditions, it may be difficult to predict the outcome of such searches. The general rule, however, is that the program gives AND searches higher priority and OR searches lower priority. Therefore it will seek to satisfy the AND conditions, and then will look for the OR conditions. The recommended option is to run either pure AND or OR searches.



Note that you can build a search from one to five conditions. To select the first parameter, scroll through the list of fields using the arrow keys to highlight the desired field. Press Enter. At the prompt, enter the contents of that field for which you would like to search. Press Enter. Note that the display returns to the multiple query screen and the search parameter you have just specified is displayed in the box marked "Search Conditions:". This is to enable you to keep track of the various parameters you have selected.

To choose a second search parameter, scroll through the list of fields again, pressing Enter to select the desired field. Note that you can reselect search parameters several times. For example, you could search for entries where LANG ORIG = English or LANG ORIG = German. Before being prompted for the contents of the second parameter, however, the program will prompt you to enter either 'AND ' or 'OR', as the operand which will link parameters one and two. Finally, you will be prompted to enter the

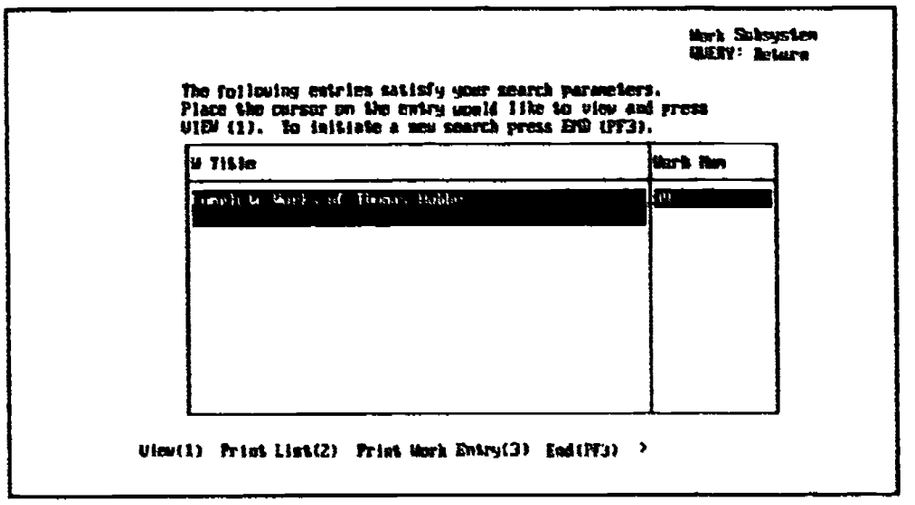
contents of the second search parameter. Repeat this process until you have specified all the desired search parameters (up to five).

Note that when you select each parameter, the program will automatically display a lookup list if available. Scroll through the list, highlight the value for which you want to search and press **Select (Enter)**. If there is no lookup list available for that field, or if you select **Fill In (2)** from the lookup list, then a prompt will be displayed, asking you to type in the value for which you want to search. If you want to use wild cards in your search, you should use the **Fill In** option. See page 9 for information on using wildcards.

When all search parameters have been specified, select **Run Search (2)**. The program will run the search and display the results in a list.

Query Returns

Once you have run either a simple or a multiple query, the program displays the results on the **QUERY: Return** screen.



From this screen you can perform several functions:

- To VIEW a particular entry in the list, scroll through the list, highlighting the entry you want. Then select **View (1)**. The scenario is the same as the work detail screen displayed from the PRODUCT subsystem. Once the entry has been viewed, the screen returns to the **QUERY: Return** listing.

- To PRINT a LIST of returns, select **Print List (2)**. The program will generate a report file. The file is automatically mailed to your electronic mail address when you log out of the program. Once the report has been generated, the screen returns to the **QUERY: Return** listing.

- To PRINT a WORK ENTRY, scroll through the list, highlighting the entry you want to print, and select **Print Work Entry (3)**. The program will then display a special report screen. You can cancel the report at this point by pressing **End (PF3)**. To go ahead and print the report, select **Report (1)**. The program will generate a report file. The file is automatically mailed to your electronic mail address when you log out of the program. Once the report has been generated, the screen returns to the **QUERY: Return** listing.

APPENDICES

APPENDIX A:
Field Type and Length

The following table lists the various fields contained in the catalogue, indicating the type of field and the field length.

Field Name	Field Type	Field Length
Address1 (contact)	text	50
Address2 (contact)	text	50
Address3 (contact)	text	50
Apparatus	text	35
Author First Name (work)	text	20
Author Last Name (work)	text	20
Availability	text	300
Bib Source	text	300
Center	text	100
Center Acronym	text	10
Century	integer	-128 to 128
City (contact)	text	20
Compatible Programs	text	20
Contact First Name	text	20
Contact Last Name	text	20
Country (contact)	text	15
Create Year (work)	integer	up to 32,767
Discipline	text	20
Electronic Address (contact)	text	40
Hardware	text	50
IBM	text	1
Inquiry Date	date	
Language	text	20
Language (Elec) (work)	text	20
Language (Orig) (work)	text	20
Last Contact	text	40
MAC	text	1
Mainframe	text	1
Media	text	10
Media Comment	text	300
Operating System	text	100
P Bibliography	text	300
Phone (contact)	text	40
Product Acronym	text	10
Product Description	text	300
Product Num	integer	up to 2.147 million
Product Title	text	100
Response Date	date	
Size	text	50

Field Name	Field Type	Field Length
Software	text	100
Source	text	10
Source Page	integer	up to 32,767
Spoken/Written	text	1
State (contact)	text	10
Text Format	text	10
University	text	50
Work Description	text	400
Work Number	integer	up to 2.147 million
Work Title	text	30
Zip (contact)	text	15

APPENDIX B:
Electronic Documentation

This manual is also available in an electronic ascii version for quick and easy on-line reference. The ascii version of this manual is identical to this manual except that it does not include illustrations. To receive a copy of the manual in electronic form, send a mail message to:

James A. Wilderotter II
Project Assistant
Center for Text and Technology
Academic Computer Center
Reiss Science Building, Room 238
Georgetown University
Washington, DC 20057
(202) 687-6096

BITNET: wilder@guvax
Internet: wilder@guvax.georgetown.edu

Please include a current Bitnet or Internet address.