

Set of ADVANCED-BASIC programs for literature storage and retrieval with minicomputers

RAY T. STERNER and CHARLES P. BREIDENSTEIN
Denver Wildlife Research Center, Denver, Colorado 80225

Most researchers store reprints of scientific papers for planning projects and preparing reports. Typically, however, the indexing and filing work needed to maintain traditional card-reference systems has made such files prone to disuse. The increased availability of small in-house computers with peripheral memory devices now makes automated storage and retrieval of this information practical.

This paper describes a reprint-reference system adaptable to a variety of minicomputers having auxiliary mass storage capabilities (cf. Sterner & Breidenstein, Note 1). The software is based on a modified version of the Uniterm Coordinate Index System (Broadhurst, 1962), and is offered as an example of literature storage and retrieval with minicomputers.

Description. Preparation of the reprint-reference system requires that each article be assigned an identifier, multiple keywords, and a set of keyword codes. Keywords are established by the user and describe the informational content of reprints, and each keyword is designated by a number code for purposes of memory storage. A software program is used to enter the identifier, keyword codes, and select reference data for each reprint onto memory disks. These are termed reference disks, and serve as the main storage location for reprint information. Subsequently, another program is used to duplicate the identifier and keyword codes on a second disk (the accession disk) which facilitates the rapid search and retrieval of these data.

In order to initiate a search, the user loads a retrieval program into computer memory, specifies the type-of-search (AND-OR) and keyword codes to be scanned, and obtains a printout of the reprint identifiers plus keyword codes and/or reference data which satisfy these search criteria. The user then selects articles of interest from reprint-storage files.

Figure 1 is a schematic diagram of the reprint-reference system. As shown, the system is comprised of three elements: system hardware, system software, and reprint storage.

System Hardware. The currently used equipment consists of a Wang 2200B computer with ADVANCED-BASIC language, CRT display, 16K memory, input keyboard, Wang 2243 removable-flexible disk drive, and

30-cps printer (Qualterm X-145). The Wang 2243 has three Memorex 651 flexible disk drives. Each drive provides 262,144 bytes of storage divided into 1,024 sectors on 19-cm-diam magnetic disks. Separate disks and drives are used for program statements, accession data, and reference data.

To operate the system, the user simply mounts the respective program, accession, and reference disks into the 2243 disk drives (cf. Figure 1) and "loads" a designated program into computer memory via the input keyboard. The user then enters information as directed by the program. Although the system is designed for the Wang 2200B and 2243 disk units, the system could be easily converted to other computers and disk devices.

System Software. A set of five ADVANCED-BASIC programs, plus reference- and accession-disk information, comprise the software portion of the system. Specifically, the programs are: INTREF (INTialize REFerence disk), BLDREF (BuiLD REFerence disk), BLDACC (BuiLD ACCession disk), LSTREF (LiST REFerence data), and RETRIV (RETRIEVe reference data). Programs INTREF, BLDREF, and BLDACC involve storage operations, whereas LSTREF and RETRIV provide data retrieval functions. All of the programs are question-answer routines, requiring little user knowledge of either program or computer functions. A short description of each routine follows:

(1) INTREF (19 statements): This program "initializes" the reference disks. That is, it assigns a two-character alphanumeric identifier (e.g., A1, B2) to the reference disks, as well as addresses to the disk-storage sectors (i.e., 1 to 1,024). The identifier is input by the user. Program INTREF allows sequential use of multiple reference disks.

(2) BLDREF (60 statements): BLDREF is the main reference-storage program of the system, and governs the input of reference data into disk storage. The following data are entered for each reprint: (a) a five-character alphanumeric reprint identifier, (b) a maximum of five three-digit numeric keyword codes, and (c) 215 alphanumeric characters of author, title, journal, volume, year, and page data. With the Wang 2243 disk drive, a total of 999 citations can be stored per reference disk, and disks can be added as required.

(3) BLDACC (66 statements): This program duplicates the reprint identifier and keyword codes of each reference-disk citation onto an accession disk. The accession disk thus provides for rapid, efficient search and retrieval of reference information. Accession data for 6,000-8,000 references can be stored on a single Memorex-651-type disk. Although multiple accession disks can be used, the system is basically designed for searches involving a single accession and multiple reference disks (i.e., search and retrieval times will increase greatly with multiple accession disks).

We thank Mr. Ivan S. Alkes for programming assistance. Use of the equipment cited in this report was selected solely on the basis of availability. Reference to trade names does not imply U.S. Government endorsement of commercial products.

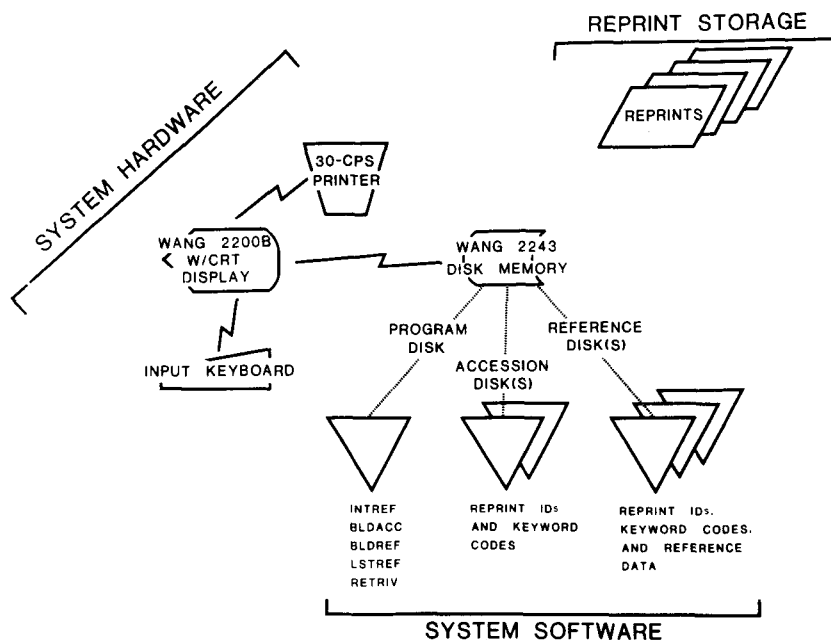


Figure 1. Schematic diagram of reprint-reference system. Note the three types of removable disks utilized with the Wang 2243 disk memory—program disk, accession disk(s), and reference disk(s).

(4) LSTREF (36 statements): Program LSTREF provides a printout of the information stored on reference disks. This is useful for book-keeping operations associated with the system. The reprint identifier, keyword codes, and reference data for all of the citations stored on a disk are listed as 126-character/line output.

(5) RETRIV (161 statements): RETRIV governs the search and retrieval of accession and reference data. Two search parameters are used with the routine: (a) type of search (i.e., AND-OR) and (b) keyword codes (i.e., ≤ 5 , three-digit numbers). The type-of-search parameter designates either a conjunctive (AND) or disjunctive (OR) search strategy, whereas the keyword codes specify the search topic. The program scans the accession disk for keyword codes which match the selected strategy (e.g., codes 2 AND 33 AND 181, codes 2 OR 33 OR 181). Input includes: (a) entry of the word "AND" or "OR," (b) entry of one to five three-digit keyword codes, and (c) entry of the word "YES" or "NO" (i.e., the user specifies whether or not retrieved references are to be printed). Output includes: (a) a list of the reprint identifiers which match the search criteria and (b) an optional list of the reference data for each retrieved identifier.

Reprint Storage. Storage of reprints requires that each be assigned an identifier (i.e., a five-character alphanumeric identifier: one-letter file label and four-digit reprint number). Reprints are then stored numeri-

cally under each alpha-file label; new reprints are added to the file(s) as desired.

Costs. Costs of operating the system are limited mainly to charges associated with external mass storage (i.e., about \$5/disk with our system) and salaries for personnel required to code and enter reference data.

Limitations. Retrieval times vary as a function of file size, search criteria, and print speed (e.g., detection and print times for a five-category search of a 300 reference file have averaged 8 sec/citation). Storage and retrieval is limited to ≤ 5 keyword codes per reprint and ≤ 349 separate keyword categories, limitations easily modified by redimensioning the BLDREF, BLDACC, and RETRIV programs.

Availability. A user's manual and program listing are available without charge from Library Services, Denver Wildlife Research Center, Building 16, Federal Center, Denver, Colorado 80225.

REFERENCE NOTE

1. Sterner, R. T., & Breidenstein, C. P. *Reprint-storage and reference-retrieval system for the Wang 2200B computer with 2243 removable-flexible disk memory*. Denver Wildlife Research Center Technical Report, in press.

REFERENCE

- BROADHURST, P. L. Coordinate indexing: A bibliographic aid. *American Psychologist*, 1962, 17, 137-142.