Time-Shared Computing

Implications for Medical Libraries*

BY CHARLES J. AUSTIN, *Director of Computer Services and Assistant Professor of Preventive Medicine and Comprehensive Health Care*

University of Colorado Medical Center
Denver, Colorado

ABSTRACT

Many library data processing systems are unresponsive to the needs of librarians because of the necessity to batch-process transactions in a computer center. Such systems tend to be report-oriented rather than information-oriented with resultant problems in the timeliness of information. Time-shared computing permits multiple users to process jobs simultaneously through on-line interaction with the computer. Such systems offer to the librarian the advantages of immediate access to information, costs shared with other users, and direct man-machine interaction. This tutorial paper describes time-shared systems with applications in the library. Problems concerning the cost and present state-of-the-art of time-sharing are discussed.

INTRODUCTION

THE importance of automation in the library of the future has been underscored by many authors. King has discussed the promise of automation: "There is a spectrum of functions which automation can improve, ranging from purely routine processing chores, such as circulation control and the maintenance of serial records, to advanced information retrieval techniques" (1).

Sir Frank Francis, Director and Principal Librarian of the British Museum, discussed the urgency of the problem during the Brasenose Conference on the Automation of Libraries: "Thus the essential development of the library service, which it is impossible to contemplate by present methods, becomes entirely possible by using mechanized methods. The question is not, therefore, why or what for, but how soon can this change be brought about. Without it our great libraries are in danger of atrophying through sheer bulk" (2).

Despite the many references to the potentials of automation, much of the promise of computers in libraries remains just that—promise rather than realization. Many reasons for this could be and have been cited, including the reluctance of librarians to accept change and the complexity of library information processing requirements. While not denying that these factors exist to a greater or lesser extent depending on the library involved, the major thesis of this paper is that most automatic data processing systems are not totally responsive to the needs of librarians, primarily because they are "report-oriented" rather than "information-oriented." This hypothesis leads one to a discussion of on-line, time-shared systems as a tool for the modern library.

This paper is tutorial, describing first the more traditional batch-processing computer systems, and then the newer time-sharing systems and their implications for medical libraries.

Background

In order to understand the recent development of time-shared computer systems, it is necessary to review briefly the development of automatic data processing in this country.

Punched-card equipment was developed in 1888 by Dr. Herman Hollerith for use by the U. S. Census Bureau for tabulation of the 1890 census. Punched-card or unit-record systems were used extensively during the period 1900–1950 by governmental and industrial organizations—primarily for accounting and statistical

* Presented as part of a general session on "National and Regional Library Systems" at the Sixty-seventh Annual Meeting of the Medical Library Association, Denver, Colorado, June 11, 1968.

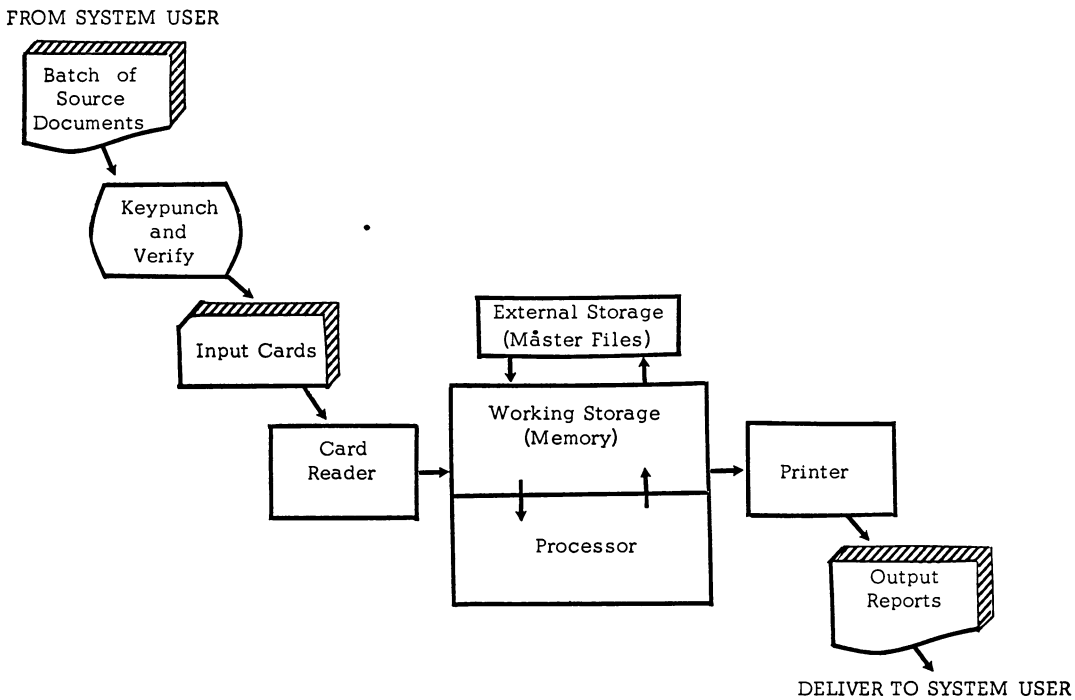

FIG. 1.—Typical batch-processing system

applications. These systems operated entirely on a *batch-processing* basis; that is, source documents providing input data for the systems were collected, punched, and processed on a periodic, sequentially-scheduled job basis. Hence, output information would be delivered to the "customer" from the data-processing department on a periodic basis—daily, weekly, monthly, and so forth.

Digital computers became commercially available about 1950, and computer technology can be described as developing in three phases. *First generation computers* (1948–1960) employed vacuum tubes in their electronic circuitry and were employed strictly in a batch-processing mode, with the computer replacing some of the functions formerly performed by tabulators and other punched-card devices. The development of transistors led to the construction of *second generation computers* (1960–1965) which greatly improved operating speed and efficiency. However, machines of this class continued to be operated mostly as batch processors, although some limited experimentation with on-line, man-machine communication was beginning. *Third generation computers* (1965–) employed new miniaturized circuit elements (referred to as solid logic technology), and emphasis was placed in the design of these systems on flexible input, output, and storage devices to permit more on-line utilization of the equipment. Precise definitions of terms such as "on-line" and "time-sharing" are given later in the paper.

Typical Batch-Processing System

Figure 1 shows the flow of information through a typical batch-processing system. Input documents are collected periodically and keypunched for processing by the computer at a scheduled time. Internal computer files will normally be set up in sequential order according to an identifying code, and input transactions will be sorted into the same sequence for matching to these master files.

The computer center which operates in this batch-processing environment will process jobs according to a master schedule showing the time when each job should be run. If a given job on a given day runs longer than anticipated, if input data is delayed for any reason, or if

machine down-time occurs, the schedule is disrupted and must be readjusted.

The major problems faced by users of batch-processing computer systems are usually related to scheduling—the fact that they cannot process information on a demand basis. Computer centers are frequently faced with problems of job queuing and schedule variations from day to day as described above. It should be noted, however, that good management practices in the computer center coupled with cooperation from users can help to minimize these problems. It should also be noted that batch-processing will be the *most efficient* method for processing many types of jobs, and on-line access will not be required for every task in the library. Examples of library systems employing both on-line and batch-processing principles are given in a later section of this paper.

Library Batch-Processing Applications

The great majority of library data processing applications developed to date are batch-processing systems, many of these using punched-card equipment (particularly in smaller libraries). Examples of these batch-processing systems include:

(1) Printed Catalogs.
(2) Lists of Serial Publications.
(3) Circulation Lists and Statistics.
(4) Budget and Financial Reports.
(5) Mailing Lists.

In general, these systems are "report-oriented" as mentioned above. The systems often suffer from the fact that reports are not current and are not readily accessible to a large enough number of people who may wish to reference them at the same time. Currency of bibliographic information is an essential ingredient of most library systems.

Description of Time-Sharing Systems

Basic Definitions

The computer literature dealing with on-line or time-sharing systems has suffered severely from lack of precision and agreement on definitions of commonly used terms. An article by Colilla written in 1966 presents good technical definitions of some of the more important terms (3). A similar, but less technical, set of definitions is given as follows:

Basic Terminology*

On-Line: A human user or device is serviced by a computer through direct communication with an operating program. For human users, this involves "conversational" interaction.

Real-Time System: A system in which program execution must satisfy a particular operational response time which could range from microseconds to minutes.

Time-Sharing: That operation of a computer which permits more than one user to operate the system simultaneously (or apparently simultaneously) in such a way that each is or can be unaware of the use of the facility by others.

Processor: That component of a computer system which can execute a set of instructions serially (one at a time).

Multiprogramming: That operation of a (serial) processor which permits the execution of more than one program in such a way that none of the programs needs to be completed before another is started or continued.

Multiprocessing: A computer system consisting of more than one processor, each having access to common storage area of data.

Note that a user may be on-line to a computer without operating in a time-shared system although this would usually not be economical.

Time-sharing systems will require some combination of multiprogramming and/or multiprocessing capabilities in order to be able to serve simultaneously several users with different jobs (programs) to be executed.

Typical Time-Sharing Configuration

A typical machine configuration for a time-sharing system is shown in Figure 2.

Several on-line users share the central computer through remote terminals; signals to and from these terminals are serviced by a multiplexor communications unit. The remote terminals may be any required distance away from the computer, communicating with the computer via data transmission lines (the cost, of course, increases with distance).

* An article by Rosenberg provided valuable information in compiling these definitions (4).

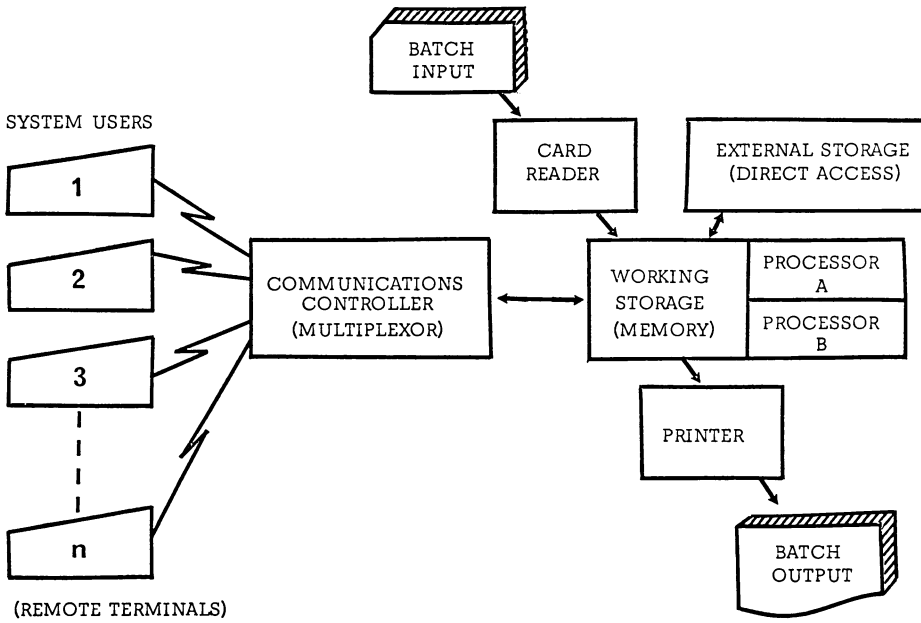

FIG. 2.—Typical time-shared system

The central computer also has standard input-output devices for running batch-processing jobs to fill up the computer when terminals are not utilizing the full capacity of the machine. These are referred to as "background" jobs.

In order to provide on-line service to users, the configuration must have direct-access storage units so that records on data files can be immediately referenced from the remote terminals.

Finally, the typical configuration has two processors, each of which can operate with multiprogramming—that is, several programs can be executed simultaneously in both of the two processors.

This type of configuration requires a very sophisticated system of scheduling and monitoring jobs in the computer. This system, usually referred to as an "operating system," is provided by the computer manufacturer and is one of the most critical elements of time-sharing.

Desirable Characteristics of Time-Sharing Systems

An effective time-sharing system must be responsive to the demands placed on it by multiple users. Some of the desirable characteristics of such a system include:

(1) Simultaneous use by multiple terminals—in order to be economical the system must be able to be used by several customers simultaneously.
(2) Instantaneous response—the slightest delay in responding to incoming signals from a terminal will disturb the user and could result in inefficient use of the system.
(3) Independent operation—each user must be completely unaware that anyone else is sharing the system; the terminal should represent the user's "own private computer."
(4) General purpose—the system should minimize the restrictions placed on individual users with regard to the types of jobs they may run on the computer.

Storage Allocation in Time-Shared Systems

Drattell describes four primary ways of allocating internal computer storage in a time-sharing system (5).

Users may operate the system with *reserved storage* for their own exclusive use, with file protection built into the operating system in such a way that no other user may access that part of the file which is reserved. Thus, for example, the chief librarian might reserve stor-

age for his own personal file of library work load and budget statistics.

A *data bank* is a file available to all users of the system for reference purposes, but which is updated by a central supplier. An on-line library catalog, for example, could be referenced by circulation, reference, and acquisitions personnel, but would only be updated by catalogers.

Some files may be available to a restricted group of users on a *data exchange* basis. Thus individual files may be "borrowed" by others who are qualified to use the information. For example, the serial record file might be available to the reference librarian when required, but would not be readily available to all users of the system.

A *cooperative file* is a pool of data which can be both interrogated and updated by the same group of users. For example, a work-in-process file would be freely available for updating and interrogation by all members of the acquisitions and cataloging staff.

Present State-of-the-Art of Time-Sharing

Time-shared computer systems are still largely developmental, and only limited applications have been made to date of this kind of system. R. L. Patrick, in a recent review of time-sharing development, takes a somewhat critical point of view: "Most of our time-sharing systems were designed without a thorough analysis of the problems to be solved, the services to be offered, or the costs to be encountered. Since these are the only systems that we have available to us, the remaining choice is to exercise a moderate degree of caution and select from those available systems the ones which solve problems you actually have" (6).

It must be remembered, however, that time-sharing systems are still in their infancy, and the amount of developmental work in this field (obvious to a reviewer of the computer literature) portends well for the future of user-oriented systems.

There are still many problems to be overcome before time-sharing systems will be widely used. These problems may be characterized as: technical, sociologic, and economic.

(1) Technical problems are found both in the equipment and programming systems required for time-sharing. Time-sharing "hardware" requires greater reliability than ever before, since multiple users are affected by a single equipment failure. The problems of "software" or programming systems are even more severe. Complex job-scheduling systems and easy-to-use conversational languages must be provided by the manufacturer. The time-sharing software systems developed to date are still very primitive and will require continual development.

(2) Sociologic (or human-engineering) problems relate to methods of insuring efficient, and not just casual, use of the system by personnel not trained in the art of programming. The ideal terminal device has not yet been developed, and user education will represent a major investment for the prospective time-sharing customer.

(3) Economics constitute the third problem area to be considered. Time-sharing is still very expensive. The high cost of data-communications equipment and random–access storage are major contributors. Unless the system is efficiently scheduled, excessive overhead costs may result from equipment not used.

The prospective user of time-sharing systems is advised to investigate these potential problems carefully before a final decision is reached.

Implications of Time-Sharing for Medical Libraries

General Effects on Library Operations

An obvious and important implication of time-sharing for medical libraries is the ability of a single library to share a central computer with other users, and hence cut down the costs of the service. However, in considering the possible use of a commercial time-sharing service, the small library with limited funds may find the total costs of the service to be prohibitive. A major factor will be the high cost of direct-access storage which is required for almost all library applications. Even a library with relatively low day-to-day activity will require large files for storage of catalogs, serial records, and circulation records. The advantages of sharing will be greatly enhanced when several libraries can begin to share *common files*—catalogs and serial lists, for example. Thus, the cooperative file approach described above seems

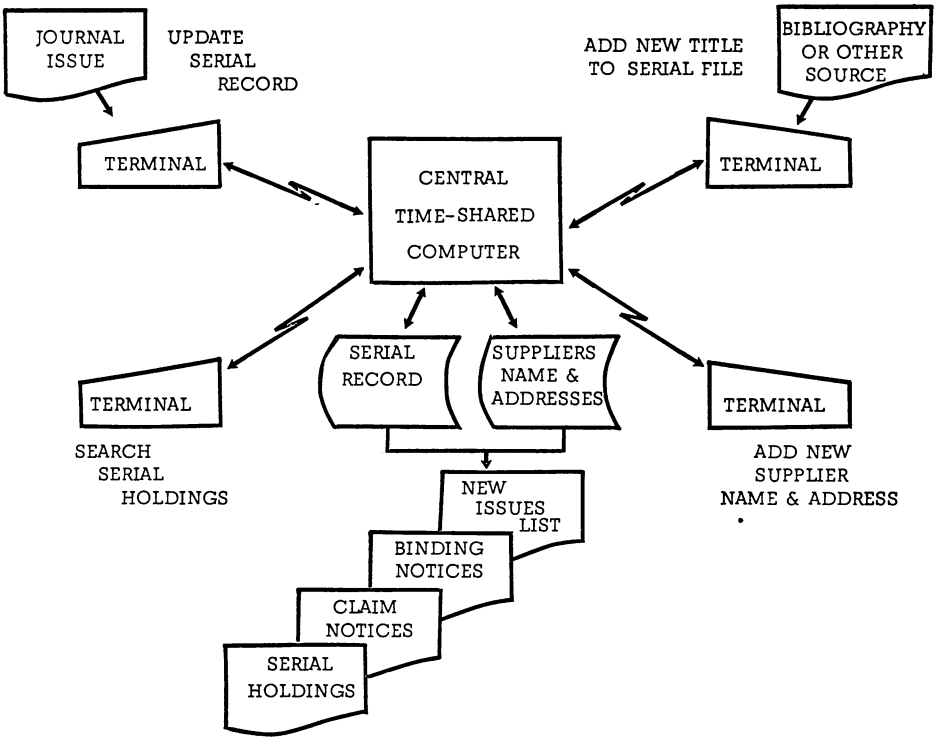

FIG. 3.—Serials system

to offer promise if several libraries can agree on standards for classification and cataloging. The centralized cataloging activities of the National Library of Medicine and the Library of Congress will play an important role in such developments.

A second major effect of time-sharing on medical libraries will result from computer systems becoming more information-oriented and less report-oriented. Information will be readily available on demand; the librarian will no longer have to wait "... until the printed catalog is delivered next Monday." The computer files will be updated directly in the normal course of business, and data will be immediately accessible to all qualified users.

An on-line system will require more active involvement of the librarian in a direct man-machine interaction. The librarian will become an integral part of the system, rather than someone who sits on the periphery of the system waiting for a report and hoping that it will be correct. This interaction should lead to a better understanding of the computer by the librarian.

In summary, the environment of conversational interaction in which the computer becomes a direct tool in the work process should lead to more effective and meaningful library systems.

Some Typical On-Line Systems

The implications of time-sharing discussed above can best be illustrated by some examples.

Figures 3 through 5 describe schematically three potential library systems which could be implemented on a time-shared computer: a Serials System, a Circulation System, and an Acquisitions and Cataloging System. Note that all three of these conceptual systems employ both on-line and batch-processing techniques. Those tasks involving updating and interrogation of the files are generally performed on-line through remote terminals (e.g., searching serial holdings, adding new records to the circulation file, subject cataloging of newly received monographs). The three systems also require the production of periodic reports which are best handled on a batch-processing basis, using information added to the files by the remote ter-

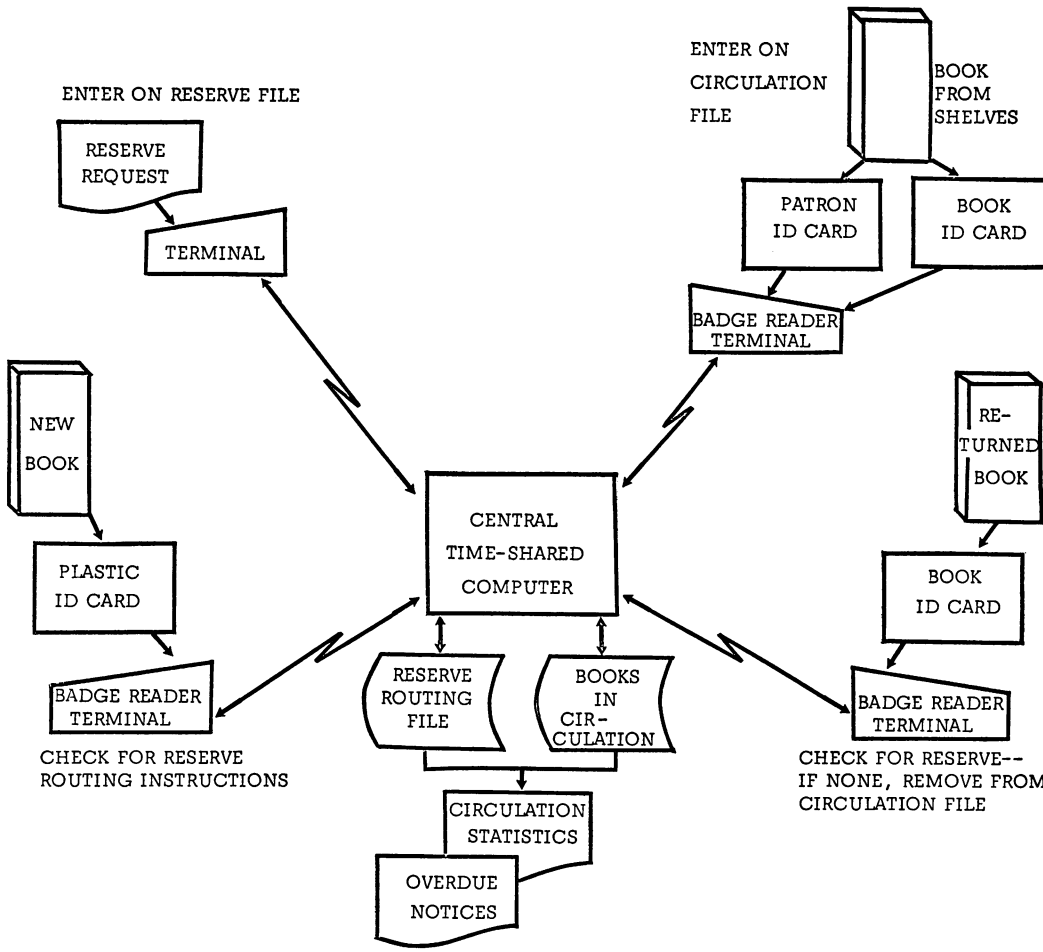

FIG. 4.—Circulation system

minals (e.g., Circulation Statistics, Book Catalog, Binding Notices). These batch-processing jobs would be run as background to the on-line tasks performed through the terminals. Master files are stored in the computer on direct-access storage devices; note that the master file of Suppliers Names and Addresses is used by two of the systems—Serials System and Acquisition and Cataloging System.

Operating Library Time-Sharing Systems

As mentioned earlier, time-sharing systems are still largely developmental, and for this reason not many library applications have yet been described in the literature. The work done by the State University of New York in developing a regional medical library network has been described by Mr. Pizer (7).

The Illinois State Library, as a department of the office of Secretary of State, has successfully implemented an on-line circulation system. The computer files are on-line five and one-half days each week for handling 1,300 to 2,500 transactions per day (charge-outs, renewals, reservations, and returns). The transactions are drawn from a collection of nearly 500,000 items. The success of the system is described as follows: "After nearly one year of full-time operation, the system has exceeded the expectation of the library staff. Work loads have been shifted and the circulation records staff has been reduced fifty percent, resulting in staff availability for many projects and therefore better and faster library service to the residents of the state of Illinois" (8).

The Illinois State Library System described

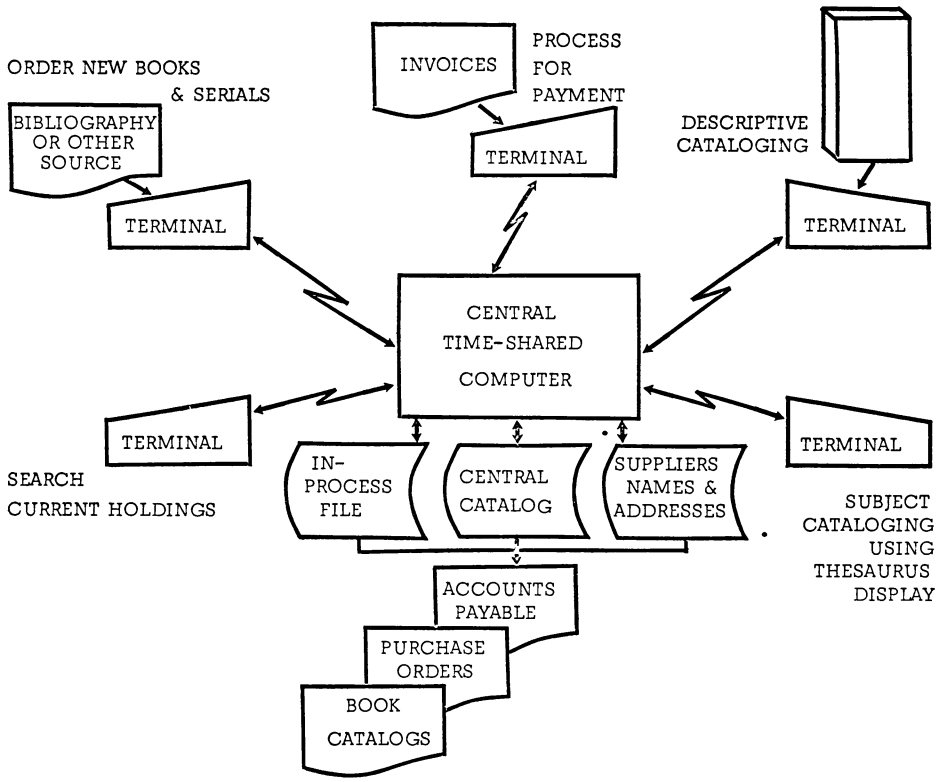

FIG. 5.—Acquisitions and cataloging system

above represents a good example of a well-defined, self-contained system which lends itself well to on-line, time-shared computing capabilities. Searching of the computer files in this system involves looking for a specific item or subscriber's name. Considerably more difficult is the problem of catalog searching by computer. One attempt to develop this kind of capability is the BOLD (Bibliographic On-Line Display) system designed by the staff of System Development Corporation in Santa Monica.

BOLD is a document storage and retrieval system designed to allow users to search for information in a file of coded and computer-stored document abstracts (9). The system permits browsing through the collection, examining documents stored under each major subject category. If the user is not sure of correct procedures to follow in searching, he can receive help and instructions which are displayed on his terminal device. Search requests are stated in natural English. The system was operated experimentally using abstracts of approximately 6,000 documents obtained from the Defense Documentation Center (DDC), with the documents grouped according to the DDC Classification system.

SUMMARY

The developing field of on-line access to time-shared computers opens new possibilities for effective use of such systems in medical libraries. Batch data-processing systems are generally report-oriented, and the resultant delay in providing completely current bibliographic data is a major drawback to the use of these systems in libraries.

On-line systems involve direct, conversational interaction between the human user and the machine, so that the computer system becomes information-oriented and a more useful tool in the work process. Time-sharing of a central computer with cooperative files or data banks permits multiple users within one library or from several libraries to have direct access to large stores of information.

Time-sharing systems are still in an early and developmental stage, and many problems

still must be solved. These include technical problems with hardware and software, sociologic or human-factor problems, and the present high cost of time-sharing.

A few libraries are beginning to experiment with time-sharing. As the field develops and the above problems are solved, the impact of time-sharing in medical libraries will depend directly on the extent of active participation by librarians in the development of effective systems.

REFERENCES

1. KING, G. W. Automation and the Library of Congress. In The Growth of Knowledge, Readings on Organization and Retrieval of Information. New York, Wiley, 1967, p. 105.
2. FRANCIS, SIR FRANK. Libraries Call in Computers To Catch Up on Knowledge. In Brasenose Conference on the Automation of Libraries. London, Mansell, 1967, p. xv.
3. COLILLA, R. A. Time-sharing and multiprocessing terminology. Datamation. March 1966, p. 49–51.
4. ROSENBERG, A. Time-sharing: A status report. Datamation. Feb. 1966, p. 66–77.
5. DRATTELL, A. Big computers for small users. Business Automation. Aug. 1966, p. 26–30.
6. PATRICK, R. L. Time-sharing tally sheet. Datamation. Nov. 1967, p. 42–47.
7. PIZER, I. H. The State University of New York Computerized Biomedical Information Resource. In Brasenose Conference on the Automation of Libraries. London, Mansell, 1967, p. 151–164.
8. HAMILTON, R. E. Illinois State Library computer system. Wilson Libr. Bull. March 1968, p. 721–722.
9. BORKO, H. Utilization of On-Line Interactive Displays, SP–2575. Santa Monica, Calif., System Development Corp., Aug. 22, 1966.

ADDITIONAL REFERENCES ON TIME-SHARING SYSTEMS

BAUER, W. F., AND HILL, R. H. Economics of time-shared computing systems. Datamation. Nov. 1967, p. 48–55.

BETYAR, L. A user-oriented time-shared online system. Communications of the ACM. July 1967, p. 413–419.

BURGESS, E., ed. On-Line Computing Systems; Proceedings of the Symposium Sponsored by the University of California at Los Angeles and Informatics, Inc., Feb. 1965. Detroit, American Data Processing, 1965.

FANO, R. M., AND CORBATO, F. J. Time-sharing on computers. Sci. Amer. Sept. 1966, p. 129–140.

GLASER, E. L., AND CORBATO, F. J. Introduction to time-sharing. Datamation. Nov. 1964.

KLEINROCK, L. Time-shared systems: a theoretical treatment. J. Ass. Computing Machinery. Apr. 1967, p. 242–261.

NIELSEN, N. R. The simulation of time sharing systems. Communications of the ACM. July 1967, p. 397–412.

OPLER, A. Requirements for real-time languages. Communications of the ACM. March 1966.

OVERHAGE, C. F. J., AND HARMAN, R. J., eds. INTREX; Report of a Planning Conference on Information Transfer Experiments. Cambridge, Mass., MIT Press, 1965.

SCHATZOFF, R. T., AND WIIG, R. An experimental comparison of time sharing and batch processing. Communications of the ACM. May 1967, p. 261–265.

SUMMIT, ROGER K. DIALOG: an Operational On-Line Reference Retrieval System. In Proc. ACM 22d National Conf. Washington, D. C., Thompson Book Co., 1967, p. 51–56.