IBM ADVANCED SYSTEMS DEVELOPMENT LIBRARY IN TRANSITION

Marjorie Griffin

The IBM Advanced Systems Development and Research Library in San Jose became a pioneer in mechanization for the sake of expediency. It was organized conventionally in 1952, with its main purpose to give dynamic information service to the personnel in the Research Laboratory. Then, during 1955-56, an expansion program within the company increased employees from 80 to 1,000, and locations from 1 to 16 buildings. The demands of the new users on the small staff created a need for faster processing and greater control of library material. As a result of this demand, the second function of the library was introduced: experimentation in the mechanization of library routines to speed the flow of processing. We decided to mechanize those routines which would immediately minimize our problems of backlogs and of control of dissemination of library material. We used the IBM machines which were then available, including a keypunch, a sorter, a collator, and an accounting machine—the IBM 407. The underlying philosophy throughout this development was to provide greater service with as much economy as possible.

Routing of Periodicals

Accordingly, the procedure of routing periodicals to personnel in different buildings received first emphasis. The objective was to reduce time spent in typing the routing slips, by making machine-preprinted routing lists on a monthly basis, to which names of personnel could be easily added or deleted. To accomplish this objective, we took the following steps:

Marjorie Griffin is Librarian, IBM Advanced Systems Development Division, San Jose, California.

1. The periodicals were arranged in alphabetical order and assigned a four-digit code number from the IBM 10,000 Division Code for Proper Names, with 99 unused numbers left between titles. For example, Bell Lab. Record was number 2,100, and Bell System Technical Journal was 2,200. This identified each periodical numerically, for machine purposes, and permitted the addition of new titles in alphabetical order.

2. Routing slips for each periodical were analyzed to determine what information on them remained the same (and were machine usable) for every issue, and what varied and thus would have to be added by the periodical assistant at every routing. Only the date of the issue and the date the periodical is sent on routing had to be added.

3. To simplify the addition and deletion of users from a routing slip, all bibliographic periodical information was placed on one card (header), and employee information was punched on another (detail). A detail card was made for every person on the routing list. These card sets were linked by code and copy numbers of the periodical. Figures 1 and 2 show a card set (header and detail cards) and the routing slip printed from them.

4. The routing slip was designed in a double form: one half to be used as the circulation card, while the periodical was out on routing, and the other half to be attached to the periodical as a mailing slip. An IBM 407 control panel was wired to print the same information on both halves of the form from a single deck of cards (Fig. 3).

The advantages of this system over the manual one were: (1) It was easily updated; (2) The correct number of copies of the routing slip were made for the number of copies of the periodical received; (3) It saved time in processing--only the date of issue and the date of routing had to be added manually for correctly routing each periodical; (4) The exact reproduction onto the circulation card portion of the slip of the names on the mailing slip portion permitted tracing the issue if needed; and (5) Upon return of the issue, the mailing slip was destroyed and its other half could become the permanent circulation card. After seven years of usage this successful system remains substantially the same.

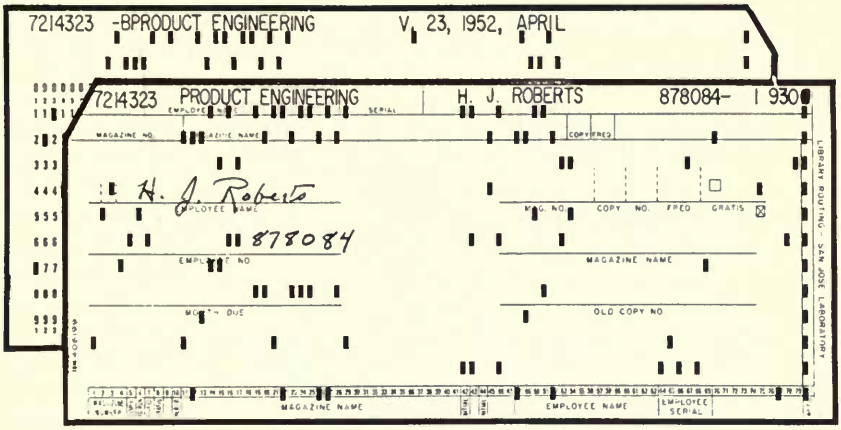

Figure 1
Routing Card

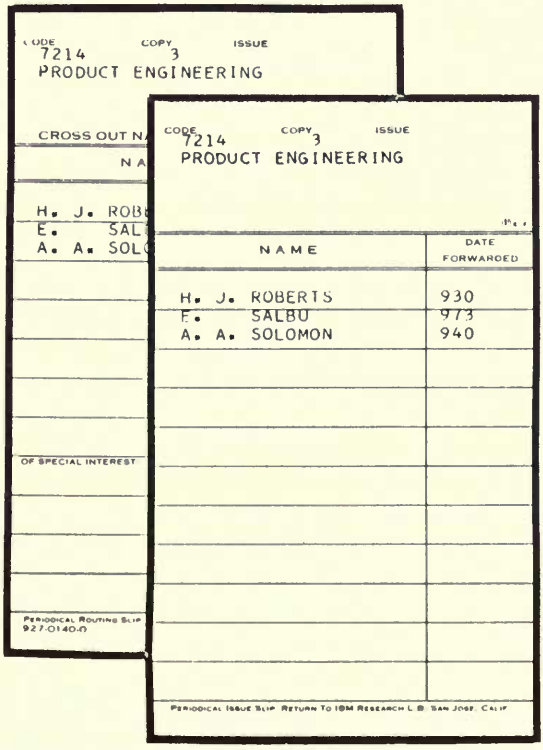

Figure 2
Routing Slip

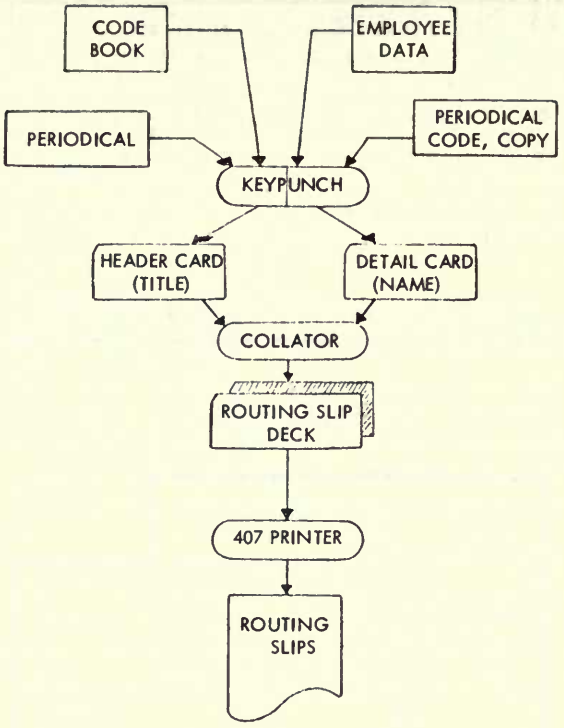

Figure 3
Routing

Circulation

The next routine to receive attention was circulation. In order to keep tighter control of books and pamphlets, we needed to inform borrowers regularly of all material charged to them and to inform a terminating employee on a demand basis of all material charged to him. To achieve these goals we took the following steps:

1. We identified the information we would need to know for machine-producing these notifications. This was:
   a. for the employee—his name, department, serial number, and the date he borrowed the material.

b. for the material—author, title, volume, part, call number, copy, and year of publication.

2. We punched a master deck of requisite bibliographic information for each item from the shelf list onto one IBM punched card. This meant compressing the information from a Library of Congress card into 80 columns or characters. Fixed fields were assigned to each kind of information.

3. We decided upon a type of circulation card that would be used for a single transaction and then placed in an inactive file, to be used later for statistical purposes.

4. We designed a circulation card which would incorporate bibliographic information from the master deck and still leave room for the borrower information to be punched. Compromises had to be made in this compression, and we decided not to punch the title, but to type it on the card for visual reference.

5. The circulation system was set up as follows. When an item was charged out, the borrower wrote his name, serial number, department, and the date on the card. This information was keypunched into the card, and the card was duplicated with the exception of the borrower information. The two cards were placed in the circulation file. When the item was returned, both circulation cards were pulled from the file. The card containing the borrower information was placed in an inactive file, from which circulation statistics could later be obtained. The other card was placed in the book pocket, ready for the next transaction.

6. At the request of some of the library users, overdue notices were designed as single sheets, one item per sheet. This sheet could be trimmed to provide a 3"× 5" record of a book borrowed. We tried to send these notices once a month.

7. To print the overdue notices, the completed circulation cards were merged with the master deck so that the title would be printed. This required a rearrangement of the active circulation file for processing.

This system helped us meet some of our original goals of tighter control. It gave us these advantages over the manual system: (1) We could prepare overdue notices more quickly than by typing them, even though we fell short of our goal of monthly reminders. (2) We could easily gather a variety of statistics for periodic reports to management. (3) We could print lists of loans to terminating employees, on demand.

However, unsuspected weaknesses were discovered while the system was in operation. Weaknesses found were: (1) Typing the title on the circulation card took too much time; (2) To print overdue

notices, the master deck had to be sorted and merged with circulation —a laborious card-handling operation; and (3) A single-sheet overdue notice for every item was costly in printing, and time-consuming in handling. Therefore, we changed the procedure as follows:

1. From the master card, the bibliographic information is now reproduced by an IBM 519 onto a redesigned circulation card (Fig. 4).

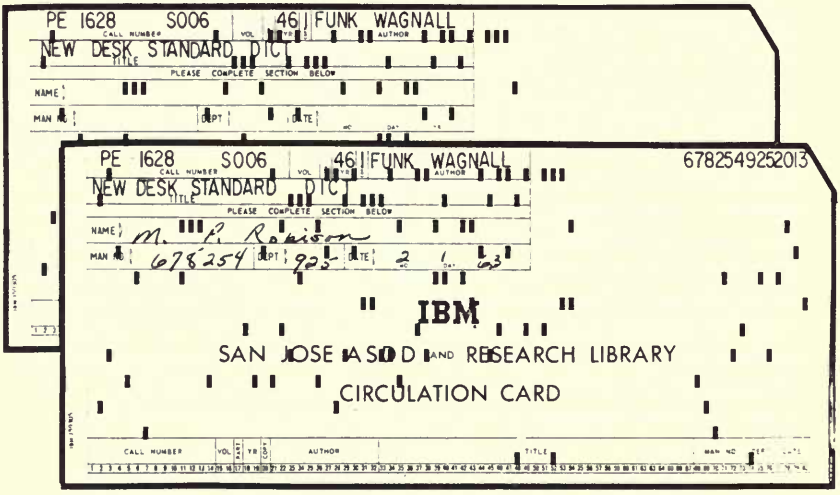

Figure 4
Circulation Card

2. This redesigned circulation card includes a truncated title which allows us to use the circulation file to print overdues and termination lists without merging it with the master deck. The title no longer has to be typed for visual reference.

3. A second duplicate of the completed circulation card is now made at the time the borrower information is punched. This duplicate is filed by the employee's serial number, and termination listings can be made directly from this file.

4. Overdue notices, now called inventory control lists, are also printed from this file, which is kept up-to-date daily. The file is merged with a personnel deck to print borrower's name at the top of the inventory.

5. The individual inventory control lists are printed on single sheets, so that the borrower can see at a glance what is charged to him (Fig. 5).

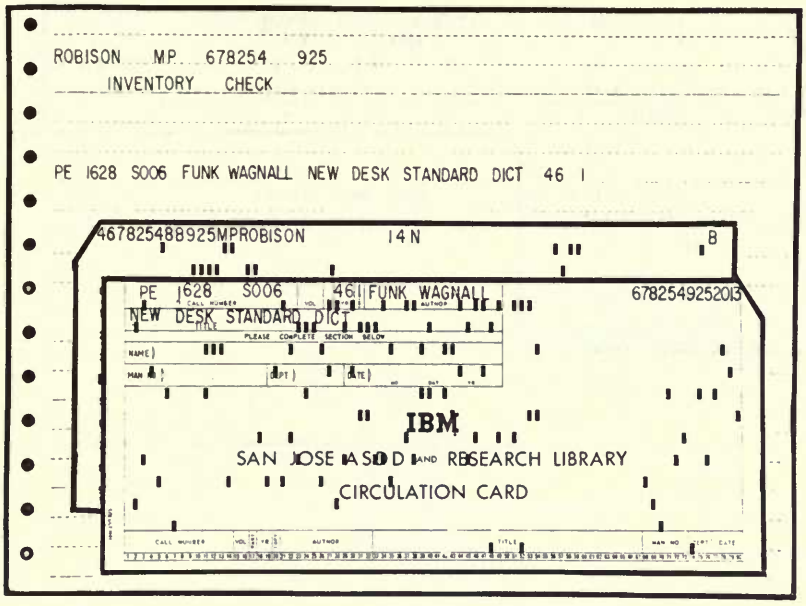

Figure 5
Inventory Control

Author and Title Book Catalogs

By 1957, in this rapidly developing location, several small libraries had been started to accommodate the engineers in scattered buildings. Book catalogs seemed the solution to the problem of keeping up-to-date and comprehensive catalogs in every library in this complex.

We decided to make this catalog a finding list, rather than a complete catalog with full bibliographic entry. The Research Library already had, for the circulation procedure, a master deck of punched cards prepared from shelf list cards. These could be used to prepare a short-entry type of finding list (Fig. 6). In addition to being partially prepared, this was by far the cheapest way to process the information, allowing faster punching, and faster printing, and producing smaller catalogs.

The system was designed as follows. Each individual library maintained its own card catalog and supplemented this by using the book catalogs which were reprinted monthly. The assistant in each

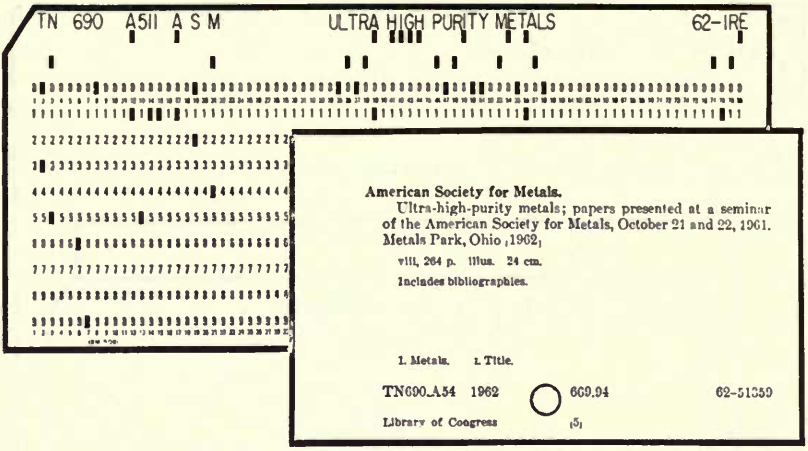

Figure 6
Master Card

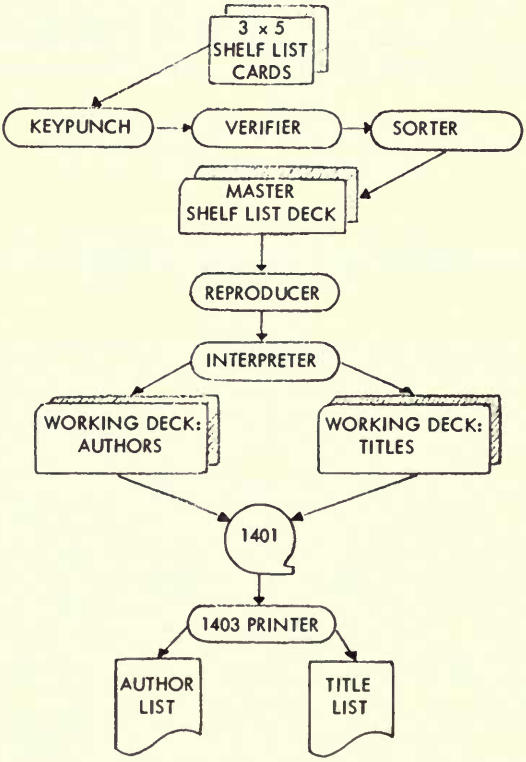

Figure 7
Author-Title Catalog

library forwarded her shelf list cards to the Research Library for processing. Three book catalogs were made—for authors, titles, and subjects.

The production of the author and title catalogs followed these steps (Fig. 7):

1. Two columns were designated on the master card to identify location of book; e.g., RE was Research, PD was Product Development, etc.
2. Each master card was reproduced twice. One set of master cards became the author deck, the second the title deck, and the original remained the master source deck (Fig. 8). The author deck was alphabetically sorted on the author field; the title deck on the title field.

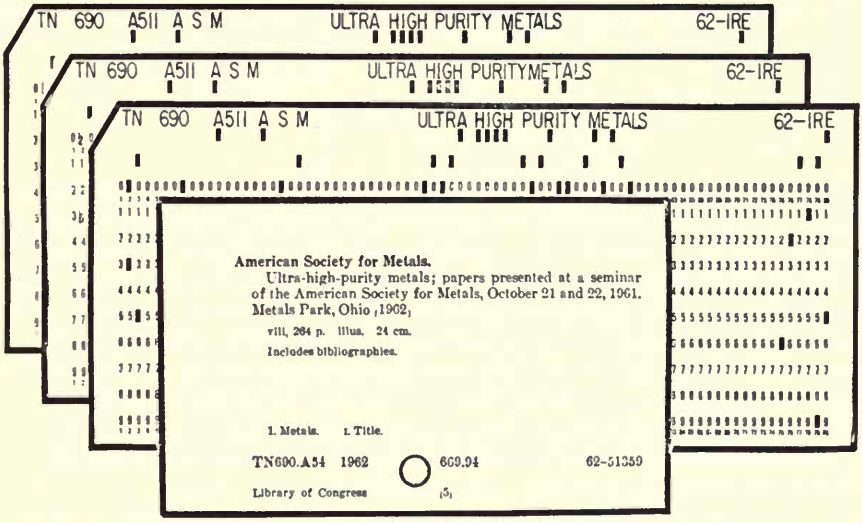

Figure 8
Author-Title Catalog

3. Two panels on the IBM 407 were wired to rearrange the information on the cards when printing the catalogs, with the titles first in one case, and the authors in the other. The complete entry remained the same; only the order of entry differed.
4. For updating these catalogs, the cards made for the additions were sorted into their respective sequences and merged with their main decks on a collator. The new decks were then reprinted.

We learned that it was necessary to establish an authority file for abbreviations in the titles and corporate entries. In manual filing one can poke a card into its correct place, but a machine will only sort the card as it is punched, and a collator will only merge like names into sequence. A major improvement was made in our catalog procedure when we normalized the language used in the titles and corporate entries.

When we improved the processing procedure for the title and author catalogs, we were able to take advantage of the greater capabilities of the IBM 1401, which had become available in our laboratory. We transferred the author and title decks to magnetic tape, and now, instead of using a sorter and collator for updating, we merge new additions directly into the existing files on the tape. The IBM 1401 sorts the decks into their correct sequence, and the IBM 1403 prints the new catalogs. The total process is accomplished in less than 15 minutes—a saving of two days of sorting, merging, and printing.

Subject Book Catalog

The subject catalog was a more complicated procedure to initiate. It required establishing the subject-heading code dictionary, as well as adding to the master card the number assigned to each book's Library of Congress subject heading. To compile the dictionary (1) each Library of Congress subject heading we had used, and each *see* and *see also* reference was punched on an IBM card; (2) the deck was sorted into alphabetical order; (3) to each subject heading and *see* and *see also* reference a five digit code number was assigned, leaving 99 numbers between entries for interfiling future subjects. From this deck, the subject code dictionary was printed by an IBM 407.

The subject headings on Library of Congress shelf list cards were numbered from the new dictionary, and our next problem was to find space for these numbers on the already-full master cards. Rather than compress the titles to accommodate the extra data, we transferred information from the single master card into two cards, one containing the title, and call number, date, location and subject-heading code, and the other the author and the same common information. This two-card-per-entry deck allowed us to print subject catalogs, but these catalogs required two lines of print per entry, and lost in readability and usefulness. This handicap proved severe enough that we revived the single entry card for the subject catalog deck. We reduced the title field on the master card, but only by six columns, so that the title was still identifiable. This left a field into

which we could put the subject-heading code number when we made the entry card from the master card. Fortunately, we were still a small enough library that such experimentation was possible (Fig. 9).

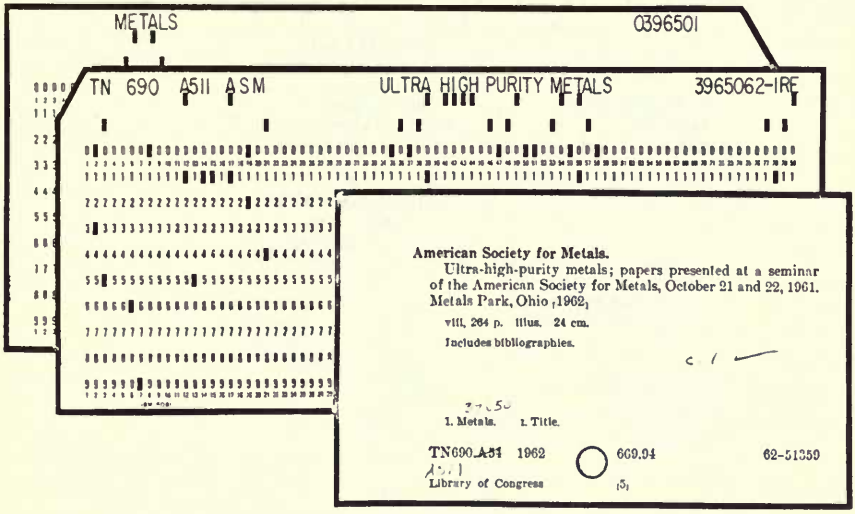

Figure 9
Subject Catalog

The subject card is produced from the master card as follows (Fig. 10). The subject card is made at the same time that the master card is keypunched, unlike the author and title cards which are reproduced on the IBM 519. By using the alternate programming on the keypunch, we automatically duplicate the bibliographic information from the master into the subject card, until we come to the fixed field for the subject code. The code is then punched, and the remainder of the information duplicated. This single subject card format produces a smaller and more readable book catalog.

At present, a program is being written for the IBM 1401 which will arrange the entries first by subject and then by alphabetical author sequence under the subject. This will save days of processing time, for this computer sorts in blocks and will print the catalog at 600 lines per minute, instead of the 150 lines per minute of the IBM 407.

In analyzing the book catalogs system we found that although weaknesses had grown from too many compromises, advantages remained over a manual system. Advantages found were: (1) A complete card catalog in every library was not required. (2) Therefore,

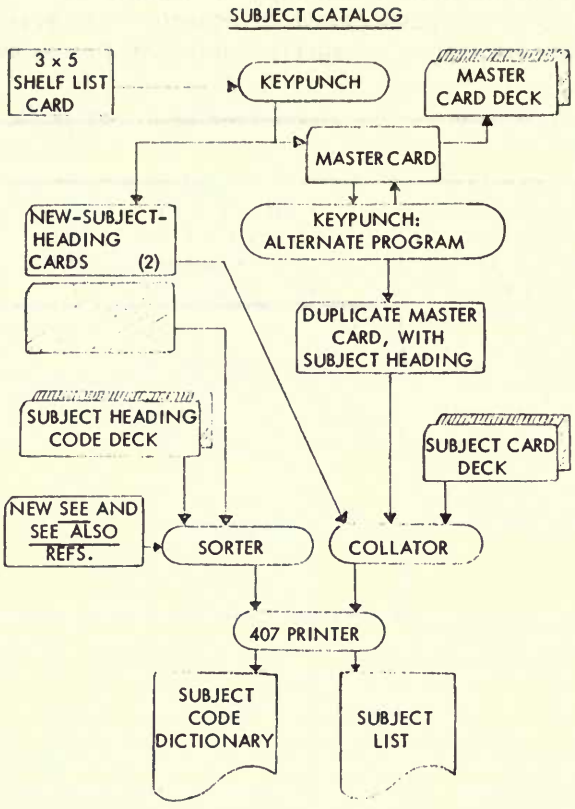

Figure 10
Subject Catalog

many sets of duplicate cards did not have to be typed. (3) The branch assistants were able to locate books in other branches without telephoning. (4) Users preferred to scan a page of text rather than use a card catalog. (5) Once organized, the system was simple to update. (6) Because complete lists were readily available, there was less duplication in ordering material. (7) Users and librarians alike enjoyed having portable catalogs that could be carried to the shelves.

Ordering

The next system to be analyzed for mechanization was ordering. There were two specific purposes in mind: (1) We wanted to be able to trace the progress and standing of an order—at any time, and (2) We were frequently asked by management and accounting for many and varied statistics on orders.

A 3" × 5" order card for library users had not provided sufficient room for the information added by the order librarian, which would trace the progress of the order; and in addition, statistics could only be obtained by hand counting. Accordingly, we listed the information we would need to know for a multiple-purpose order card which would contain all data about an order and which could be quickly processed to give us the statistics we needed.

This was: (1) Book information--author, title, series, publisher, year, edition; (2) Additional report and pamphlet information—issuing agency, code number, verification, and source; (3) Requester information—employee's name, serial number, department, division, and the date of the request; (3) Library information—number of copies received, call number or Library of Congress number; and (5) Purchasing information—whether item was received or not, gift or purchase, price, agent, and dates of order, receipt and payment.

Benefits we aimed to accomplish were: (1) To keep a single-card, complete record of each order; (2) To obtain statistics on total number of orders by year, month, division, department, employee, and subject; (3) To obtain total costs; (4) To be able to follow-up back orders with minimum effort; (5) To assist the Accounts Payable Department with invoice approval; (6) To have our "order-received" cards provide a complete shelf list; and (7) To obtain accession listing from the punched "orders-received" cards.

When these goals had been determined, we designed the order card. To assist the users, we attached two carbon copies to the main card. The requester took the first so that he would have a copy of his order. The other was sent to him by the library when the requested item was ordered. This was his justification for complaint if the item was not received within ten days. Its progress could be checked easily on the order card.

By punching the information from this card onto other cards, we could manipulate the data to print lists in any order we chose. All the ends we planned to accomplish with this system have been achieved with one exception: again we compromised on the title, and have omitted it from listings. With the success of the system, we have established an interlibrary loan procedure using similar techniques and cards.

Information Retrieval

While these procedures were being developed, another area of our library was receiving experimental attention, that of subject searching. To speed our answers to the queries of engineers, we now have an information retrieval system on the IBM 7090, which stores information from 33,000 documents. However, this retrieval system has evolved from its beginning on the IBM 101, through the RAMAC 305, and the IBM 704. Each time we have converted the system to be able to utilize the greater capabilities of the machines, as they have become available in our laboratory.

The documents in the system include reports, pamphlets, bibliographies, and selected periodical articles. Keywords from the title, abstract, summary, first and last paragraphs, and paragraph headings are manually selected from the documents. This technique of indexing was chosen because of the speed of input it allows, and because we felt it would be more efficient to have the intellectual structuring (which is what classifying is) occur at the time of the search. This took advantage of the searcher's knowledge of his subject field and its vocabulary, and his ability to pinpoint the terminology that would produce a tailor-made individual bibliography. Even in 1959 we were thinking in terms of eventual machine indexing and wanted to begin with a system that would be compatible with it.

Material is processed as follows for the IBM 7090 system (Fig. 11):

1. The documents are sent to the library machine room where a document number is assigned.
2. Two card sets are punched directly from each document. One contains the index terms, hereafter called the descriptor cards, and the other the bibliographic information, called the document cards. The machine number links the two sets together.
    a. The document card set contains the machine number, the author, title, date, and library call number.
    b. Each descriptor card contains only the machine number and a single descriptor.
3. From these sets of cards stem by-products:
    a. The document cards are printed on the IBM 407 to make the master reproduction copy for the Daily News—our medium for getting word to the laboratory personnel of material received in the library.
    b. The first document card in a set is reproduced and interpreted to produce the circulation cards.

c. Document cards are reproduced and interpreted for two files, author and report number, used manually for quick reference.

4. The document and descriptor cards are sent to the Computation Laboratory where the information on them is transferred to two tapes for the 7090. The information retrieval program compiles bibliographies by searching these tapes, tying document to descriptor by means of their common machine number. By-products also result from this process:

a. At the time of updating, the descriptor tape is printed to produce the vocabulary listing for visual reference in framing a search.

b. The descriptor and document cards are combined by the IBM 7090, and a listing is printed of all titles under each single descriptor. The chief use of this lengthy listing is for browsing under single terms.

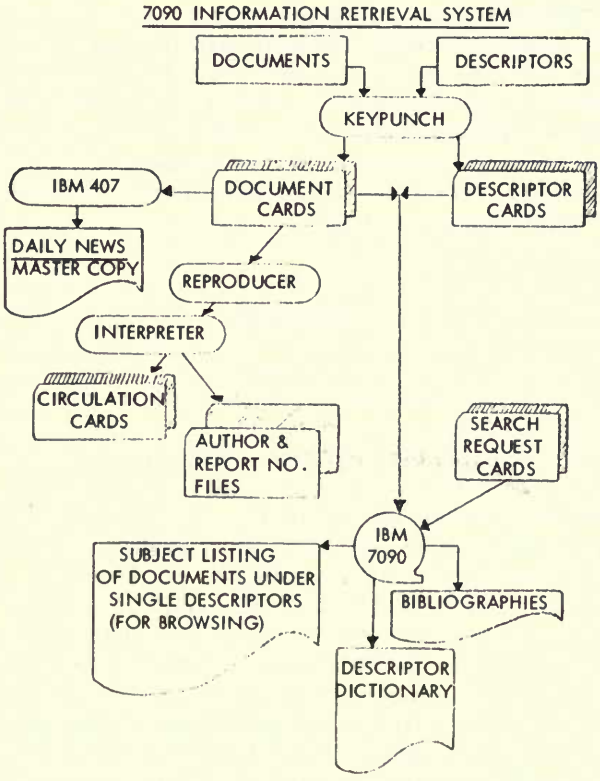

Figure 11
7090 Information Retrieval System

When a requester initiates a search, a librarian fills out a search worksheet giving the requester's name, department, extension, and date; and then the descriptors and the logic to be used in the search. The logical connectives accepted by this program are AND, OR, and NOT. This worksheet is sent to the keypuncher who punches search request cards. The program will permit six requests per pass of the descriptor tapes through the 7090 reading station. The machine numbers located on this pass are looked up on the document tape to obtain the complete reference. The resultant bibliography lists title, author, call number, and date for each item.

The machine searching program allows us great freedom in manipulating and combining the single descriptor terms to construct tailor-made searches. We can specify a precise inquiry by the ability of the program to combine terms with the logical connectives, and we can make a generic search by masking the grammatical endings of the descriptors.

The advantages of this information retrieval system are several: (1) It is a time saver—we can run a search in about two minutes which, by manually searching the published indexes, would take up to several days. (2) We can locate more current information, for published indexes are at least three months late, and we update biweekly. (3) The users find the printed listing easy to use and the reading of them analogous to browsing. (4) We can expand or further delimit a topic by making a second search on descriptors located in a pertinent document from the first search.

New Goals

Each of the previous procedures was mechanized to circumvent an immediate problem, usually to secure a particular benefit. Through our experience in using the machines and through discovering our early systems' limitations, we became aware that the many machine produced records needed should stem from one complete entry and that the comprehensive and interdependent system should be planned in advance before any of it is mechanized. We know, too, that hindsight gives a perspective not discernible while developing a system.

Fortunately, we have been given another challenge—to plan a new library for a major move in January 1964. Benefiting from the experience gained from mechanizing the present library procedures, the new library organization is being based on a total systems concept, with no piecemeal approaches. Each of the present mechanized procedures has been studied in detail to determine data required for efficiently manipulating and retrieving information and each procedure has been analyzed, step by step, for the integrated system.

Whereas in our piecemeal approach, we implemented procedures in a practical but illogical progression with the ordering procedure at the end of the library mechanization, in this rebirth we plan a logical, systematic flow in which all procedures will originate from a basic initial request, whether for an interlibrary loan or to order a book, pamphlet, or periodical. This information will be checked for completeness and for accuracy and then stored in computer memory. Through different programs we will instruct the computer to manipulate these data.

The single record of complete, accurate information on every item will be the basis of all future operations on that item. Around this one source record, and the ability of the computer to withdraw sections of information from it, we are planning services in addition to those now in effect. To assist in accomplishing our objectives we shall use such devices as remote station typewriter input, mechanized format-correction of bibliographic entries, and remote terminal displays to facilitate on-line information availability.

We expect the system to do the following new operations automatically: (1) Print orders to vendors, (2) Print claims for periodicals, (3) Print the Daily News, (4) Print lists of recent accessions to match an employee's personal interest profile, (5) Print bindery information, (6) Produce index terms for documents from title and text input, (7) Print union lists of periodicals, (8) Print routing slips for periodicals, (9) Print a check-in card for each periodical issue, (10) Print call number labels for book spines and card pockets, (11) Print circulation cards, (12) Print book catalogs, (13) Print overdue or inventory notices, and (14) Handle circulation transactions.

Since the library staff will be freed from these clerical tasks, they will be able to render more intellectual service; for no matter how much we accomplish in library automation, there is always need for intrinsic intellectual effort.