A comprehensive automated bibliographic control system has been developed by the New York Public Library. This system is unique in its use of an automated authority system and highly sophisticated machine filing algorithms. The primary aim was the rigorous control of established forms and their cross-reference structure. The original impetus for creation of the system, and its most highly visible product, is a photocomposed book catalog. The book catalog subsystem supplies automatic punctuation of condensed entries and contains the ability to produce cumulation/supplement book catalogs in installments without loss of control of the cross-referencing structure.

BACKGROUND

In 1965 studies confirmed what much of the New York Public Library’s administration had long felt: the public card catalog of the Research Libraries, containing entries dating back to 1857, was rapidly deteriorating. It was estimated that 29 percent of the cards were illegible, damaged, or in some other way unusable. Further, cataloging and card filing arrearages were monotonically increasing at an alarming rate. Increases in labor costs were eroding all efforts to cope with these problems manually. In addition, the deputy director at that time (now director), John M. Cory, realized that a wider base of support was absolutely essential to the survival of the New York Public Library as an institution.

As a result of these disquieting observations, three logical conclusions followed. First, the existing card catalog would have to be closed off, rehabilitated, and photographically preserved. Second, available technology should be explored as a possible solution to some of the spiraling arrearage problems. In particular the applicability of computer technology was to be explored. This exploration appeared to offer some most attractive long-term solutions. The capture of all future cataloging in a machine-readable form would obviate for all time the deterioration problem. This strategy could also provide a basis for a check against spiraling costs, since traditionally unit costs have tended to increase in manual and decrease in
automated systems. Seen within the context of the MARC project at the Library of Congress (LC), the economies were becoming manifestly obvious. The long-term benefits to the entire library community of a national network of shared machine-readable bibliographic data could not be denied. Capture of data in machine-readable form for use by information retrieval systems which might become economically feasible in the near future had to be viewed as a matter of great value. Third, wider access to the resources of the New York Public Library had to be provided if a wider base of support for the library’s operation was to be sought.

The solution decided upon was the development of an automated bibliographic control system capable of producing photocomposed book catalogs. The book catalog would then serve as the prospective catalog and augment the retrospective card catalog, which would also appear in book form following photographic duplication of the cards.

This solution, at one stroke, addressed itself to all three of the major problems, and showed great promise as a future investment. Reproducible book catalogs could be widely distributed. A machine-based system would eliminate manual filing, would take full advantage of cataloging available from MARC, and would begin at the earliest possible time the establishment of an invaluable machine-readable bibliographic data base.

Photographic techniques had already been employed in producing book catalogs, e.g. the National Union Catalog, the Book Catalog of the Free Library of Philadelphia, and the Enoch Pratt Free Library catalogs, among others. Computer-produced book catalogs embodying various techniques (computer line printing, photo-typesetting, etc.) and levels of sophistication were being produced by many institutions, e.g. Harvard University’s Widener Library Shelflist, Stanford University’s Undergraduate Library Catalog, Baltimore County Public Library’s catalog, among others. An extensive review of various types of book catalogs including typical pages of each is given by Hilda Feinberg.

Following extensive studies conducted by Messrs. Henderson, Rosenthal, and Nantier of the NYPL Research Libraries, the Systems Analysis and Data Processing Office (SADPO) was formed, staffed by EDP and library specialists, to be completely dedicated to the solution of problems of automated bibliographic control and library automation. From the beginning it was decided that if EDP technology were to be utilized, it should be utilized in a manner which took full advantage of the properties of the medium. The computer was not to be used as an ultrasophisticated and costly printing press. The application of new technology to a field will invariably lead to waste and awkward results if the intrinsic properties of the technology are not fully utilized.

The fundamental properties of EDP technology lie in its abilities to:

1. Reorganize and combine data;
2. Select items meeting a set of predefined conditions;
3. Maintain a permanent but flexible correlation between items;
4. Transform a set of conditions into data;
5. Perform all of the above with remarkable speed and accuracy;
6. Perform all operations with a merciless consistency.

Thus, it was realized, at the outset of the project at NYPL, that technology could provide a great deal more than the maintenance of a machine-readable record and its reorganization for display. A rigorous control of bibliographic data was possible, and would extract maximum utility from any investment in EDP technology. It was with these ideas in mind that machine-based authority control and filing systems were developed. The authority control file provides the fundamental utility of the system. Control of data usage has always been of paramount concern to the professional bibliographer. It becomes even more important in a machine-based system in which the data lie in an essentially invisible form until a fairly complex display operation is performed.

**Advantages of an Authority File**

Another bibliographic aid which the computer could provide through an authority control system was the maintenance and integrity of a cross-reference structure. In addition, one of the classical functions of cross-referencing could be eliminated: it would no longer be necessary to direct a user from one classification which has been used extensively to a newer one when terminology changes. Consider the problems which might arise if the Library of Congress were to change its current usage of the heading *Aeroplane* to *Airplane*. It would be virtually impossible, under a manual system, for a library to attempt to locate, alter, and refile all cards bearing the tracing *Aeroplane*. With a central authority file the problem is reduced to a single transaction and a fraction of a second of effort by the computer. The change is effected with an accuracy unattainable in a manual system. Finally, the common nuisance of a cross-reference leading to yet another cross-reference is automatically obviated.

The presence of a machine-readable authority file and the ability to verify use of all forms against this central authority, with machine accuracy, eliminates all clerical errors in the usage of names and headings to which a manual system is susceptible. The problem of consistent usage is greatly compounded in a machine-based system which does not provide mechanical verification. Inconsistencies in any automated system generally tend to diminish its utility, and invariably lead to ludicrous results. Nonetheless, inconsistencies of usage in an automated system are more readily corrected than those in a manual system. The existence of a central authority file, however, reduces the operation to maximum simplicity and allows no deviation from established standards.

While maximum rigor in machine control was attempted, an attempt was also made to shield the professional librarian, who would be using the system, from as much of the tyranny imposed by the machine as possible. In the system finally adopted, the librarian need only exercise care when
establishing a form. Following establishment of the form, the cataloger need not be concerned with any of the details of the entry, such as punctuation, accent marks, MARC delimiting or categorization. The authority subsystem supplies all such details. In short, the cataloger is only required to spell the form correctly. The machine will identify any incorrect usage; thus a great deal of tedious and time-consuming (and thereby costly) manual searching is eliminated.

At the same time that work began on the automated system at NYPL extensive activity in library automation was also in progress in many other parts of the country, involving virtually all areas of library operation: cataloging, acquisitions, serials control, circulation, and reference services (information retrieval). Since, at NYPL, it was assumed that the bibliographic data base and its control would form the cornerstone of each of these systems, cataloging was given first priority. This approach differed from that taken at other institutions; others, Columbia University for example, chose to develop an acquisitions system first. Still others developed highly sophisticated circulation systems, Ohio State University being notable among these.

Even among those institutions which chose to address themselves to the problems of automated cataloging, important differences in approach were evident. These differences were largely a result of attempts to solve different types of problems related to cataloging. Among the many projects initiated at that time two will be mentioned, as they are representative of the differences in approach to automated cataloging.

The first is represented by the University of California Union Book Catalog project, undertaken by the Institute of Library Research (ILR). This system is characterized by an attempt to minimize, via computer programming, manual intervention in data preparation. Employing the technique of Automatic Format Recognition, the ILR staff attempted to find the most economical means of rendering a vast amount of retrospective data into machine-readable form. In converting such a large amount of data they had to also concern themselves with the statistical error levels to be expected from keying. Having decided that extensive manual edit was too time-consuming and costly, and itself prone to statistical error, they attempted to create computer programs which would use the massive amounts of data as a self-editing device. In a sense, ILR used the nature of the problem as its own solution. The goal of the project was the production of a book catalog representing a five-year cumulation (1963–1967) of materials on the nine University of California campuses, and a MARC-like data tape clean enough for print purposes. NYPL, on the other hand, decided to consider only prospective materials in a continuously published catalog, and the creation of a MARC-like record which would approach in completeness, as closely as was economically feasible, that created by the Library of Congress. To this end manual tagging and editing were absolutely essential.
The second system to be considered is the shared cataloging system developed by the Ohio College Library Center.\textsuperscript{12} The primary emphasis here is on the economy to be derived by instantaneous access to the combined cataloging efforts of a cooperating group of libraries. At OCLC the primary emphasis was placed on on-line bibliographic data input and access. The major bibliographic product to be produced was a computer printed card set. The overriding consideration of OCLC was the sharing of resources among many users, while at NYPL the major concern was the content integrity of a single user’s file.

Advantages of a Book Form Catalog

A book form catalog has several advantages over a card form catalog: it is portable, compact, more readily scanned and extremely simple to reproduce. When coupled with an automated system for maintenance and production the advantages are greatly magnified, as manual filing is virtually eliminated. The format, sequencing, and usage of terms in a book catalog may be varied at will to accommodate users’ needs and library service policies. Advantages and disadvantages to book catalogs are summarized in the introduction to Tauber and Feinberg’s collection of articles on book catalogs.\textsuperscript{13} Comparisons of book versus card catalogs are presented by Catherine MacQuarrie and Irwin Pizer in articles reprinted in the work cited above.\textsuperscript{14,16}

The most obvious advantage of the book catalog is its portability. Wide availability of the catalog of a library’s collection makes possible a level of service not economically feasible under any other system. Access to the complete collection of a library system can be made available economically to every educational institution in the region served by the system. Access to a highly valuable research collection can be made available to a much wider geographic region than was hitherto possible. The concept of a union catalog for a region becomes much more viable, making possible regional cooperation in acquisitions policies and relieving the burden of heavily duplicated collections currently borne by library systems within manageable geographic regions. Such cooperative ventures allow the cost of maintaining the catalog to be defrayed among the various members of the consortium. Thus, a book form catalog would appear to provide groups of libraries with the possibility of operating economies, while increasing the overall level of service to the public they serve.

The utility of a book form union catalog has already been demonstrated by the experience of the Mid-Manhattan Libraries in New York. Mid-Manhattan, a central circulating library, consists essentially of five libraries in two locations. Provision of complete bibliographic access with a traditional card catalog would require the manual maintenance of five individual and two union catalogs. The utility of the Mid-Manhattan catalog has been further increased with the inclusion of the entire NYPL Branch Library system in January 1973.
A library's internal operation benefits by wide availability of the catalog, as individual copies of the catalog can be made available to the acquisition division, the cataloguing division, and each of a library's special collection administrators, making references to the traditional Official Catalog more efficient; such has been the experience of NYPL. Baltimore County Public Library reports a similar finding.15

A perhaps hidden advantage of a book catalog lies in its compactness. A book catalog requires neither the space nor the expensive furniture required by a card catalog. The problem of space becomes more and more acute as the "information explosion" continues to mushroom. An ironic squeeze is encountered in that the collection yearns more and more for the space occupied by the catalog, while the catalog, in growing, continues to make its own demands on available space.

DESCRIPTION OF THE NYPL BIBLIOGRAPHIC SYSTEM FILES

Before attempting to describe the book catalog subsystem, we shall briefly describe the nature of the files from which the bibliographic data are drawn.

The complete bibliographic system consists of four major files and computer programs for their control and maintenance. The files are:

1. Complete MARC data base (updated weekly with all changes and additions) from which cataloging may be drawn;
2. Bibliographic master file;
3. Authority master file;
4. Bibliographic/Authority linkage file.

For the purpose of this discussion we shall take the existence and maintenance of these files for granted, and concern ourselves solely with their use in the production of photocomposed book catalogs.17

Bibliographic Master File

This file contains unit records for each bibliographic item in the collection;† books and book-like materials, monographs, serials, analytics and in-

---

* The system actually consists of three independent sets of such files (MARC is common to all)—one each for the Research Libraries, the Branch Libraries, and the Dance Collection.

† Separate data bases are maintained for the Research and Branch Library systems. The Research Libraries file contains all book and certain book-like material added to its collections since January 1971. The Branch Libraries' file contains all holdings of books and book-like materials of the Mid-Manhattan Library collections. This file currently duplicates to a large extent the holdings of the rest of the Branch System, and will eventually encompass the entire system.
The information content is identical to that of MARC records. Tagging and delimiting adhere to the MARC conventions except in those cases in which it was necessary to expand delimiting in order to enhance the functional utility of the MARC coding structure. Some data distinctions which MARC has since dropped, but which are nonetheless useful, have been retained. The expansions consist of the addition of several delimiters not used by MARC in order to provide filing forms (which are automatically generated, but which may be manually overridden) for titles, and sequencing information for volume numbers of series and serials. Transformations from a MARC II communications format to the NYPL format and vice versa are possible due to the isomorphism of the two records. The transformation of MARC II format records into NYPL processing format is carried out in the normal course of processing, in which MARC records are selected for addition to the NYPL files.

**Authority Master File**

This file is the central repository of all established forms. Names (personal, corporate, and place), series titles, uniform titles, conventional titles, and topical subject headings are all established on this file. Categorization of each form is controlled by this file. No form is accepted for use in a bibliographic record unless it matches a form already established on the authority file, and is used consistently with the categorization assigned to it, e.g. a form categorized as a topical subject is never permitted as an author, a series title may only match a form categorized as a title, etc.

The cross-reference and note structures are maintained on this file. An additional heading employed by NYPL, which falls conceptually half way between a cross-reference and a subject heading, the dual entry, is also controlled here. The dual entry heading serves to bring together, under a non-LC heading, bibliographic items which NYPL considers unique, by virtue of the nature of its collection. An example might be found in the genealogy division which contains a very extensive collection dealing with New York City. Use of the dual entry allows a sequencing under both a subject heading indirectly regionalized to New York City (LC heading) and at the same time a drawing together of all items about New York City into a single sequence headed by New York City. Take, for example, the LC established heading *Elections-New York (City)*; NYPL automatically causes all

---

4 At NYPL a distinction is made between analysis and indexing of a work. The latter refers to selective analysis, used when it is desired to provide, for example, subject access to a significant article in a periodical without creation of the series added entry. There are two types of indexing provided by the NYPL system. The first creates only a subject tracing; such treatment might be accorded an article of topical significance by a staff writer of a popular periodical. The second would create both an author and subject entry; this might be used in the case of an author of note writing on a significant subject in a popular periodical, e.g. Norman Mailer writing on political conventions for *Esquire Magazine*. 
items traced to the above heading to appear under both the LC heading, and the dual entry New York (City)-Elections (Figure 1). The dual entry merely provides an alternate form of organization for display. No bibliographic tracing is permitted directly to a dual entry. The additional entry point is automatically created when a catalog is printed. Manual effort by the cataloger in order to provide the additional entry point is prevented; in addition, the bibliographic record remains rigorously MARC-compatible.

Automatic control of cross-references, dual entries, and the en masse alteration of classification are facilitated by the authority subsystem together with the correlative and reorganizational capabilities of the computer. There is some irony in the relative ease with which the computer allows such individualized organization of data to be effected and the computer's reputation—richly deserved—for imposing a bland uniformity on its victims.

The authority file provides one other invaluable service: it controls, in a single location, filing forms to be associated with a heading. Consistency of filing is assured and, again, extreme simplicity of alteration is possible. Only one record need be changed in order to alter the filing of the entire
body of material associated with a heading. Filing forms are automatically generated, with provision made for a manual override. Automatic filing has been found to be correct in better than 95 percent of the cases currently in use. The remaining 5 percent required manual intervention.

The machine filing algorithms are based on language and on MARC categorization and delimiting. Initial articles are dropped in each of thirty-eight languages, including the major languages transliterated into a romanized alphabet (those employing Cyrillic alphabets, oriental languages, Hebrew, and Yiddish). Chronological subdivisions are filed automatically observing rules regarding inclusiveness of dates, etc. Important chronological periods (currently fifty-four such periods) are recognized and filed automatically, e.g. American Revolutionary and Civil Wars, French Revolutions, Chinese dynasties, Middle Ages, etc. Roman enumeration is automatically filed in correct decimal sequence.

**Bibliographic/Authority Linkage File**

The basic function of the bibliographic/authority linkage file is to provide a communications channel between the two major files by assigning to each authority form a neutral unique number. The linkage file then provides access to the established form regardless of the metamorphoses which it may have undergone since its original use (the number remains inviolate).

Each authority upon addition to the file is assigned a unique number; however, the authority file is sequenced by an alphabetic sort key. This sort key bears no logical relationship to the filing form of the heading; it is constructed by dropping punctuation and accent marks, converting to upper case, dropping multiple blanks, and appending a hash total. The linkage file maintains the correspondence between authority control number and alphabetic sort key. Only the authority control numbers, determined by the first bibliographic/authority file match for each field, are carried in the bibliographic records.

In addition, information is provided to the book catalog subsystem regarding changes in the authority file (alteration of established forms, etc.) which would cause an entry exhibiting such alterations to be immediately regenerated for inclusion in a book catalog supplement. Appropriate action is taken against the bibliographic file when activity to an authority heading is sensed by the book catalog subsystem. The presence of a dual entry form, which will require the creation of an additional entry under the associated variant form, is also indicated here.

**Alternative Input Files**

It should be mentioned that the full set of files described above is not a mandatory requirement for creation of a book catalog. A bibliographic file in a MARC II communications format alone will suffice. We have performed tests using both another library's data file and the MARC file as
sole input to the system. Using unmodified file update software we have generated from these MARC II format data bases complete authority files, and thence book catalogs. No cross-references or scope notes are possible in this mode of operation, since MARC makes no provision for them. A further experiment was performed using another library's data base (in MARC II format) in combination with the cross-reference structure of the NYPL authority file. This led to highly satisfactory results, demonstrating that a photocomposed catalog could be created, and exhibiting the utility of the input file enhanced by cross-references. §

THE PHOTOCOMPOSED BOOK CATALOG SUBSYSTEM

The system for production of book catalogs represents only the visible tip, albeit a large and complex tip, of the entire bibliographic system. It consists, in all, of ten computer programs and several score modules. The system was designed with thought toward production of catalogs with a variety of output options. In most cases, these options can be attained by the elimination of entire programs or modules. Space does not permit a consideration of all possible variations; the most important will be mentioned in the course of the discussion.

One consideration which was deemed of paramount importance was to remain as independent of photocomposition hardware as possible. Photocomposition is yet in its infancy; hence, an inextricable commitment to a particular device, it was decided, was to be avoided. The final approach taken was the design, by SADPO, of generalized photocomposition software which is responsive to device-independent typographic commands. The only function of this software is to accept, as input, completely defined text data and typographic instructions from which it generates formatted pages. This task is accomplished via a translation of device-independent into device-particular commands in the form of a photocomposition device driver tape. Should a new or more desirable photocomposition device become available, or significant advantage be found in employing a different photocomposition vendor, only one program need be altered. The photocomposition software is completely generalized and can be used to generate anything from book catalogs to typeset prose, in virtually any format (see the section on the pagination program for a discussion of the formatting options provided). Figures 2, 3, and 4 demonstrate some of the possibilities. The creation, organization, and control of data to appear in the catalog was undertaken as a completely distinct set of programming tasks.

Design Objectives of the Book Catalog System

Before embarking upon a discussion of the technical aspects of each

§ The September 1972 Hennepin County Public Library book catalog was published with a bibliographic data base produced by the Hennepin County Library combined with the NYPL Research Libraries' authority file.
The page contains a catalog for the NYPL Research Libraries Dictionary Catalog Supplement, November 1972. The catalog is arranged alphabetically by author and title. The page includes sections for different areas of study such as mathematics, physics, and philosophy. The catalog is in English and uses standard bibliographic notation. The page is from the Mid-Manhattan branch of the New York Public Library.
Fig. 4. The Mid-Manhattan Library Catalog Supplement, July 1972. This page was created as a test utilizing 4 pt. type on 4 pt. body. The actual supplement created for use by the public was created in 6 and 8 pt. type.
processing step, we shall state the objectives which we set out to meet, and the constraints—generally economic—under which they were met.

**Method of Publication**

As it is economically impractical to publish the entire catalog on a very frequent basis, a cumulation/supplement scheme was adopted. Two basic types of supplements are possible: (1) a supplement containing only new items for the period represented; or (2) a cumulative supplement containing all items new to the system since the last appearance of a cumulation of the entire collection, automatically replacing all previous supplements. The latter is more costly than the former. The economic desirability of the former was eschewed in favor of convenience to the user. Under the scheme adopted, a user has, at any time, only three sources to consider—the retrospective catalog, the prospective cumulation, and the cumulative supplement.\(^1\)

We have derived several optimization formulae for reaccumulation schedules.\(^1^9\) Application of these formulae indicated a reaccumulation cycle of approximately one year, assuming that supplements would appear monthly. The formulae also indicated that a small premium would have to be paid for the administrative convenience of spreading the printing and processing load of the cumulation over the span of the entire reaccumulation period, compared to the cost of a complete printing at the beginning of each period. The adopted publication scheme calls for the publication each month of \(\frac{3}{4}\) of the cumulation, together with a supplement containing all items which have not yet appeared in the cumulation and those which have been altered since their appearance in a cumulation. The division into twelve segments is table-controlled; the number of segments may be varied from one to sixteen. For example, in January a cumulation is published for the alphabetic span A–B; a supplement is published for the remaining letters of the alphabet. A similar situation would occur the following month, etc. Thus, at any given time the public is presented with a set of volumes representing the cumulated catalog and a supplement which contains all material not found in the former. The public is unaware of the fact that the cumulation is being cyclically updated. They are only aware of the fact that they have no more than three sources to consult: (1) the old card catalog, (2) the basic cumulative book catalog, and (3) the cumulative supplement. The fact that entries are migrating from the supplement to the basic cumulation each month is of no consequence from the standpoint of catalog usage.

The decision governing representation of an item in a cumulation or supplement is made on an entry by entry basis. For example, one of the

\(^1\) All material in the card catalog has become known as the retrospective collection, and all material entered into the automated system after January 1972 has become known as the prospective collection.
subject added entries may have migrated into the cumulation; hence, it will no longer appear in a supplement. However, the main, and all other added entries, falling into different filing ranges, will continue to appear in a supplement until they too can be absorbed into the cumulation. Similarly, alterations to a bibliographic record will cause only those entries whose text or sequencing is affected to reappear in a supplement. A change to or an addition of a subject tracing will cause only that subject added entry to be regenerated for inclusion in a supplement. The main, and all other added entry citations, which remain unaltered, need not reappear in a supplement (assuming they have previously migrated into the cumulation).

Condensed Added Entries

In order to keep printing costs to a minimum, all added entries are condensed; title page extension, publisher, and bibliographic notes do not appear under any of the added entries, the assumption being that the user who is interested in such data will take the trouble to refer to the main entry, which contains the complete bibliographic citation. This type of back-and-forth reference, while quite awkward in a card environment, is extremely simple in a book catalog. Economic considerations also led to the decision to suppress tracings from the main entry. The system was designed to allow these decisions not to be irreversible. The choice of data which are to appear with an entry is governed by a set of tables which may be readily altered should it be desired to change the format or context of an entry.

Punctuation of condensed entries is accomplished automatically. This is not a trivial problem, and one that only a cataloger can truly appreciate. Consider, for a moment, the myriad ways in which bracketing may occur within the title or imprint statement, and the ways in which these may span the two fields. Add to these factors the rules which do not permit the appearance of double punctuation.

We have found that punctuation of added entries is effected correctly in 98 percent of catalog entries. In those instances in which ALA punctuation rules are observed in the complete record, correct punctuation is assured (this is not true of cataloging obtained from European sources).

Control of Cross-references

It is in the realm of cross-references that the mindless consistency of the computer is most effectively employed. The goal to which we addressed ourselves was the absolute integrity of cross-referencing. Under no circumstances—short of erasing a cross-reference from a previously published catalog—were cross-references to refer the user to a heading which did not have an associated bibliographic citation. All meaningful cross-references providing alternate access points to a citation must appear. By the same token, in order to minimize costs, cross-references which appear in a
cumulation available to the public are not to be repeated in a supplement. Cross-references to a heading would be considered valid entry points to the catalog when bibliographic citations appear under a subdivision of that heading. For example, the appearance of bibliographic citations under Negro Art—Exhibitions would cause all cross-references to Negro Art to be generated (Figure 5). The same rules concerning appearance in supplements and cumulations are observed for these secondary cross-references.

Alterations to cross-references which have appeared in a cumulation will cause the altered forms to reappear immediately in a supplement, provided the referenced heading is still in use in the catalog. Similarly, alteration of the referenced heading would cause the reference to the new form to be automatically generated.

Fig. 5. The NYPL Research Libraries Dictionary Catalog Supplement, October 1972: L–Z page 120 and 121. These pages demonstrate the generation of the cross reference Negroes–Art see Negro–Art even though only subdivisions of Negro–Art appear in the catalog.

A further consideration extends to cross-references which have migrated into a cumulation. When a cumulation segment is updated, all cross-references which previously appeared in it should continue to appear if, and only if, the referenced heading is still in use in either the same segment of the cumulation, another segment of the cumulation, or a supplement; if not, its use is discontinued. Subsequent use of the referenced heading would then call up the cross-reference for reuse. Each of the above desid-
erata requires rather intricate logic when the cumulation is being produced in monthly installments, as any of the following is possible:

1. Cross-reference in a supplement, referenced heading in a supplement;
2. Cross-reference in a supplement, referenced heading in a cumulation;
3. Cross-reference in a cumulation, referenced heading in a supplement;

In each case, the cross-reference must be suppressed whenever the referenced heading disappears from the catalog available to the public, but must be retained when it refers to a heading existing in any part of the catalog. The cross-reference and referenced heading may easily appear in catalog segments published as much as eleven months apart, making it absolutely essential that both the authority and book catalog subsystems maintain strict control of the cross-reference structure.

Control of Hierarchies

It was decided that the appearance of cataloging under a subdivision of a heading which contains associated notes should cause the higher level heading with its attendant notes to appear. Such a heading would be forced to appear regardless of whether or not it itself headed a bibliographic citation, under the assumption that notes concerning a heading might be valuable to a user interested in a subdivision of that heading (See Figure 6 for an example).

Fig. 6. The NYPL Research Libraries Dictionary Catalog Supplement, August 1972: page 2. The heading Actors is caused to appear due to the presence of a scope note and the use of a subdivision of the heading.
Dictionary and Divided Catalogs

The same system was required to serve two divisions of the New York Public Library, each of which has different traditions and philosophies of service to identifiably different users. Therefore, an additional flexibility was required of the system: the ability to produce both dictionary form and divided catalogs.

The Research Libraries, which have traditionally used a dictionary form of catalog, wished to continue that practice. The Branch Libraries, on the other hand, felt that their public could be better served by a divided catalog, separated into Titles, Subjects, and Names. The system was designed in such a manner that the modification of a single parameter in the final sort would produce either form of catalog.

BOOK CATALOG SUBSYSTEM—TECHNICAL DESCRIPTION

The entire subsystem consists of ten separate programs, each of which will be described below. The flow charts in figures 7, 8, and 9 depict the processing flow of the subsystem. The system was designed to operate on an IBM 360 model 40 (which has since been replaced with a 370 model 145) with 256K bytes of core storage. The programs were written exclusively in BAL for a DOS configuration. A conversion to full OS has recently been completed. Each processing step described below is executed sequentially.

Significant peripheral devices required are: five tape drives, one disk drive in addition to those required by the operating system, and a line printer.

Please refer to figures 7, 8, 9, and 10 for the programs and files referenced by symbols P1, T1, D1, etc.

Entry Explosion and Construction—Program P1

This program serves as the driver for the entire subsystem. In this step entries are selected for inclusion in a supplement or cumulation segment. Requests for data required from the authority file are initiated. The format and data content of each entry are defined by this program via a set of tables. These tables may be altered at will, allowing redefinition of the format and content of any entry. The bibliographic master file is updated to indicate the appearance of an entry in a cumulation, preventing its subsequent appearance in a supplement. In addition, this program is charged with accepting communication of activity to the authority file and taking the appropriate action with respect to the bibliographic file. This activity may take several forms: alteration of a heading, change of delimiting, change to a filing form, posting or removal of a cross-reference or dual entry, change of categorization, or the complete transfer of all cataloging from one valid heading to another.

Evidence of activity to an authority heading is carried on the authority/bibliographic linkage file (DØ). When such activity has affected a head-
Fig. 7. Subsystem flow chart.
Fig. 8. Subsystem flow chart.
Fig. 9. Subsystem flow chart.
Fig. 10. Subsystem flow chart.
ing used by a bibliographic record as an authority field, the field is tagged for verification by the authority file in the next file update/authority-interface run. The indicator for the field in question, denoting previous appearance in a cumulation, is turned off. At the same time, the indicators for all other catalog tracings which require that authority field as data are turned off. When a transfer from one heading to another has occurred, the new linkage number is inserted into the authority directory of the bibliographic record. This is not absolutely necessary, as the authority/bibliographic linkage file provides the link via a chain when a transfer has occurred. Nonetheless, the insertion of the true authority control number into the bibliographic file eliminates the necessity of a chained search in all future accesses of the tracing, space on the linkage file is conserved, and no additional indicators are required to make note of the fact that the entry has been caused to reappear in a supplement as a result of the transfer. In all cases of activity to an authority record, reverification is forced for the associated tracing field in order to guarantee correct usage of the altered authority.

Each bibliographic record is examined to determine whether it will contribute to the catalog. This is done on an entry by entry basis. Each field of the bibliographic record capable of defining a catalog entry is examined. All fields which define a catalog entry (tags 1--, 245, 4--, 6--, 7--) carry a set of indicators denoting appearance in the cumulation, and a number defining the cumulation segment into which the entry should file. An additional indicator for authority fields denotes the presence (or absence) of an associated dual entry on the authority file. Appearance in the cumulation and filing segment number of the dual entry are also carried in the bibliographic record, allowing independent control of the dual entry citation. As may be readily seen, the dual entry acts as a phantom tracing in the bibliographic record and will thus not be specifically mentioned in the discussion of selection criteria below.

An entry is selected for construction on the basis of the following criteria:

1. The bibliographic record is in a valid status, i.e. has passed all editing tests, and sufficient time for proofreading has elapsed.
2. All authority fields required for construction of the entry have been verified against the authority file in the weekly bibliographic file update/interface production runs.
3. It files in the segment being produced that month.
4. The indicator denoting appearance in the cumulation is NOT set. Thus, any alteration to the content of a bibliographic record, warranting immediate reappearance of an entry, may be communicated to the book catalog subsystem by the extinction of the cumulation indicator.

Both cumulation and supplement entries are created in the same run. The entries are separately collated by causing the highest level of the final
sort to be a code denoting supplement or cumulation.

It will prove fruitful at this point to draw a distinction between a cata-
log entry—the printed bibliographic citation—and the machine record
which is created by the system prior to phototypesetting. The machine rec-
ord is nothing more than a highly organized print record. The final merg-
ing of such print records from various processing steps completely define
the text, typography, and sequencing of the final printed catalog. The ma-
chine print records created by the system up to step P8 will be referred to
as Text Entry (TE) records.

When an entry is to be included in a particular month's catalog segment
or supplement, a table for the particular type of entry is consulted in or-
der to determine the data and the typographic commands which will gov-
ern the entry's format. At this point only a skeleton Text Entry record is
constructed, as all authority data will be obtained from the authority file.

The sequencing information is contained in the sort key of each TE
record, which defines six levels of sorting:

1. Collation—catalog or supplement. This is further refined when a di-
   vided catalog is being produced.
2. Level I sort, and sort code.
3. Level II sort, and sort code.
4. Level III sort, and sort code.
5. Publication date.
6. Publisher.

In the case of certain series entries, level II and III may be split into two
half-size levels by the program in order to further refine the sort sequence.

As an example of the use of sort levels I, II, and III, we might consider
a subject added entry. In that case, the level I sort is defined by the filing
form of the subject tracing, level II by the filing form of the author's
name and level III by the filing form of the title of the work. The sort
codes are used to separate entries which would result in the same sort keys
but are conceptually different, e.g. a name which might simultaneously de-
fine a title added entry, a main entry, and a subject added entry. A similar
situation exists at the second sort level where conventional titles are to be
separated from titles or subject title entries.

Sort key levels, as all other data elements required in a TE record, will
be directly inserted into the record under construction if they consist of
nonauthority data, and will be identified by linkage codes for later inser-
tion when the filing form data is returned from the authority file. The
final TE record will not be completed until step P7, to be described below.

Following construction of the sort key (or indications to complete a sort
key) typographic commands and text data are inserted into the TE record.
The typographic commands are contained as binary bit settings in a record
directory. The directory also defines the location and length of each data
element, or gives a linkage code when the data are to be obtained from the
authority file, and hence cannot be inserted until program P7. The order
of entries in the directory defines the printing sequence of text data. Thus, when text data are available, true locations and lengths are provided in the record. When they are not, linkage codes replace them in the directory. These linkage codes are simply replaced by true locations and lengths when the authority text is added to the end of the record by another program (P7).

It will suffice at this point to mention that all typographic commands are present in the record. The function of the commands will be discussed in detail below when the pagination program (P8) is discussed.

Having constructed a set of skeleton TE records, the program initiates requests to the authority file for authority text data and filing forms. Requests are also made to the authority file for headings which are to print above the bibliographic citations. These headings will be constructed in the same manner as catalog entries, i.e. as TE records. They will then be merged with the respective TE records as citation entries. These heading requests also initiate a sequence of processing steps culminating in the location and formatting of all relevant cross-references. The necessary cross-references are formatted into TE records, and are likewise merged to form the complete catalog.

When an entry is chosen for inclusion in a cumulation segment, indicators to that effect are set in the bibliographic master record; it is then written onto the updated bibliographic master file.

**Locate Authority Data and Select Headings—Program P2**

All inquiries to the authority file are sorted into authority sort key sequence and matched with the authority file. All inquiries will result in a match to a valid authority record. A match for each inquiry is assured by the weekly file update/interface processing programs.

Inquiries to the authority file result in any combination of the following actions: (1) authority text and filing data are supplied, via a response record, to program P7 for the completion of TE records created by program P1; (2) authority records are selected to serve as headings above bibliographic citations (these same records will also cause cross-references to be selected); (3) authority records are selected in order to initiate a search for the associated dual entry, as per instructions contained in the inquiry record.

The selected headings consist of complete authority records with instructions regarding their eventual use and routing. Headings are routed, via a collation code, into cumulation segments or supplements. Since a single authority heading may appear as both a main entry and subject heading, indicators are set defining its eventual use as one, the other, or both. These indicators will be called usage indicators. Usage decisions made by P1 are passed to this step as part of the inquiry records. The results of these decisions are then transmitted as a set of codes inserted into the selected authority records.
This program is further charged with the responsibility of keeping current the catalog status indicators for cross-references by maintaining two binary indicators with every cross-reference. A cross-reference record with multiple see fields will have a pair of indicators for each see field. The first binary indicator denotes prior appearance of a cross-reference in a cumulation segment. The second indicates that the referenced heading currently appears in some part of the catalog. In passing through the entire authority file, this program will note that a heading which falls in the current month's filing range has had no requests for its use lodged against it. When this is the case, transactions are created for every cross-reference, defined by see froms in the heading record, extinguishing the second binary indicator described above. The cross-reference will then not be used again until it is required. The need for this operation will become more evident when we discuss program P6.

The maintenance of the physical linkage between cross-references and headings is performed by the authority file update subsystem. This subsystem guarantees that the linkage is kept current regardless of alterations to headings and cross-references. Hence, all see froms are guaranteed to refer to a cross-reference (direct see) record on the file.

**Explode Hierarchies, Cross-references and Dual Entries—Program P3.1**

The selected authority records are examined for the presence of see from fields. If any are found, they are used to create further inquiries to the authority file for cross-references. A similar operation is performed for dual entries with the exception that the dual entry inquiry is not created unless it was requested by program P1. The request is passed via indicators in the inquiry record (as discussed above in the description of program P2).

All records which are subdivisions of headings, e.g. Sculpture—Technique, will cause inquiries for all significant higher level headings (Sculpture in this case) to be created. Higher level headings will supply additional entry points via cross-references to them, or may themselves appear if they contain notes.

Cross-reference requests are separated for later processing. They will be processed with requests for secondary cross-references to be generated by program P3.2 below.

**Exclude Duplicate Headings and Separate Inquiries—Program P4**

This program is nothing more than a sort with exits. The input tape of selected headings and higher level heading requests is sorted, and if a request for a higher level heading has already been filled by a heading selected in P2, the request is dropped. All usage information carried by the request is logically added to the matching heading. When multiple requests for the same higher level heading are discovered, all but the first are
dropped. Usage information from all duplicates is added to the retained request by a logical OR operation.

The authority records which were selected by P2 for use as headings are formatted into complete text entry (TE) records for later input to the pagination program. TE heading records are formatted by a single module invoked by this step and again in P5.

The surviving hierarchy requests, and all dual entry requests are separated for processing in the next step.

Format Headings Module

All heading records selected for print are processed by this module, which converts the input text and filing data of authority records into TE records. At times quasi-duplicates of the TE record are constructed with different filing and typography codes for use as main entry and subject headings. At times portions of the data are encoded as nonprinting because it is known that the print data will be provided by other heading records. This is the case with author/conventional title records. The author heading is assured because of the explosion of higher level headings; hence, a simple method is provided for insuring its appearance only once regardless of the number of associated conventional titles.

When a subject heading record is created, the heading is made to appear twice in the record, once in upper case for printing, and once in its normal upper and lower case form, encoded as nonprinting, for possible use as a dictionary heading by the pagination program. The conversion to upper case is effected via a translate table, because of the presence of control information within the text for floating diacritics. Also, diacritics and many special characters do not have a simple upper case equivalent due to the use of the complete ALA character set.

Punctuation of cross-references is effected in this module. The complexities by no means approach those encountered in punctuating condensed added entries; nonetheless, they do exist. For example, terminal periods in headings referenced in a cross-reference must be replaced with semicolons when more than one heading is referenced, a blank must be inserted following the hyphen and preceding the semicolon in open ended dates, the final referenced heading in a string must end in a period unless it terminates with a hyphen, quote mark, exclamation point, question mark, parenthesis, etc.

Typographic codes which apply to headings, notes associated with headings, and phrases in cross-references are inserted by this program when TE records are created.

Locate Hierarchies and Dual Entries—Program P5

All heading requests are applied to the authority master file. When the heading corresponding to a request is located, the entire authority record is written onto an output file for further processing. This process is sim-
ilar to that executed when the original heading requests were processed in program P2. Higher level headings are encoded for use in accordance with their categorization and filing form.

When a requested dual entry heading is located, a TE record is written for later processing by the pagination program. A response record containing the filing form of the dual entry is also written onto an indexed sequential disk file. A direct access file is necessary since the catalog record contains only a link to the primary heading, and all requests for the dual entry come via a request against the primary heading in program P2. Rather than attempting a complex scheme for keeping track of all bibliographic items requiring the dual entry data, only one copy of the dual entry response is isolated and indexed by the control number of the primary heading. It is then retrieved on that basis when needed.

**Explode Secondary Cross-references, Separate and Select Hierarchical Headings—Program P3.2**

This program is simply a phase of program P3.1 described above. The major difference lies in its handling of the authority records which it accepts as input. They are written out as TE records, but only if they meet one of two conditions: if the authority record matching the heading request contains notes, it is selected for eventual formatting into a heading; or if it represents an author, required of an author/conventional title combination. In all other instances higher level headings are not selected for printing. The Format Headings module is invoked by this step for all higher level headings selected for print. If secondary cross-references are not desired, the explosion module which creates the requests is simply bypassed. Similarly, higher level headings may be suppressed.

No further attempt is made to generate higher level headings, as they have all been exploded in P3.1. The exploded cross-reference requests are separated in this program, just as they were in P3.1.

**Locate Cross-references—Program P6**

Prior to execution of this step tapes T3.1, T3.2, T3.3 are sort/merged into a single tape T3.4 (Figure 9). T3.4 now contains all of the transactions generated by program P2, and all cross-reference requests. Recall that P2 has created transactions extinguishing the indicator carried by cross-reference headings, denoting that the referenced heading appears somewhere in the catalog. The sort causes all of these transactions to be applied before any cross-reference requests are processed. It might appear a bit paradoxical that a request should be made to a cross-reference whose referenced heading was not selected in P2; however, recall that a cross-reference may be invoked as the result of the use of a subdivision of the referenced heading (secondary cross-reference).

At this point some discussion of the cross-reference record is in order. A cross-reference may point to several headings simultaneously, e.g. **Ani**
mals see Aardvarks/Bears/Cats/...Zebras. Each referenced heading is controlled individually. Only the required references are extracted as needed. In the example above, if Aardvarks and Cats appeared in the catalog those two references would have been selected, and no others.

Hence, the discussion which follows will be greatly simplified if we consider each cross-reference transaction to apply to only a single reference. This is effected operationally by carrying the control number of the heading which gave rise to the cross-reference request within the request.

Following the application of transactions, if any, to extinguish indicators, the selection for print logic is executed. Cross-references are selected for printing when the indicator specifies that the cross-referenced heading appears somewhere in the catalog available to the public, regardless of whether there is a specific request for it, and the cross-reference is filed in the segment being produced. A request for a cross-reference which already appears in a cumulation segment currently in use is ignored. A request for a cross-reference which is not already in the catalog is honored. The actual logic is somewhat complex; however, the end result is as described above.

Cross-references to be printed are routed to either a supplement or cumulation installment depending upon the filing range in which they fall. When a divided catalog is being produced cross-references are further routed into the appropriate catalog on the basis of categorization. Following the selection of, or refusal to select, a heading, the indicators denoting prior appearance in the catalog and linkage to a heading in use are updated. Continuing integrity of the cross-reference structure for future printings of the catalog is thus assured.

**Complete Citation Text Entry Records—Program P7**

Prior to execution of this processing step, response records emanating from P2 are sorted into bibliographic item number sequence. Sequencing is necessary since the skeleton TE records are in the same sequence as the bibliographic master file. Identification of authority response data required by a TE record is via bibliographic item number and a sequence number assigned to each authority field within a bibliographic record. Subfields of a response record are identified by delimiter.

Response records are matched to skeleton TE records bearing the same item number. Following the match, all required data are inserted into the skeleton TE record. Codes are carried in the TE record directing this program to perform certain formatting functions not possible in step P1. These functions include insertion of certain combinations of parentheses and brackets required by series notes, addition of a series note to certain call numbers, and the replacement of the author portion of an author-title combination series note with His:, Her:, In His:, In Her:, etc. None of the above could have been accomplished in a typographically acceptable manner in program P1.

Dual entry data are obtained from the indexed sequential file (D1). The
Automated Book Catalog Subsystem

Identification of such data is via the authority control number of the primary LC subject heading carried in the bibliographic record. This number is used to access file D1 for the required text and filing data.

Pagination—Program P8

Prior to execution of this step a set of page initialization records is created for the particular type of catalog being produced. These records are prepared by a program not shown in the subsystem flow. Initialization records govern the overall format of the book to be produced. There are six such initialization records, all of which must appear at the beginning of the input tape. They may also appear embedded anywhere among the TE records in various combinations.

The first initialization record, known as a Page Dimension (PD) record, defines the physical dimensions of the page to be printed. Parameters carried in this record also determine the dimensions of inner and outer page margins, head and foot margins (independently for recto and verso pages), number and width of columns, body size on which to set type, and spacing between entries. When an embedded PD record is encountered the program will terminate any page currently being formatted, begin a new page, and continue formatting in accordance with the redefined dimensions.

The second initialization record defines the starting page number, and indicates whether paging is to start with a recto or verso page. The pagination program may also be directed via this record to place a black square at the edge of a page, at a location defined by the record, to serve as a thumb index. This record may also appear anywhere else on the tape. When it does appear as an embedded record it commands the program to terminate the page being formatted at that point, to begin a new page, and possibly provide a number of blank pages. This allows volumes to be broken at predefined sort points. In this manner we may separate alphabetic segments, the various volumes of a divided catalog, or cumulation and supplement volumes, and move the thumb index.

Subsequently four records define caption and legend text (independently for recto and verso pages). Any one or combination of these records may also occur elsewhere on the tape. When they do occur as embedded records, the program terminates the page currently being formatted, alters the appropriate caption and/or legend text, and continues to format text. Interfiling of these records with TE records allows captions to be changed automatically between volumes of a divided catalog, or between supplement and cumulation volumes, or at any other desired sort point.

The six records described above control those aspects of page format which are common to a large class of entries. Individual TE records carry typographic commands which are specific to the entry, or to an element of the entry. A code carried by each TE record (Entry Format Code) defines typographical rules for the entry as a whole. This code is used to identify
data to be used in the formation of dictionary and column headings when page breaks occur. Certain widow rules affecting the entire entry are specified, e.g. entry may not span columns, entry may not form the last line of a column, etc. Line advance commands, defining the amount of space (if any) to be left between entries, are carried in this code.

Data elements within an entry may require different typographic rules. Format codes for each such element are carried within a record directory. The directory also serves to identify the location and length of text data to be typeset in accordance with the typography specified by Element Format Codes. Element Format Codes consist of 32 bit fullwords. Groups of bits within the word define separate typographic rules. These bits may be set in any combination, defining a complete spectrum of typography. The major typographic parameters governed by these bit settings are:

1. Starting indention ("continue on the previously used line" is included).
2. Overflow indention to be used if the element must be continued onto another line.
3. Space to be left on a line before adding any additional text to a previously used line.
4. Justification—left, right, center of column, and center of page.
5. Type size height.
6. Type size width relative to height.
7. Type face—bold or light.
8. Type style—Roman or Italic.
9. Element widow rules—restrictions which do not allow text to: span columns, form the first line of a column, span from a verso to a recto page, or span from a recto to a verso page.
10. Line break—indicating whether lines may be broken at blanks only, or may be broken at blanks and certain special characters. Line break decisions observe a hierarchy of rules, e.g. if the indicator is set to break at blanks only and no blanks are found within the entire line, the program automatically reverts to the second option (break at blanks and special characters); should that also fail, the line will be broken arbitrarily at the last character which fits on the line.
11. Hyphenation indicator—due to the great number of foreign languages used in the NYPL catalog no hyphenation routine is employed. Allowance has been made, however, for the inclusion of a hyphenation module should it be desired in the future and an indicator provided in order to invoke it.

Other rules of lesser importance exist, but space does not warrant their discussion.

The entire ALA character set plus several additional characters specified by NYPL may be typeset via this program on an III Videocomp. Diacritics are floated onto the characters they accent. The coding structure adopted by NYPL consists of two unique codes preceding a pair of characters to be
overprinted. The first code indicates to all processing programs that the
data to follow must be interpreted in a unique manner. The second de­
fines the unique treatment to be accorded. We currently employ only two
such functions codes; both imply a form of overprint. Coding in this
manner allows unlimited expansion of the character set. A function code
has been assigned but not yet utilized for overprinting of triplets. This
would be necessary in handling doubly accented characters, such as are
found in Vietnamese. Functions codes have been assigned defining escapes
to nonroman alphabets.

The character set includes two blanks in addition to the normal word
space. One of these will provide a word space on printed output but will
fail line break tests. Such a character is of great utility as a separator in
abbreviations and as a word space preceding such terminal characters as a
close parenthesis. Conversion of the NYPL data base to utilize this super
blank will be effected following definition of sufficiently reliable rules for
its automatic generation at input. The second blank is a zero set width
character. This character, when present in a machine record, is assigned a
null width by the phototypesetting device. Its utility lies in areas in which
it is required to remove only one or two characters from a record, but it is
not desired to expend the programming or processing time in restructuring
the record.

All of the input text data and format codes are translated into com­
mands to an III Videocomp 830 and written onto a driver tape. The driver
tape is then delivered to a photocomposition vendor who mounts it on a
Videocomp to produce camera ready copy for catalog pages. The camera
ready copy is then delivered to a printer who produces multilith plates, and
thence, pages which are bound into monthly supplements and cumulation
segments.

CONCLUSION

Photocomposed book catalogs have been in use at NYPL since January
1971. The effectiveness of the system can, perhaps, best be judged by the
only adverse reaction received thus far: in the case of material which must
pass through the bindery after cataloging, entries appear in the catalog be­
fore the materials reach the shelves, thereby causing annoyance to users.

Judged by more serious criteria, the system has been proven to be an op­
erational success. The processing budget for the Research Libraries is now
insignificantly higher than it was under the manual system, but cataloging
volumes have increased dramatically: 7,500 titles/mo. cataloged vs. 5,500
titles/mo. under the old manual system. The increase in productivity can­
not be solely attributed to the automated system. Some of it is attributable
to the revision, by the head of Preparation Services, of manual procedures.

Expansion of Book Catalog Coverage

The entire bibliographic system is currently in the final stages of revi-
sion for production of a multimedia catalog of the Dance Collection of
the Research Library of the Performing Arts. The organization of cita-
tions referring to material in diverse media will be accomplished by
providing separate sequences under appropriate headings, denoting: Works by,
Works about, Visual works, Music, Audio materials. Listed under each of
these headings will be the following types of materials:
1. Works by—Written works by an author.
2. Works about—Written works about an author, performer, etc. (The
subheading is not used under topical subjects.)
3. Visual works—Photographs (original and indexed), prints and origi-
inal designs, motion pictures and videotapes, filmstrips and slides.
5. Audio Materials—Phono records and phonotape.
These headings are not as specific as those suggested by Riddle, et al.,
however, they do provide the early warning function discussed by Virginia
Taylor. This catalog is due for publication in early 1974. Pending the success
of this venture, a study will be made of the means of extending the scope of
the Research Libraries' catalog to include nonbook materials.
In late fall 1973, an extremely exciting and bold step will be taken by
the Jewish Division of the Research Libraries. They will begin data input
of material in Hebrew, using the recently defined ANSI correspondence
scheme for Hebrew characters. Within this scheme roman and special
keyboard characters have been assigned to each character of the Hebrew
alphabet. Book catalog display of Hebrew text will utilize these characters
in a left to right print mode until such time as development money is
found for the digitization of Hebrew character fonts, and for modifications
to the pagination program in order to display mixed roman and He-
brew text. All Hebrew entries will be filed in accordance with conventions
for sequencing Hebrew text. The Hebrew entries will be interfiled with
entries in romanized forms by conceptually assuming the sequencing al-
phabet to contain 57 characters: blank, A, B, ..., Z, 0, 1, ..., 9, N, T, ...
If we have an author who has written several titles in roman alphabet languages,
and others in Hebrew, we would create a sequence of
main entries under his name interfiled according to the alphabetic se-
quency shown above. All Hebrew or variant title added entries would be
found in a sequence starting at the end of the roman alphabet.
The primary reasons for adopting such a scheme as opposed to the more
traditional romanization are:
1. A nationally endorsed correspondence schedule has been provided by
ANSI.
2. It is desired to enter this data into the automated system and end the
manual operation at the earliest possible time.
3. It is desired not to have to revise all cataloging when true Hebrew
text may be economically displayed. It is virtually impossible to re-
cover the true form of nonroman text from its romanized form.

These two areas, nonroman alphabet display and inclusion of nonbook materials, represent the only areas in which further development of the book catalog system is planned. Future efforts will be directed to conversion of the batch-oriented processing system to one with on-line file maintenance capability.

It should be stressed again that the primary aim of the bibliographic system is not production of book catalogs. The system was designed to create a highly controlled data base which could be used in conjunction with whatever display medium is technologically and economically feasible. On-line access to the catalog will require extreme control of the data, as automated retrieval techniques require very precise definition of access points. The problems of data organization become greatly magnified when CRT display devices are used, as the visual scan range produced is severely limited.

The extensive development effort to produce book catalogs was undertaken at NYPL since it was felt that for at least the next decade book catalogs in printed or microform would provide the only economically viable form of access to the collection. Book catalogs will, no doubt, also serve as backup forms of display for a considerable time after introduction of electronic access techniques.

REFERENCES


