

# dBASE III Plus and the MARC AMC Format: Problems and Possibilities

RONALD J. ZBORAY

**Abstract:** As archives with smaller holdings adapt to the MARC AMC data format, they will need to explore the utility of various commercially-available, microprocessor-based, data base management systems. This article presents four approaches to using the AMC format in dBASE III Plus, the best selling of these systems. The discussion of the limitations and possibilities of the four approaches demonstrates that while there is no inherent incompatibility between dBASE and the MARC format, the most efficient use of a dBASE-MARC system requires considerable programming.

*About the author:* Since 1984 Ronald J. Zboray has been the Microfilm Editor of the Emma Goldman Papers at the Institute for the Study of Social Change, University of California at Berkeley. He received his B.A. in history from the University of Bridgeport and the A.M. and Ph.D. in American Civilization from New York University. He has taught American history at the University of Connecticut at Stamford, Pace University, the University of Hartford, and Post College (Waterbury, Conn.). His articles have appeared in *The International Journal of Micrographics and Video Technology*, *Film and History*, *American Quarterly*, and *Southwest Review*. For the latter he won the 1986 DeGolyer prize for best nonfiction essay in American Studies.

WITH THE INCREASING ACCEPTANCE of the MARC AMC data format by the community of archivists, archives with smaller collections which are able to afford only microcomputers will seek to use the new format for automated control of their holdings. Vendors, such as Cactus Software, have already recognized this small emerging market with records management packages that have the MARC AMC format already built in. Researchers at Michigan State University, with funding from the National Historical Publications and Records Commission, have developed yet another such package. Yet, by and large, archivists have been unhappy with the overall inflexibility and occasional unreliability of these specialized offerings. Some have limited or no networking capability, and none contain a flexible internal programming language as extensive as those in best-selling data base management systems. This dissatisfaction is not surprising; few small software publishers can afford the costs and labor time necessary to develop state-of-the-art data base management systems. Certainly the limited market of archivists who control their collections with microcomputers does not warrant larger publishers undertaking this development.<sup>1</sup>

For the foreseeable future, archives with small holdings will probably rely upon mass market, commercial data base

management systems for document control. Many archivists, before they knew much about MARC AMC, used such systems as Revelation, R:Base 5000, and, principally, the industry leader dBASE. Because of the large, highly competitive market these products address, they are generally speedy and dependable, and offer the flexibility of a wide variety of customized reporting options and data manipulations and searches—particularly necessary to the frequently unique characteristics of archival collections.<sup>2</sup>

These limited local applications of data base management systems scarcely required the descriptive depth and standardization demanded by MARC AMC. Few archivists on their own would create a data base file with MARC AMC's seventy-seven variable data fields with each having from one to twenty subfields. Through foresight or sheer luck, a handful of archivists can output data they collected in earlier systems in a form compatible with MARC AMC; far too many others cannot. To ensure compatibility and, above all, to use the new data format most effectively, archivists using microcomputer-based data base management systems should plan for AMC at the input stage when files are created.<sup>3</sup>

Some considerations for archivists planning to use AMC within a commercial data base management system will be presented below. Though the focus is on

<sup>1</sup>MARC (Peoria, Ariz.: Cactus Software, Inc.); MicroMARC:amc (East Lansing, Mich.: Michigan State University Archives and Historical Collections). See also Micro Archives and Research Collections On-Line (MARCON) (Baltimore: Automated and Information Retrieval Systems); Mark H. Kibbey, "Marcon II: Information Control on a Microcomputer," *Library Hi-Tech News* 14 (Summer 1986): 11-21.

<sup>2</sup>Revelation (Seattle: Cosmos, Inc.); R:Base Series 5000 (Bellevue, Wash.: Microrim, Inc.); dBase III Plus (Torrance, Cal.: Ashton-Tate). A good starting point for archivists thinking about implementing data base management systems is Jeanne C. Raudenbusch and Robert B. Shaklee, "Selecting Records Management Software," *Records Management Quarterly* 19 (July 1985): 22-31.

<sup>3</sup>The MARC AMC format appears in Library of Congress, *MARC Formats for Bibliographical Data*, updates nos. 10 and 11 (Washington, D.C.: Library of Congress, 1984, 1985). A more usable guide is Nancy Sahli, *MARC for Archives and Manuscripts: The AMC Format* (Chicago: Society of American Archivists, 1985). Helpful for implementing MARC is Max J. Evans and Lisa B. Weber, *MARC for Archives and Manuscripts: A Compendium of Practice* (Madison: State Historical Society of Wisconsin, 1985). The literature on MARC AMC grows annually; solid introductions to key issues surrounding MARC AMC are presented in three articles in the Winter 1986 edition of the *American Archivist* (vol. 49: 9-40): Nancy A. Sahli, "Interpretation and Application of the AMC Format"; Katherine D. Morton, "The MARC Formats: An Overview"; and Steven L. Hensen, "The Use of Standards in the Application of the AMC Format."

dBASE III Plus, most of the viewpoints can be applied to other systems, since most of them have similar features and, for archivists, similar limitations. dBASE was chosen because it is the best-selling commercial data base management system for microcomputers. In the discussion below, examples will be drawn from the experiences of the Emma Goldman Papers Project at the University of California at Berkeley as it faces the challenge of making dBASE compatible with MARC AMC.<sup>4</sup> After a brief overview of dBASE's capabilities, four different approaches to using dBASE for AMC will be explored. Each approach requires an increasing level of programming sophistication yet yields a more efficient use of hard disk storage space.

The programming required to use dBASE for AMC most efficiently may well be used as an argument against using dBASE at all for this purpose. As the discussion below will demonstrate, dBASE can be adapted to the needs of the AMC format only with difficulty or within severe limitations of disk space usage.

Nevertheless, the installed base of dBASE users among archives will inevitably continue to grow, and many archivists may have little alternative to dBASE for collection control. Though dBASE has a steep learning curve, the program encourages experienced users to dabble in

the internal programming language where the true flexibility of the system emerges; the expertise for more sophisticated adaptations of dBASE to the AMC format may thus eventually be found among the staff of an archive. Archivists may also use dBASE for a host of different applications, from keeping addresses of correspondents to keeping track of volunteers or donors. Staff members must master only one system for these different purposes. Because dBASE, despite its limitations, will continue to be used by archivists, the suitability of the four approaches will be discussed below.

### Capacity and Capabilities

At first glance dBASE III Plus seems to be adequate to the needs of even moderately large archives. Up to a billion records and a total of 2 billion bytes can be controlled. A maximum of 128 fields promises to fit easily MARC AMC's requirement for 77 fields. Although the maximum record size of 4,000 bytes may barely contain the information required by the MARC format, dBASE permits variable-length memo fields that can run to 512 kilobytes.<sup>5</sup>

In practice, for use on microcomputers, dBASE's specifications make it suited only in archives with small to moderate-sized holdings. dBASE demands that a single, entire data base file

<sup>4</sup>The Emma Goldman Papers, in existence since 1980 under the directorship of Dr. Candace Falk, is an NEH- and NHPRC-sponsored project collecting Goldman's published and unpublished works and correspondence and government documents pertaining to her. Supported by a guide and in-depth, three-volume index, a microform edition of these materials will be published by Chadwyck-Healey, Inc.; completion is expected in 1991.

<sup>5</sup>dBase III Plus differs from dBase II and III in several ways. The Plus version, released in 1986, offers multi-user operation, the ability to control vastly more files, optional menu-driven operation and dozens of other new features. A clumsy copy protection in the first release has recently been removed. Most library and archival literature has not yet had time to catch up with the new upgrade. Since there are strong similarities between the various versions, some of the older literature may still be relevant. See Roger D. Palmer, *dBase II & dBase III: An Introduction for Information Services* (Studio City, Cal.: Pacific Information, 1984); and the various regular columns in *Library Software Review* beginning with Randolph Gadikian, "dBase II in Libraries" 3 (1984): 521-26. Although more and more general sources have appeared for dBase III Plus, one of the best is Luis Castro, Jay Hanson, and Tom Rettig, *Advanced Programmer's Guide Featuring dBASE II, dBASE III, and dBASE III Plus* (Culver City, Cal.: Ashton-Tate, 1985). For good introductions to dBASE III Plus, see Robert A. Byers, *Everyman's Data Base Primer Featuring dBASE III Plus* (Torrance, Cal.: Ashton-Tate, 1986) and Howard Dickler, *dBASE II Plus Trail Guide* (Torrance, Cal.: Ashton-Tate, 1986).

be kept on one disk or diskette. Since Microsoft DOS has difficulty addressing a file greater than 32 megabytes on a single storage media, the total record capacity of dBASE falls far, far short of the theoretical 1 billion limit. That each variable memo field uses a minimum of 512 bytes when data is entered even more severely restricts the total number of records that can be controlled.<sup>6</sup>

The most obvious limitation of dBASE (and nearly all other data base management systems except Revelation) is the fixed-length field. Because the variable-length memo fields cannot be searched or manipulated by dBASE in any way, the fixed-length fields remain the heart of the system. Fixed length fields waste immense amounts of disk space (about one-half to two-thirds of capacity), further limiting the maximum number of records that can be stored on microcomputer with a fixed disk. The business applications for which dBASE is designed will often set a reasonable length on, for example, a personal name field, and truncate names that go beyond that length. Standard archival practice allows for no such shortening; the longest field entry in a file will determine the length of the field, even if that entry is singularly long. Thus for archives, the potential for wasting disk space with the fixed-length field far exceeds that of other applications.

The MARC AMC format itself encourages a waste of disk storage space. Relatively few data control records will use all seventy-seven fields and perhaps only some will use more than a majority. Certainly if an archive settles the first or even second level of AACR2 description, even fewer fields will be used.<sup>7</sup> But such avoidance of the descriptive depth offered by level three in conjunction with MARC AMC may be shortsighted, since

points of access considered excessive today too often become expected in the future.

In the discussion below the fullest implementation of MARC AMC will be considered. Control will be assumed to be at the document level and the full AMC format will be used even though most archives would never have need to implement the format to such an extent.

### **Approach I: Using dBASE Directly Off the Shelf**

The ability of very novice users of dBASE to construct and manipulate data files with relative ease accounts for much of dBASE's success. Employing either the "CREATE «filename»" command at the dot prompt or the menu-driven, pull-down "Assistant," the user who wants to design a data base needs to do little more than name fields, determine whether the type of data in each field should be character, numerical, logical, date, or memo (text file), and decide how many characters each fixed-length field will contain.

Table 1 shows the structure of a viable, easily-constructed dBASE file that conforms to the requirements of MARC AMC. The estimated fixed-field lengths range from 12 characters for Tag 007 (category of materials) to 250 characters for Tags 600 and 650 (subject added personal names and topics, respectively). Since all the fields, except one, use character-type data this information has not been displayed. Because dBASE permits only field names of ten characters or less, these names have been mnemonically derived in the table (e.g., "LCNUMBER" in Tag 010 stands for Library of Congress control number).

Six variable length fields do not appear. Tag 002, the subrecord map of the directory, can be generated by the pro-

<sup>6</sup>Editing memo fields adds to their length; the old version is saved in the data base text file and even subsequent edits are saved. dBASE, however, will release the space used by these older versions if the file is copied.

<sup>7</sup>*Anglo-American Cataloging Rules*, 2d. ed. (Chicago: American Library Association, 1978), 15.



gram that outputs the information in the AMC format.<sup>8</sup> The field for character sets specified (Tag 006), not a required field in AMC, has not been used because dBASE stores data in ASCII format. The remaining four fields representing Tags 870 to 873 have not been included because these have fallen into widespread disuse since the Library of Congress's adoption of AACR2 in 1981.

Table 1 shows only forty-eight fields rather than the seventy-seven prescribed by MARC AMC. The difference arises out of combining Tags 500 through 59X—the notes section of the AMC format—into one variable-length field. These basically textual note fields use enormous amounts of space; without combining them, the total number of bytes used by the fixed length exceeds the 4,000 byte per record limit of dBASE by nearly 75 percent. The combined notes field sacrifices little flexibility since the corresponding AMC fields seldom require computer manipulation aside from free text searching. Although dBASE III Plus does not provide free text searching of memo fields, at least one very inexpensive utility program does.<sup>9</sup> One, rather than several, memo fields has been used to save disk space; as mentioned earlier, every time data is entered into a memo a minimum of 512 bytes are used. Users access the memo field by calling up the record to be edited, placing the dBASE blinking cursor on the field filled with the word "memo," and typing control page down ("Ctrl-PgDn"). A text file (in either the Wordstar-based dBASE editing mode or a word processing program of the user's choosing) appears on the screen, and data may be entered in text form with the Tag number identifying each field. Exiting the editing mode or

word processor returns the user to the memo field in the controlling data base record.

Even when using memo fields for notes, this off-the-shelf approach wastes space promiscuously. Most fields will be used rarely, yet the possibility of their use requires that they be included. For example, the three fields for conference and meeting information (Tags 111, 611, and 711), which are not always relevant to manuscripts and archives, together use 310 bytes. Shortening these fields might make it impossible to enter the full name of the rare conferences or meetings—which often have very long names—that occur in a collection. The space wasted by such uncommonly-used fields in addition to that lost due to the fixed-length field basis of the system may amount to as much as 95 percent of disk capacity.

The maximum number of records that can be stored on a hard disk using this approach can be estimated easily. The fixed-length fields use up 3,814 bytes per record and the variable-length notes field a minimum of 512 bytes. Since the notes field embraces so many fields archivists commonly use, 1,024 bytes per record is probably a more realistic estimate than the minimum 512. Thus, total number of bytes used by each record equals 4,838 (the use of the memo field permits an excess of the 4,000-byte dBASE maximum). That number divided into the thirty megabytes of the largest single storage media DOS can address suggests an absolute maximum of 6,200 records on a hard disk with that capacity. But the performance of a hard disk slows dramatically as it approaches full disk capacity. For best performance, a 30 megabyte hard disk should hold no more than 5,000 records. For many archives, especially ones con-

<sup>8</sup>Tags Leader/00 to Leader/23 and Directory/00 to Directory/11 also do not appear because they are best generated at the program level during data output.

<sup>9</sup>dBASE III Memo Searcher (San Diego: Data Based Solutions).

trolling materials only at the collection level, this maximum number of records may be ideal.

Such a scheme, however, provides only for the storing of records; to get the most out of dBASE, users must build index files based on the data base. A single index file (e.g., main entry–personal name) uses only about 10 or 20 percent of the size of the related data base, but several such files must be created. Index files are especially necessary for making a quick search for information in the main data base. Seeking indexed information (the first character and up to a hundred characters to the left) takes less than two seconds for a data base of any size; searching for character strings within fields of a moderately-sized unindexed data base can take several minutes. Fortunately, indexes can reside on a separate drive or partition, so one need only add on an extra 20 or 30 megabyte hard disk to make use of this important feature.

Using dBASE directly off the shelf has advantages and disadvantages, the balance of which changes depending upon the nature of the archive and the resources available. Within only an hour or so, users can create a workable MARC AMC-based system and begin to enter data, with little or no support from programmers or consultants. With programmers charging about \$50 an hour, such user independence may be well worth the tens of millions of wasted bytes of storage the off-the-shelf approach requires. Storage media increasingly becomes less expensive so that one can anticipate paying for every ten megabytes of storage capacity about \$250 or less—a one-time, certain cost that may compare favorably with the unpredictable and chronic costs of customized programming. The chief disadvantage of the off-the-shelf approach is the limited number of records that can be controlled. Even if

the user tinkers with field lengths to make the data records most suited to the collection and thus cuts down the number of bytes used by each record, such gains would be minimal. There would still be a limit of about 5,000 records, with no room for further growth, unless the user turns to one of the other approaches explained below.

### Approach II: Simple Dependency

The 5,000 record limit of the off-the-shelf approach can be most easily overcome by splitting the single data base outlined in Table 1 into two dependent data bases. The two data bases must be linked through a common field, indexed in at least one. The common field may be the record number or any *unique* finding number in any of the MARC format fields. Whatever the source of the linkage, the numbers or characters must appear in only one record in the data base or the linkage will break down.

Creating dependent data bases requires delving deeply into the dBASE III Plus manual. Although the menu-driven “Assistant” promises to guide the user through the steps to link data bases, there is difficulty writing programs when the data bases are on two different disks.

The first step in creating simple data base dependency consists of constructing the two data bases on two different hard disks. In the data base structure presented in Table 1, a natural division occurs between field number 27 (notes) and field 28 (subject-added personal names). Fields 1 to 27 total an estimated length of 2,478 bytes leaving 2,360 bytes for the remaining fields. If the record number will not provide the linkage between the two data bases, the slightly smaller second data base will need an extra field to duplicate the linking expression in the first. For example, if field 6, “SYSNUMBER” (the local system control number used by

many archives to store the accession number) is the link, the same field must appear in the second data base.

With the two data bases constructed on different disks the second step of creating dependency requires creating a format file. The format file will link fields from the two data bases on the CRT screen so that the existence of two data bases will become undetectable during data entry. The dBASE text editor can be used to create the format file (while dBASE runs) with the "MODIFY COMMAND «filename.fmt»" command, or a word processing program may be used as long it will not enter any invisible page formatting characters in the text.

Next the user must tell the program where on the screen to put each field name and the fixed number of spaces for each field. dBASE addresses the screen using "@" with row and column coordinates (e.g., "@ 12,40" represents "at row 12, column 40"). The field names are typed on the screen by using the "SAY" command and the blanks are painted by the "GET" command. When two data bases are being used simultaneously, the "GET" command demands that a pointer ("->") be used to identify the source of the field being read. If the first data base is named MARCAMC1 and the second MARCAMC2, then any number written in the blanks painted by the command "@ 7,12 GET MARCAMC1 -> SYSNUMBER" will be entered only in the first data base.

Figure 1 shows the text of a program that will create a screen form for entering into two data bases the fields presented in Table 1. Although the two data bases may be called MARCAMC1 and MARCAMC2, dBASE permits using a shorter name or "alias" for programming purposes. In Table 2 "ONE" represents the alias for the first data base and "TWO" for the second, thus "ONE -> SYSNUMBER" is used instead of "MARCAMC1 -> SYSNUMBER." Fi-

nally, the "SAY" and "GET" command lines are interrupted every so often by a "READ" command that creates page breaks in the data entry form; the user will be able to access the entire form at any time during entry using the "PgUp" and "PgDn" keys.

The final step in creating a simple dependent data base relationship is to write a short program to create the dependency whenever the program is loaded. Again the dBASE text editor can be used to open the program file (do not use any three-character file extension, i.e., filename.ext). Commands are entered line by line in dBASE's easy-to-use programming language. For example, assume that there are: 1) two hard disks (C and D); 2) two data base files named MARCAMC1 and MARCAMC2 on C and D respectively; and 3) a format file called MARCAMC.FMT. In this instance, the short program would read:

```
SELECT 1
USE C:MARCAMC1 ALIAS ONE
SELECT 2
USE D:MARCAMC2 ALIAS TWO
SELECT 1
SET RELATION TO RECNO( ) INTO TWO
SET FORMAT TO MARCAMC
RETURN.
```

The program selects two work areas to be open simultaneously: area 1 with MARCAMC1 (under the alias name "ONE") on disk C and area 2 that has MARCAMC2 ("TWO") on D. The "SET RELATION TO . . ." command creates the data base dependency so that when a record pointer moves in one, it moves in the other as well. The example uses the record number ("RECNO( )") as the point of interaction between the two data bases. Finally, the "SET FORMAT TO . . ." command will call up the format file when either the standard "EDIT" or "APPEND" commands are

entered at the dot prompt (the prompt reappears after the program issues the "RETURN" command). This program must be run every time the user loads dBASE; to load it at the dot prompt, the user types "DO «program filename»."

Although creating this simple dependency may seem daunting because of the amount of programming involved, most of the programming (i.e., the format file) is redundant and quite simple, even for novices. Much of this programming can be avoided if the user is willing to put up with a more obvious split between the two data bases during data entry. If so, the user would use the same short loading program given above without the "SET FORMAT TO . . ." command. The record in the first data base would be edited and written to the disk and the dot prompt will appear. The user then would change the currently selected work area by typing "SELECT 2." By entering "EDIT," the continuation of the record on the second data base appears. At the dot prompt after completion, "SELECT 1" must be entered again before beginning to edit another record.

Simple data base dependency, however achieved, has one distinct advantage over the off-the-shelf approach. Since dependency allows the forty-eight fields presented in Table 1 to be distributed between two hard disks, the maximum number of records that can be controlled doubles (if the two disks have the same capacity). Thus using the byte-wasting data base structure presented above, a user who creates a simple dependency between two data bases on two 30 megabyte disks or disk partitions will be able to control about 10,000 records (with the various required indexes stored on yet another hard disk).

The simple data base dependency approach to creating MARC AMC records

in dBASE probably is best suited to archives with smaller holdings which intend to control materials at the collection level. The archive should expect no dramatic growth since the 10,000 record limit is as fixed as the 5,000 limit of the off-the-shelf approach. Apart from hard disk space, the dependency approach is relatively inexpensive. The limited programming involved can be accomplished by anyone with a working knowledge of dBASE. Even if outside programmers must be hired to set up the dependency, the simplicity of the programming would keep costs down and limit them to the period in which the system is set up.

### Approach III: Multiple Dependency

The two approaches outlined above waste large amounts of disk space in order to set up, with minimal programming, small document control systems. More sophisticated programming offers the ability to cut down the number of bytes used by each record as well as to increase dramatically the number of records a system can control. Advanced dBASE programming becomes highly creative at this level, and a myriad of approaches can be pursued to achieve greater efficiency. Since the details of any of these approaches are so highly technical as to confuse most non-programmers, the nature of the discussion below turns from "How-to" to "What-can-be-done."

Advanced dBASE programming can significantly reduce the number of bytes used by records in the MARC AMC format by creating multiple dependent data bases. Rather than splitting a preformatted entry form such as the one presented in Table 1 (done with simple dependency), multiple dependency uses many dependent data bases to replace short codes in a main data base record.<sup>10</sup>

<sup>10</sup>See "Database Design," chap. 9 in Castro, Hanson, and Rettig, *Advanced Programmer's Guide*.

Generally, any field with standardized, recurring information can be rendered more efficiently using codes. MARC AMC has several such fields, ten of which occur in Tags 600 to 69X, the subject indexing fields of MARC. These fields use 1,510 of the 4,838 bytes in the sample record structure presented in Table 1. The fields contain, for the most part, standardized subject entries that appear repeatedly in many archives. The entries can be replaced, as has been the case at the Goldman Papers, by four-letter mnemonically derived codes with a space separating each code for clarity. For example, the Goldman Papers encodes "Theater" as "thea" and "Berkman, Alexander" as "beal" (i.e., the first two letters of the last and first names). Replacing the text-oriented subject fields with codes allows all ten subject fields to be combined into one. A subject field of 150 bytes can contain a maximum of 15 codes—probably more entries that could be contained in the 1,510 byte text-oriented field without codes.<sup>11</sup>

Figure 2 illustrates the manner in which the codes in the main data base record interact with the same codes in a dependent data base of subjects. Each record in the subject data base should contain at least four fields: the four-letter code, the entry itself, a two-letter code for type of entry (i.e., "PN" equals "Personal name"), and a logical field that queries if subfields have been used. The use of subfields requires ten data bases for each type of entry; the subfield and subject data bases also interact through the four-letter code.

One key advantage of using codes for subjects is that the dependent data base acts as an authority file. During final MARC AMC output, a single entry in the data base will replace all the corresponding codes throughout the entire main data

base. Such a find-and-replace operation ensures uniform spelling and form for subject entries. Entering a four-letter mnemonic code allows smaller room for error than correctly retyping the same full entries over and over. Still, errors may slip by and codes may be duplicated. Because of these possibilities for human error, the Goldman Papers employs a menu-driven subject assigning program that searches the dependent data base for code duplications and allows any new entries to be appended immediately to the dependent data base.

The amount of disk space such a system saves depends upon the nature of the archive. For control at the document level, as in the Goldman Papers and in most other documentary editing projects where the materials have a common focus, the great repetition of subject entries results in a dependent data base using only 10 percent or less of the bytes that full subject entries in the main record would require. Archives having a wide diversity of materials would employ a greater number of unique subject entries, thus slightly decreasing the amount of disk space saved.

In addition to reducing the space taken up by subject added entries, dBASE programming can also minimize the disk space used by rarely used fields. The greatest savings of this sort can be achieved in Tags 700 to 752, the added entries, which use up 580 bytes in the sample record structure in Table 1. The approach to rarely used fields differs from that of the ubiquitous subjects. Instead of a large field containing codes, rarely used fields can be best treated by creating a logical field (true or false) in the main data base, a field that essentially asks if added entries exist for the record. If so, a dependent data base for added entries can interact with the main data

<sup>11</sup>The Goldman Papers' computer system is described in Ronald J. Zboray, "Microfilm Editions of Personal Papers and Microcomputers: Indexing the Emma Goldman Papers," *International Journal of Micrographics and Video Technology*, 5 (1986) 213-21.



base using the record number, local system control number, or other information unique to the main record. The added entry dependent data base could have more than one record, each with the interacting code, from each main data base record; each dependent data base record should have a field to identify the type of entry (e.g., personal name or uniform title heading).

A similar approach can be followed for fields that tend toward redundancy, such as Tags 100 to 130 for main entries and Tags 240 to 245 containing various forms of titles. These fields have a total of 720 bytes in the sample form in Table 1. A single field each for main entry and title eliminates much of the redundancy. Adding a two byte field for the corresponding field in MARC AMC (e.g., personal or corporate name, uniform title or title statement) addresses any confusion about the nature of the field. Another single byte logical field can ask if other main entry or title fields have been used. If so, a dependent data base can be constructed to store the extra entries, much as was the case with the rarely used fields.

Paring down the number of fields in this manner not only makes data entry easier, but also dramatically reduces the number of bytes each record requires. Table 2 shows a record structure of 32 fields containing only 1,035 bytes in fixed length fields. With the addition of the estimated 1,024 bytes used in the variable length notes field, the total number of bytes amounts to 2,059—only 42 percent of the 4,838 in the off-the-shelf approach previously described. The reduced number of bytes doubles the maximum number of records on a 30 megabyte hard disk from 5,000 to approximately 10,000. Combined with simple dependency (i.e., splitting the main data base record between two disks), the multiple dependent approach can accommodate on two 30 megabyte hard disks 20,000 records—a

number adequate to the needs of many archives with moderate-sized holdings which control data at the collection level. The disk space required to store the multiple dependent data bases can be stored on a third hard disk. As suggested above, the number of bytes the dependent data bases require depends on how many unique entries are used in coded fields and how many times rarely used fields are employed. Even the limit of 20,000 records can be far surpassed if the data base is further split up (i.e., the notes field has its own data base) and distributed between the partitions of a 100 megabyte or greater capacity disk or between several hard disks on a local area network.

Against the advantages of increased record capacity must be placed the significant amount of programming necessary to set up and maintain the multiple data bases. Archives with a staff member adept at moderate-level dBASE programming might find the multiple dependency approach suitable to their needs. Hiring outside programmers may not be worth the expense for an archive aiming to control less than 10,000 records, for whom simple dependency might be more suited. Nevertheless, the authority file-type control offered by using multiple data bases—and the consequent savings in input effort and the decrease in the number of errors—may be attractive to even archives with less extensive holdings.

#### **Approach IV: Menu-Driven Programs**

A menu-driven approach to using the MARC AMC format in dBASE III Plus offers a wide number of advantages over the other approaches at the cost of a great deal of program development. A menu-driven program usually can be used by staff members and even volunteers with little or no computing experience. Such a program can include error and format checking routines and “help” options



(e.g., the text of relevant portions of the MARC AMC manual or AACR2) in order to guide the user with relatively little effort through the maze of rules and procedures. It takes a die-hard computer aficionado to feel comfortable entering "\$z" as subfield delimiters in the middle of otherwise text-based materials; a menu-driven program itself generates this type of input without the assistance of the user.

As another advantage, the menu-driven approach uses disk space much more efficiently than the other methods described above. Not only can this approach employ to the fullest extent multiple dependency, it also can address to a limited degree the problem of fixed fields. One way to accomplish this is to create a record with all the field names listed in Table 1 and to characterize the fields as logical (i.e., a one-byte true or false entry). The logical fields will indicate whether information exists in a dependent data base for that field. Then separate dependent data bases for all the fields, including the notes, may be created that will interact with the main record by its record number or some other number unique to it. The length of the fields in the dependent data bases is optional, because the program can combine several records to reconstruct a single entry; the smaller the length the more the system approaches the disk space usage of a variable-length field based system. A sixty-eight character field width may be adequate for most applications; this length is easy to use since it represents one line of characters across the screen (minus the twelve bytes used for a field name). At the time of data entry the program either can paint the screen with a form with very large spaces for informa-

tion or it can, in an unformatted mode, prompt the user for information for storage to a memory variable, the contents of which the program will eventually split up into standard field lengths and distribute through the relevant dependent data base. While the amount of disk space such a system will use depends upon the particular application and the length of fields in the dependent data bases, most archives could realize 60 to 80 percent usage of disk capacity, compared to an estimated 5 percent for approaches I and II discussed above.

By making the main data base a logical record and with the increase in disk use efficiency, menu-driven programs far outpace the other approaches in the maximum number of records that can be controlled. Seventy-two logical fields in the main document control record would use only seventy-two bytes; conceivably, a thirty megabyte hard disk could hold about 350,000 such records. Of course, the considerable disk space needed for the dependent data bases must be taken into account, which could limit (to various degrees depending on available hard disk storage space) the total number of possible records. Because each field has a separate data base in such a system, the distribution of disk space is flexible and may be changed over time as needs increase. Distributed hard disks on a network or a partitioned disk of over 100 megabytes certainly would guarantee that the capacity of such a system could surpass 100,000 records. In the best of all possible worlds, with an array of networked 100-megabyte-or-greater capacity disks, the 350,000 record limit would apply.<sup>12</sup>

Figure 3 illustrates one path through a menu-driven system developed by the

<sup>12</sup>Sharing disks over a network can dramatically slow down the performance of programs. With time and particularly optical fiber cabling and the adoption of the new 20386 Intel microchip, network performance will become a smaller hindrance to distributed hard disk resources on local area networks. See William Bates and Andres G. Fortino, *dBASE III Plus & Local Area Networks* (Torrance, Cal.: Ashton-Tate, 1986).

Goldman Papers during 1985 and 1986. The user goes through three menus before entering information in a main document data entry form. Data entry is selected at the first menu and customized entry at the second. The "Regularization" option in the second data entry menu deals with the problem of removing editorial interpolations (attributions in brackets and the use of question marks) in order for dBASE to index names in proper order. The selection of "B" for "Customized data entry" brings up a menu that allows various forms of authority work and error and format checking. The use of the file name extension ".dbf" among the options in the third menu indicates a search of a dependent data base, three of which appear on the menu: biographies (of correspondents), bibliography, and institutions. The customized data entry menu leads to a main document entry form, after which subjects are entered, checked in the corresponding data bases, and the field is alphabetized. A series of error checking routines completes the program, and the user is given the option to edit another record.

Such menu-driven programs reflect the most user-friendly and efficient use of dBASE III Plus for creating MARC AMC format records. The extensive, advanced programming involved in developing such a system, however, makes it viable only for projects with an advanced dBASE programmer on their staff or for larger institutions with the need to control a great many records and the funds to support the on-going development of the program. Perhaps some enterprising, commercial developer will someday cre-

ate a standardized menu-driven program in the MARC AMC format that can run from within a dBASE III Plus environment. Until then, this approach to creating MARC AMC records in dBASE will be available to only a few archives.

### Conclusion

The discussion above has attempted to show that although dBASE can be used to create data bases in the MARC AMC format, such use operates within certain restrictions. dBASE can be used directly off the shelf by archives for this purpose, but at the expense of wasting large amounts of disk storage space. Splitting up the main data base records and distributing them between two hard disks can double the number of records that can be controlled but does little to help the disk wastage problem. A system of multiple dependent data bases begins to address the problem, but it requires a commitment to moderate-level programming that may not be affordable by many archives. A full menu-driven approach, while it provides the most efficient use of disk space, also requires extensive, advanced programming.

The four approaches, however, should not be considered mutually exclusive. Elements from any of the four can be blended to make a system most suited to the needs of a particular archive. So, ironically, the MARC AMC format, which many archivists fear may lead to overstandardization, may in its actual implementation on small data base management systems lead archivists to construct, based on knowledge of their holdings, a myriad of computer applications as unique as their own collections.

TABLE 1: A MODEL OF A DBASE III PLUS DATABASE IN THE MARC AMC FORMAT

DBASE FIELD	BYTES	TAG	MARC FIELD	DBASE FIELD	BYTES	TAG	MARC AMC FIELD
1. CONTRNO	15	001	Control number				Preferred citation of described materials
2. DATETIME	16	005	Date and time of latest transaction			524	Additional physical form
3. CATEMAT	12	007/00	Category of material			530	Available note
		007/01	Specific material designation			533	Reproduction note
		007/02	Original versus reproduction			535	Location of originals
		007/03	Availability (microforms)			540	T/duplicates
		007/04	Dimensions (microforms)				Reproduction use and reproduction
		007/05-08	Reduction ratio			541	Immediate source of acquisition
		007/09	Color (microforms)			544	Location of associated materials
		007/10	Emulsion on film			545	Biographical or historical note
		007/11	Generation (microforms)			546	Language note
		007/12	Generation (film (microforms))			555	Creative index/finding aids note
4. DATALEM	40	008	[Fixed-length data elements]				Provenance
		008/00-05	Date entered on file			561	Provenance
		008/06	Type of date code			562	Copy and version identification
		008/07-10	Date 1			565	Case file characteristics note
		008/11-17	Place of publication, production, or execution			580	Publishing entry complexity note
		008/18-22	Undefined			583	Publications note
		008/23	Form of reproduction code			584	Actions
		008/24-34	Undefined				Accumulation and frequency of use
		008/35-37	Language code	28. SUBADDPN	250	59X	Local notes
		008/38	Medium code			600	Subject added entry
		008/39	Cataloging source code	29. SUBADDCN	200	610	- personal name
5. LCNUMBER	20	010	Library of Congress control number	30. SUBADDCF	150	611	Subject added entry
6. SYSNUMBER	24	035	Local system control number	31. SUBADDUT	150	630	- corporate name
7. BIBLEVEL	15	039	and coding detail	32. SUBADDOH	250	650	Subject added entry
8. CATSOURCE	40	040	Cataloging source	33. SUBADDOG	200	651	- topical heading
9. LANGCODE	25	041	Language code	34. GENRE	30	655	Subject added entry
10. AREACODE	30	043	Geographic area code	35. OCCUPTN	150	656	Subject added entry
11. GEOCODE	30	052	Geographical code or date/time	36. FUNCTION	30	657	Index term
12. GEODE	30	052	Geographical code or date/time	37. LOCUSUB	100	69X	Local subject
13. SUBCATCODE	40	072	Subject category code	38. ADPN	100	700	Added entry
14. LOCALNUMB	15	09X	Local call numbers	39. ADPCP	100	710	- personal name
15. LOCALNUMB	100	100	Main entry - personal name	40. ADDCF	80	711	Added entry
16. PERSONAME	100	110	Main entry - corporate name	41. ADDUT	100	730	Added entry
17. CONFAME	100	111	Main entry - conference or meeting	42. ADDTD	100	740	- uniform title heading
18. TITLEHEAD	80	130	Main entry - uniform title heading	43. ADPDUB	100	752	Added entry
19. UNITITLE	80	240	Uniform title	44. PHYSACCS	25	755	Physical characteristics
20. TRANSTITLE	80	242	Translation of title by cataloging agency	45. HOSTITEM	80	773	Host item entry
21. COUNITITL	80	243	Uniform title, collective	46. LOCATION	80	851	Location
22. TITLSTATE	120	245	Title statement	47. ALGRAPH	25	880	Alternate graphic representation
23. PUBDESC	80	260	Publication, distribution, etc. (imprint)				Foreign MARC information
24. PHYSDESCR	100	300	Medium				
25. MEDIUM	40	340	Medium description				
26. ORGANIZTN	150	351	Organization and arrangement				
27. NOTES	512-512-	5XX	[Notes]				
Variable- (variable-field) memo							

TABLE 2: A MODEL OF MULTIPLE DEPENDENT DBASE III PLUS DATABASES IN THE MARC AMC FORMAT

DBASE FIELD	BYTES	TAG	MARC AMC FIELD	DBASE FIELD	BYTES	TAG	MARC AMC FIELD
1. CONTROL	15	001	Control number	22. PHYSDESCRP	100	300	Physical description
2. DATETIME	16	005	Date and time of latest transaction	23. MEDIUM	40	340	Medium
3. CATEGMAT	12	007/00	Category of material	24. ORGANIZIN	150	351	Organization and arrangement [Notes]
		007/01	Specific material designation	25. NOTES	512+	5XX	General note
		007/02	Original versus reproduction aspect		500		Dissemination note
		007/03	Polarity (microforms)		502		Contents note (formatted)
		007/04	Dimensions (microforms)		505		Restrictions on access
		007/05-08	Product (microforms)		506		Citation note (brief form)
		007/09	Emulsion (microforms)		510		/References
		007/10	Emulsion on film		520		Summary, abstract, annotation, scope, etc., note
		007/11	Generation of film		521		Users/intended audience note
		007/12	Base of film (microforms)		524		Preferred citation of described materials
4. DATELEM	40	008	[Fixed-length data elements]		530		Additional physical form
		008/00-05	Date entered on file		530		Additional reproduction note
		008/06	Type of date code		533		Reproduction note
		008/07-10	Date 1		535		Location of originals
		008/15-17	Place of publication, production, or execution code		540		/Duplications
			Undefined		540		Terms governing use and reproduction
		008/18-22	Form of reproduction code		541		Immediate source of acquisition
		008/23	Undefined		544		Location of associated materials
		008/24-34	Language code		545		Biographical or historical note
		008/35-37	Modified source code		546		Language note
		008/38	Cataloging source code		555		Cumulative index/finding aids note
5. LCNUMBER	20	010	Library of Congress control number		561		Provenance
6. SYSNUMBER	24	035	Local system control number		562		Copy and version identification
7. @BIBLEVEL	1	(035)	Level of bibliographic detail and coding detail		565		Case file characteristics note
8. @CATSOURCE	1	(040)	Cataloging source		580		Linking entry complexity note
9. @LANGCODE	1	(041)	Language code		581		Publications note
10. AREACODE	30	043	Geographic area code		584		Actions
11. CHRONCODE	30	045	Chronological code of date/time		584		Accumulation and frequency
12. GEOCODE	30	052	Geographic classification code		59X		Local notes
13. @SUBCATCODE	1	(072)	Subject category code	26. %SUBJECTS	150	(600-9X)	Subject added entry
14. LOCALNUMB	15	09X	Local call numbers	27. %ADDENRY	1	(700-52)	Added entries
15. #MAINENTRY	150	(100-90)	Main entry type: personal name, corporate name, conference or meeting	28. PHYSACCS	25	755	Physical characteristics access
16. #METITLE	2		Main entry title	29. %HOSTITEM	4	(773)	Host item entry
17. #ADMMAIN	1		Main entry - additions	30. %LOCATION	10	(851)	Location
18. #TITLE	150	(240-45)	Title	31. %ALTGRAPH	1	(880)	Alternate graphic representation
19. #TITLTYPE	2		MARC AMC type of title	32. %FORGNMARC	1	(886)	Foreign MARC information
20. #SODTITLE	1		Title - additions				field
21. #PUBDESC	1	(260)	Publication, distribution, etc. (imprint)				

DBASE field names beginning with @, # and % differ from those presented in Table 1. Those fields with @ are one byte logical fields (true or false) which indicate if information appearing in the field has been entered in a dependent database. Fields with # are combinations of potentially redundant fields and occur in series of three: 1) the field 2) the field 3) the field 4) the field 5) the field 6) the field 7) the field 8) the field 9) the field 10) the field 11) the field 12) the field 13) the field 14) the field 15) the field 16) the field 17) the field 18) the field 19) the field 20) the field 21) the field 22) the field 23) the field 24) the field 25) the field 26) the field 27) the field 28) the field 29) the field 30) the field 31) the field 32) the field 33) the field 34) the field 35) the field 36) the field 37) the field 38) the field 39) the field 40) the field 41) the field 42) the field 43) the field 44) the field 45) the field 46) the field 47) the field 48) the field 49) the field 50) the field 51) the field 52) the field 53) the field 54) the field 55) the field 56) the field 57) the field 58) the field 59) the field 60) the field 61) the field 62) the field 63) the field 64) the field 65) the field 66) the field 67) the field 68) the field 69) the field 70) the field 71) the field 72) the field 73) the field 74) the field 75) the field 76) the field 77) the field 78) the field 79) the field 80) the field 81) the field 82) the field 83) the field 84) the field 85) the field 86) the field 87) the field 88) the field 89) the field 90) the field 91) the field 92) the field 93) the field 94) the field 95) the field 96) the field 97) the field 98) the field 99) the field 100) the field 101) the field 102) the field 103) the field 104) the field 105) the field 106) the field 107) the field 108) the field 109) the field 110) the field 111) the field 112) the field 113) the field 114) the field 115) the field 116) the field 117) the field 118) the field 119) the field 120) the field 121) the field 122) the field 123) the field 124) the field 125) the field 126) the field 127) the field 128) the field 129) the field 130) the field 131) the field 132) the field 133) the field 134) the field 135) the field 136) the field 137) the field 138) the field 139) the field 140) the field 141) the field 142) the field 143) the field 144) the field 145) the field 146) the field 147) the field 148) the field 149) the field 150) the field 151) the field 152) the field 153) the field 154) the field 155) the field 156) the field 157) the field 158) the field 159) the field 160) the field 161) the field 162) the field 163) the field 164) the field 165) the field 166) the field 167) the field 168) the field 169) the field 170) the field 171) the field 172) the field 173) the field 174) the field 175) the field 176) the field 177) the field 178) the field 179) the field 180) the field 181) the field 182) the field 183) the field 184) the field 185) the field 186) the field 187) the field 188) the field 189) the field 190) the field 191) the field 192) the field 193) the field 194) the field 195) the field 196) the field 197) the field 198) the field 199) the field 200) the field 201) the field 202) the field 203) the field 204) the field 205) the field 206) the field 207) the field 208) the field 209) the field 210) the field 211) the field 212) the field 213) the field 214) the field 215) the field 216) the field 217) the field 218) the field 219) the field 220) the field 221) the field 222) the field 223) the field 224) the field 225) the field 226) the field 227) the field 228) the field 229) the field 230) the field 231) the field 232) the field 233) the field 234) the field 235) the field 236) the field 237) the field 238) the field 239) the field 240) the field 241) the field 242) the field 243) the field 244) the field 245) the field 246) the field 247) the field 248) the field 249) the field 250) the field 251) the field 252) the field 253) the field 254) the field 255) the field 256) the field 257) the field 258) the field 259) the field 260) the field 261) the field 262) the field 263) the field 264) the field 265) the field 266) the field 267) the field 268) the field 269) the field 270) the field 271) the field 272) the field 273) the field 274) the field 275) the field 276) the field 277) the field 278) the field 279) the field 280) the field 281) the field 282) the field 283) the field 284) the field 285) the field 286) the field 287) the field 288) the field 289) the field 290) the field 291) the field 292) the field 293) the field 294) the field 295) the field 296) the field 297) the field 298) the field 299) the field 300) the field 301) the field 302) the field 303) the field 304) the field 305) the field 306) the field 307) the field 308) the field 309) the field 310) the field 311) the field 312) the field 313) the field 314) the field 315) the field 316) the field 317) the field 318) the field 319) the field 320) the field 321) the field 322) the field 323) the field 324) the field 325) the field 326) the field 327) the field 328) the field 329) the field 330) the field 331) the field 332) the field 333) the field 334) the field 335) the field 336) the field 337) the field 338) the field 339) the field 340) the field 341) the field 342) the field 343) the field 344) the field 345) the field 346) the field 347) the field 348) the field 349) the field 350) the field 351) the field 352) the field 353) the field 354) the field 355) the field 356) the field 357) the field 358) the field 359) the field 360) the field 361) the field 362) the field 363) the field 364) the field 365) the field 366) the field 367) the field 368) the field 369) the field 370) the field 371) the field 372) the field 373) the field 374) the field 375) the field 376) the field 377) the field 378) the field 379) the field 380) the field 381) the field 382) the field 383) the field 384) the field 385) the field 386) the field 387) the field 388) the field 389) the field 390) the field 391) the field 392) the field 393) the field 394) the field 395) the field 396) the field 397) the field 398) the field 399) the field 400) the field 401) the field 402) the field 403) the field 404) the field 405) the field 406) the field 407) the field 408) the field 409) the field 410) the field 411) the field 412) the field 413) the field 414) the field 415) the field 416) the field 417) the field 418) the field 419) the field 420) the field 421) the field 422) the field 423) the field 424) the field 425) the field 426) the field 427) the field 428) the field 429) the field 430) the field 431) the field 432) the field 433) the field 434) the field 435) the field 436) the field 437) the field 438) the field 439) the field 440) the field 441) the field 442) the field 443) the field 444) the field 445) the field 446) the field 447) the field 448) the field 449) the field 450) the field 451) the field 452) the field 453) the field 454) the field 455) the field 456) the field 457) the field 458) the field 459) the field 460) the field 461) the field 462) the field 463) the field 464) the field 465) the field 466) the field 467) the field 468) the field 469) the field 470) the field 471) the field 472) the field 473) the field 474) the field 475) the field 476) the field 477) the field 478) the field 479) the field 480) the field 481) the field 482) the field 483) the field 484) the field 485) the field 486) the field 487) the field 488) the field 489) the field 490) the field 491) the field 492) the field 493) the field 494) the field 495) the field 496) the field 497) the field 498) the field 499) the field 500) the field 501) the field 502) the field 503) the field 504) the field 505) the field 506) the field 507) the field 508) the field 509) the field 510) the field 511) the field 512) the field 513) the field 514) the field 515) the field 516) the field 517) the field 518) the field 519) the field 520) the field 521) the field 522) the field 523) the field 524) the field 525) the field 526) the field 527) the field 528) the field 529) the field 530) the field 531) the field 532) the field 533) the field 534) the field 535) the field 536) the field 537) the field 538) the field 539) the field 540) the field 541) the field 542) the field 543) the field 544) the field 545) the field 546) the field 547) the field 548) the field 549) the field 550) the field 551) the field 552) the field 553) the field 554) the field 555) the field 556) the field 557) the field 558) the field 559) the field 560) the field 561) the field 562) the field 563) the field 564) the field 565) the field 566) the field 567) the field 568) the field 569) the field 570) the field 571) the field 572) the field 573) the field 574) the field 575) the field 576) the field 577) the field 578) the field 579) the field 580) the field 581) the field 582) the field 583) the field 584) the field 585) the field 586) the field 587) the field 588) the field 589) the field 590) the field 591) the field 592) the field 593) the field 594) the field 595) the field 596) the field 597) the field 598) the field 599) the field 600) the field 601) the field 602) the field 603) the field 604) the field 605) the field 606) the field 607) the field 608) the field 609) the field 610) the field 611) the field 612) the field 613) the field 614) the field 615) the field 616) the field 617) the field 618) the field 619) the field 620) the field 621) the field 622) the field 623) the field 624) the field 625) the field 626) the field 627) the field 628) the field 629) the field 630) the field 631) the field 632) the field 633) the field 634) the field 635) the field 636) the field 637) the field 638) the field 639) the field 640) the field 641) the field 642) the field 643) the field 644) the field 645) the field 646) the field 647) the field 648) the field 649) the field 650) the field 651) the field 652) the field 653) the field 654) the field 655) the field 656) the field 657) the field 658) the field 659) the field 660) the field 661) the field 662) the field 663) the field 664) the field 665) the field 666) the field 667) the field 668) the field 669) the field 670) the field 671) the field 672) the field 673) the field 674) the field 675) the field 676) the field 677) the field 678) the field 679) the field 680) the field 681) the field 682) the field 683) the field 684) the field 685) the field 686) the field 687) the field 688) the field 689) the field 690) the field 691) the field 692) the field 693) the field 694) the field 695) the field 696) the field 697) the field 698) the field 699) the field 700) the field 701) the field 702) the field 703) the field 704) the field 705) the field 706) the field 707) the field 708) the field 709) the field 710) the field 711) the field 712) the field 713) the field 714) the field 715) the field 716) the field 717) the field 718) the field 719) the field 720) the field 721) the field 722) the field 723) the field 724) the field 725) the field 726) the field 727) the field 728) the field 729) the field 730) the field 731) the field 732) the field 733) the field 734) the field 735) the field 736) the field 737) the field 738) the field 739) the field 740) the field 741) the field 742) the field 743) the field 744) the field 745) the field 746) the field 747) the field 748) the field 749) the field 750) the field 751) the field 752) the field 753) the field 754) the field 755) the field 756) the field 757) the field 758) the field 759) the field 760) the field 761) the field 762) the field 763) the field 764) the field 765) the field 766) the field 767) the field 768) the field 769) the field 770) the field 771) the field 772) the field 773) the field 774) the field 775) the field 776) the field 777) the field 778) the field 779) the field 780) the field 781) the field 782) the field 783) the field 784) the field 785) the field 786) the field 787) the field 788) the field 789) the field 790) the field 791) the field 792) the field 793) the field 794) the field 795) the field 796) the field 797) the field 798) the field 799) the field 800) the field 801) the field 802) the field 803) the field 804) the field 805) the field 806) the field 807) the field 808) the field 809) the field 810) the field 811) the field 812) the field 813) the field 814) the field 815) the field 816) the field 817) the field 818) the field 819) the field 820) the field 821) the field 822) the field 823) the field 824) the field 825) the field 826) the field 827) the field 828) the field 829) the field 830) the field 831) the field 832) the field 833) the field 834) the field 835) the field 836) the field 837) the field 838) the field 839) the field 840) the field 841) the field 842) the field 843) the field 844) the field 845) the field 846) the field 847) the field 848) the field 849) the field 850) the field 851) the field 852) the field 853) the field 854) the field 855) the field 856) the field 857) the field 858) the field 859) the field 860) the field 861) the field 862) the field 863) the field 864) the field 865) the field 866) the field 867) the field 868) the field 869) the field 870) the field 871) the field 872) the field 873) the field 874) the field 875) the field 876) the field 877) the field 878) the field 879) the field 880) the field 881) the field 882) the field 883) the field 884) the field 885) the field 886) the field 887) the field 888) the field 889) the field 890) the field 891) the field 892) the field 893) the field 894) the field 895) the field 896) the field 897) the field 898) the field 899) the field 900) the field 901) the field 902) the field 903) the field 904) the field 905) the field 906) the field 907) the field 908) the field 909) the field 910) the field 911) the field 912) the field 913) the field 914) the field 915) the field 916) the field 917) the field 918) the field 919) the field 920) the field 921) the field 922) the field 923) the field 924) the field 925) the field 926) the field 927) the field 928) the field 929) the field 930) the field 931) the field 932) the field 933) the field 934) the field 935) the field 936) the field 937) the field 938) the field 939) the field 940) the field 941) the field 942) the field 943) the field 944) the field 945) the field 946) the field 947) the field 948) the field 949) the field 950) the field 951) the field 952) the field 953) the field 954) the field 955) the field 956) the field 957) the field 958) the field 959) the field 960) the field 961) the field 962) the field 963) the field 964) the field 965) the field 966) the field 967) the field 968) the field 969) the field 970) the field 971) the field 972) the field 973) the field 974) the field 975) the field 976) the field 977) the field 978) the field 979) the field 980) the field 981) the field 982) the field 983) the field 984) the field 985) the field 986) the field 987) the field 988) the field 989) the field 990) the field 991) the field 992) the field 993) the field 994) the field 995) the field 996) the field 997) the field 998) the field 999) the field 1000) the field 1001) the field 1002) the field 1003) the field 1004) the field 1005) the field 1006) the field 1007) the field 1008) the field 1009) the field 1010) the field 1011) the field 1012) the field 1013) the field 1014) the field 1015) the field 1016) the field 1017) the field 1018) the field 1019) the field 1020) the field 1021) the field 1022) the field 1023) the field 1024) the field 1025) the field 1026) the field 1027) the field 1028) the field 1029) the field 1030) the field 1031) the field 1032) the field 1033) the field 1034) the field 1035) the field 1036) the field 1037) the field 1038) the field 1039) the field 1040) the field 1041) the field 1042) the field 1043) the field 1044) the field 1045) the field 1046) the field 1047) the field 1048) the field 1049) the field 1050) the field 1051) the field 1052) the field 1053) the field 1054) the field 1055) the field 1056) the field 1057) the field 1058) the field 1059) the field 1060) the field 1061) the field 1062) the field 1063) the field 1064) the field 1065) the field 1066) the field 1067) the field 1068) the field 1069) the field 1070) the field 1071) the field 1072) the field 1073) the field 1074) the field 1075) the field 1076) the field 1077) the field 1078) the field 1079) the field 1080) the field 1081) the field 1082) the field 1083) the field 1084) the field 1085) the field 1086) the field 1087) the field 1088) the field 1089) the field 1090) the field 1091) the field 1092) the field 1093) the field 1094) the field 1095) the field 1096) the field 1097) the field 1098) the field 1099) the field 1100) the field 1101) the field 1102) the field 1103) the field 1104) the field 1105) the field 1106) the field 1107) the field 1108) the field 1109) the field 1110) the field 1111) the field 1112) the field 1113) the field 1114) the field 1115) the field 1116) the field 1117) the field 1118) the field 1119) the field 1120) the field 1121) the field 1122) the field 1123) the field 1124) the field 1125) the field 1126) the field 1127) the field 1128) the field 1129) the field 1130) the field 1131) the field 1132) the field 1133) the field 1134) the field 1135) the field 1136) the field 1137) the field 1138) the field 1139) the field 1140) the field 1141) the field 1142) the field 1143) the field 1144) the field 1145) the field 1146) the field 1147) the field 1148) the field 1149) the field 1150) the field 1151) the field 1152) the field 1153) the field 1154) the field 1155) the field 1156) the field 1157) the field 1158) the field 1159) the field 1160) the field 1161) the field 1162) the field 1163) the field 1164) the field 1165) the field 1166) the field 1167) the field 1168) the field 1169) the field 1170) the field 1171) the field 1172) the field 1173) the field 1174) the field 1175) the field 1176) the field 1177) the field 1178) the field 1179) the field 1180) the field 1181) the field 1182) the field 1183) the field 1184) the field 1185) the field 1186) the field 1187) the field 1188) the field 1189) the field 1190) the field 1191) the field 1192) the field 1193) the field 1194) the field 1195) the field 1196) the field 1197) the field 1198) the field 1199) the field 1200) the field 1201) the field 1202) the field 1203) the field 1204) the field 1205) the field 1206) the field 1207) the field 1208) the field 1209) the field 1210) the field 1211) the field 1212) the field 1213) the field 1214) the field 1215) the field 1216) the field 1217) the field 1218) the field 1219) the field 1220) the field 1221) the field 1222) the field 1223) the field 1224) the field 1225) the field 1226) the field 1227) the field 1228) the field 1229) the field 1230) the field 1231) the field 1232) the field 1233) the field 1234) the field 1235) the field 1236) the field 1237) the field 1238) the field 1239) the field 1240) the field 1241) the field 1242) the field 1243) the field 1244) the field 1245) the field 1246) the field 1247) the field 1248) the field 1249) the field 1250) the field 1251) the field 1252) the field 1253) the field 1254) the field 1255) the field 1256) the field 1257) the field 1258) the field 1259) the field 1260) the field 1261) the field 1262) the field 1263) the field 1264) the field 1265) the field 1266) the field 1267) the field 1268) the field 1269) the field 1270) the field 1271) the field 1272) the field 1273) the field 1274) the field 1275) the field 1276) the field 1277) the field 1278) the field 1279) the field 1280) the field 1281) the field 1282) the field 1283) the field 1284) the field 1285) the field 1286) the field 1287) the field 1288) the field 1289) the field 1290) the field 1291) the field 1292) the field 1293) the field 1294) the field 1295) the field 1296) the field 1297) the field 1298) the field 1299) the field 1300) the field 1301) the field 1302) the field 1303) the field 1304) the field 1305) the field 1306) the field 1307) the field 1308) the field 1309) the field 1310) the field 1311) the field 1312) the field 1313) the field 1314) the field 1315) the field 1316) the field 1317) the field 1318) the field 1319) the field 1320) the field 1321)

FIGURE 1: A DBASE III PLUS PROGRAM TO CREATE MARC AMC FORMAT  
DATA ENTRY FORMS USING SIMPLE DATABASE DEPENDENCY

<pre> @ 0, 0 SAY "CONTR" @ 0, 12 GET ONE-&gt;CONTR @ 1, 0 SAY "DATETIME" @ 1, 12 GET ONE-&gt;DATETIME @ 2, 0 SAY "CATEGMAT" @ 2, 12 GET ONE-&gt;CATEGMAT @ 3, 0 SAY "DATAELEM" @ 3, 12 GET ONE-&gt;DATAELEM @ 4, 0 SAY "LCNUMBER" @ 4, 12 GET ONE-&gt;LCNUMBER @ 5, 0 SAY "SYSNUMBER" @ 5, 12 GET ONE-&gt;SYSNUMBER @ 6, 0 SAY "BIBLEVEL" @ 6, 12 GET ONE-&gt;BIBLEVEL @ 7, 0 SAY "CATSOURCE" @ 7, 12 GET ONE-&gt;CATSOURCE @ 8, 0 SAY "LANGCODE" @ 8, 12 GET ONE-&gt;LANGCODE @ 9, 0 SAY "AREACODE" @ 9, 12 GET ONE-&gt;AREACODE @ 10, 0 SAY "CHRONCODE" @ 10, 12 GET ONE-&gt;CHRONCODE @ 11, 0 SAY "GEOCODE" @ 11, 12 GET ONE-&gt;GEOCODE @ 12, 0 SAY "SUBCATCODE" @ 12, 12 GET ONE-&gt;SUBCATCODE @ 13, 0 SAY "LOCALNUMB" @ 13, 12 GET ONE-&gt;LOCALNUMB @ 14, 0 SAY "PERSNAME" @ 14, 12 GET ONE-&gt;PERSNAME @ 16, 0 SAY "CORPNAME" @ 16, 12 GET ONE-&gt;CORPNAME @ 18, 0 SAY "CONFNAME" @ 18, 12 GET ONE-&gt;CONFNAME @ 20, 0 SAY "TITLEHEAD" @ 20, 12 GET ONE-&gt;TITLEHEAD @ 22, 0 SAY "UNITITLE" @ 22, 12 GET ONE-&gt;UNITITLE @ 24, 0 SAY "TRANSTITLE" @ 24, 12 GET ONE-&gt;TRANSTITLE READ @ 1, 0 SAY "COLUNITITL" @ 1, 12 GET ONE-&gt;COLUNITITL @ 3, 0 SAY "TITLESTATE" @ 3, 12 GET ONE-&gt;TITLESTATE @ 5, 0 SAY "PUBLDESCR" @ 5, 12 GET ONE-&gt;PUBLDESCR @ 7, 0 SAY "PHYSDESCRP" @ 7, 12 GET ONE-&gt;PHYSDESCRP @ 9, 0 SAY "MEDIUM" [CONTINUED IN NEXT COLUMN] </pre>	<pre> [CONTINUED FROM PREVIOUS COLUMN] @ 9, 12 GET ONE-&gt;MEDIUM @ 10, 0 SAY "ORGANIZTN" @ 10, 12 GET ONE-&gt;ORGANIZTN @ 13, 0 SAY "NOTES" @ 13, 12 GET TWO-&gt;NOTES @ 14, 0 SAY "SUBADDPN" @ 14, 12 GET TWO-&gt;SUBADDPN @ 18, 0 SAY "SUBADDCN" @ 18, 12 GET TWO-&gt;SUBADDCN @ 21, 0 SAY "SUBADDCF" @ 21, 12 GET TWO-&gt;SUBADDCF @ 24, 0 SAY "SUBADDUT" @ 24, 12 GET TWO-&gt;SUBADDUT READ @ 2, 0 SAY "SUBADDTH" @ 2, 12 GET TWO-&gt;SUBADDTH @ 6, 0 SAY "SUBADDGN" @ 6, 12 GET TWO-&gt;SUBADDGN @ 9, 0 SAY "GENRE" @ 9, 12 GET TWO-&gt;GENRE @ 10, 0 SAY "OCCUPTN" @ 10, 12 GET TWO-&gt;OCCUPTN @ 13, 0 SAY "FUNCTION" @ 13, 12 GET TWO-&gt;FUNCTION @ 14, 0 SAY "LOCSUB" @ 14, 12 GET TWO-&gt;LOCSUB @ 16, 0 SAY "ADDPN" @ 16, 12 GET TWO-&gt;ADDPN @ 18, 0 SAY "ADDCP" @ 18, 12 GET TWO-&gt;ADDCP @ 20, 0 SAY "ADDCF" @ 20, 12 GET TWO-&gt;ADDCF @ 22, 0 SAY "ADDUT" @ 22, 12 GET TWO-&gt;ADDUT @ 24, 0 SAY "ADDTTD" @ 24, 12 GET TWO-&gt;ADDTTD READ @ 1, 0 SAY "ADDPUB" @ 1, 12 GET TWO-&gt;ADDPUB @ 3, 0 SAY "PHYSACCS" @ 3, 12 GET TWO-&gt;PHYSACCS @ 4, 0 SAY "HOSTITEM" @ 4, 12 GET TWO-&gt;HOSTITEM @ 6, 0 SAY "LOCATION" @ 6, 12 GET TWO-&gt;LOCATION @ 8, 0 SAY "ALTGRAPH" @ 8, 12 GET TWO-&gt;ALTGRAPH @ 9, 0 SAY "FORGNMARC" @ 9, 12 GET TWO-&gt;FORGNMARC </pre>
--	---

FIGURE 2: A MODEL OF MULTIPLE DATABASE DEPENDENCY OF MARC AMC SUBJECT FIELDS

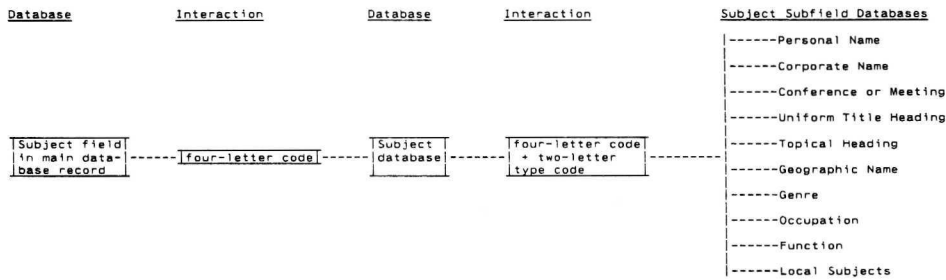


FIGURE 3: SAMPLE SESSION FROM THE EMMA GOLDMAN PAPERS MENU-DRIVEN PROGRAM

