

A "software-free" reprint file management system under RT-11

C. ROY GENTER II

State University of New York, Buffalo, New York 14226

The maintenance of reprint files is a formidable and unpleasant task. A number of ingenious mechanical or index-card systems have been described (Broadhurst, 1962, 1963; Hornik, 1963), but they cannot match the potential power and flexibility of a system implemented on a digital computer. Unfortunately, programming a file management system in a familiar higher level computer language (e.g., FORTRAN or BASIC) is a difficult task because of the extensive string manipulations required. In addition, strict formatting may be necessary, and core storage limitations may restrict file size (see, e.g., Deni, 1977).

A powerful reprint filing and retrieval system may be implemented on a Digital Equipment Corporation (DEC) PDP-11 (or LSI-11) using only the capabilities of the RT-11 system editor.¹ The system described here allows searches and cross-referencing to be carried out by key word, by words in the article's title, by author, or by any code or acquisition number assigned to the paper for filing purposes. The system will run with no restrictions on a baseline (8K-word) machine. Only a rudimentary knowledge of the RT-11 editor is required to input the source files (DEC, 1977, Chapter 5).

Source File Format. Articles may be entered into the source file in almost any format using the RT-11 editor.² There is only one restriction: Each completed entry must be followed by a control L (ctrl L). This inserts a form feed, which the editor interprets as a page break. Thus, each article entry occupies a separate "page" in the reprint source file.

A suggested format for entering articles is shown in Figure 1. The first line contains the paper's acquisition number, and the following lines contain author, title, publication, and key-word information, respectively. As many lines of each category as are necessary may be included, and any category may be omitted. For legibility of the output files, it is a good idea to place at least one carriage return after the key words and before the ctrl L.³

The only point that requires special comment is the period that precedes the acquisition number and each key word. This period allows a search for these categories without confusing them with numbers in the publication information or words in the title. For example, if a search is carried out for the key word "COLOR" with the references in Figure 1, only the first reference will be found. A search by the word "COLOR" (without the preceding period) retrieves both references. Without the simple expedient of insert-

```
.1417  
GENTER,II,C.R., GERI,G.A., & KANDEL,G.L.  
LUMINANCE-DEPENDENT SHIFTS IN UNIQUE  
YELLOW LOCI  
INVEST. OPHTHAL. & VIS. SCI. 17:198-199 1978  
.COLOR .UNIQUE YELLOW .LUMINANCE  
.OPPONENT PROCESSES
```

```
(ctrl L)  
.1418  
GERI,G.A., GENTER,II,C.R., & KANDEL,G.L.  
STILES-CRAWFORD EFFECT AT THE UNIQUE  
YELLOW POINT  
INVEST. OPHTHAL. & VIS. SCI. 17:198 1978  
.STILES-CRAWFORD .UNIQUE YELLOW  
  
(ctrl L)
```

Figure 1. The format used to enter articles into the reprint source file. (See text for a complete description.)

ing these periods, a rigid formatting procedure would be required, along with more extensive software development.

Search Procedure. After a reprint source file has been created, the following procedure searches that file for a specified character string (e.g., an author, key word, etc.). The "P" command reads a page of text (a single reprint entry) and looks for the specified search string. If no match is found, the current page is deleted and a new page is read. When a match is found, the "N" command writes the entire current page to the output file, and the sequence is repeated.

```
R EDIT  
*ERREPRNT.REP$R$EWSSEARCH.REP$$  
*10000<P.COLOR$N$>$SEX$$
```

In this example, the reprint source file to be searched is named REPRNT.REP, and the results of the search are placed in a file named SEARCH.REP. The located search string is the key word ".COLOR".

Users with BATCH capabilities may simplify the search procedure by creating the following two files.

```
SEARCH.COM  
R BATCH  
SEARCH.BAT  
  
SEARCH.BAT  
$JOB/RT11  
(tab) LET E=33  
(tab) LET A=74  
(tab) LET B=76  
*ERREPRNT.REP'E'R'E'EWSEARCH.REP'E'E'  
*10000'A'P'ENTER SEARCH STRING'CTY'E'  
*N'E'B'E'E'  
*EX'E'E'  
$EOJ
```

These two files need be entered only once. When the

indirect command file SEARCH.COM is executed (by typing @SEARCH), a prompt will appear, asking for the search string. Enter the search string and press return. The results of the search will be placed in the file SEARCH.REP.

Cross-Referencing. Cross-referencing is carried out simply by searching the output file of a previous search with a new search string. The following example cross-references the file REPRNT.REP by the key words ".COLOR" and ".STILES-CRAWFORD."

```
R EDIT
*ERREPRNT.REP$R$EWSEARCH.REP$$
*10000<P.COLOR$N$>$$
*EF$ESEARCH.REP$R$$
*10000<P.STILES-CRAWFORD$N$>$EX$$
```

This example outputs two files. SEARCH.BAK includes all references with the key word ".COLOR," and SEARCH.REP includes only those references with the key words ".COLOR" and ".STILES-CRAWFORD" in common.

BATCH users may simplify cross-referencing by the creation of the following two files.

```
XREF.COM
R BATCH
XREF.BAT
XREF.BAT
$JOB/RT11
(tab) LET E=33
(tab) LET A=74
(tab) LET B=76
R EDIT
*ERSEARCH.REP'E'R'E'EWSEARCH.REP'E'E'
*10000'A'P'ENTER CROSS-REF STRING'"CTY'E'
*N'E'B'E'E'
*EX'E'E'
$EOJ
```

To run, enter @XREF. A prompt will appear for the string to be cross-referenced with the results of a previous search in the file SEARCH.REP. The cross-referenced file replaces the old file, SEARCH.REP. XREF may be run any number of times to cross-reference by multiple search strings.

Run Statistics. This system runs on any PDP-11 supporting RT-11. A 1,000-reference source file occupies

approximately 330 disk blocks and can be searched in approximately 60 sec on a LSI-11 system with floppy-disk drives. Better performance can be expected with a faster processor and/or hard disk mass storage.

Availability. More specific information covering the system may be obtained from C. Roy Genter II, Department of Psychology, State University of New York, Buffalo, New York 14226.

REFERENCES

- BROADHURST, P. L. Coordinate indexing: A bibliographic aid. *American Psychologist*, 1962, 17, 137-142.
- BROADHURST, P. L. Information retrieval. *American Psychologist*, 1963, 18, 664-665.
- DENI, R. BASIC-PLUS programs to manage bibliographies, reference lists, and reprint files. *Behavior Research Methods & Instrumentation*, 1977, 9, 307-308.
- DIGITAL EQUIPMENT CORPORATION. *Introduction to RT-11*. Maynard, Mass: August 1977.
- HORNIK, G. J. Information retrieval. *American Psychologist*, 1963, 18, 260-261.

NOTES

1. Although this paper describes the implementation details of the reprint filing system for the RT-11 Version 3 operating system, many other operating systems for different processors have commands equivalent to those discussed.

2. When adding references to the source file, it is wise to create short temporary files and to concatenate them to the source file, thus protecting the source against machine crashes and operator error. There are two simple ways to do this. The first method uses a monitor command:

```
COPY/CONCATENATE REPRNT.REP,TEMP.REP
REPRNT.REP
```

Unfortunately, this sequence will not execute on a baseline machine. The following sequence will:

```
R EDIT
*ERREPRNT.REP$R$EWREPRNT.REP$FIND SNARK$$
*ERTEMP.REP$R$FIND SNARK$$
*EX$$
```

This method should not be used on files containing snarks.

3. If two ctrl Ls are inserted between each entry, then new entries may be printed, cut into 3 x 5 in. cards, and filed alphabetically by the first author.

(Accepted for publication January 24, 1980.)