Lib sc. 5; 1968; PAPER Q.

CATALOGUE-ON-TAPE.

(Non-conventional methods in document retrieval. 5).

S VENKATARAMAN*, *International Computers Ltd, N R Square, Bangalore* 2. and A NEELAMEGHAN, *Documentation Research and Training Centre, Bangalore* 3.

[Describes the steps in the preparation of the Catalogue-on-Tape forming the input to the computer-based document finding system. The format of the entry for a document, preparation of the punched card, sorting, and transfer of the entries on to magnetic tape to form the store, are explained. The step-by-step procedure and flow-chart in the transfer of the entries on to magnetic tape, and amending this list of documents, are given.]

0 Scope of the Paper

This paper describes the preparation and amendment of a Catalogue-on-Tape. This work forms Phase 1 in setting up the experimental computer-based document finding system.

1 Input Documents

11 SUBJECTS

In the experiment, the subjects of the documents forming the input to the system were mainly those going with the Basic Subject 'Fountain Pen Production'. There were also a few documents dealing with related subjects, such as those going with the Host Subjects Glass Technology, Plastics Technology, and Metallurgy.

12 SCHEME FOR CLASSIFICATION

The participants in the DRTC Course 1967–68 developed a depth version of CC for the classification of subjects going with the Basic Subject 'Fountain Pen Production'. This work was done in the class as part of a demonstration and practice of the

---

* Visiting Lecturer, DRTC, 1968.

design and development of a depth schedule, using the methodology developed in DRTC (2, 4).

13 CATALOGUING

The rules in the Classified Catalogue Code, Ed 5 (=CCC) for making a Main Entry for an indexing periodical (3) and the Indian Standard for Bibliographical References (1) were used in cataloguing the documents. The Class Number (=CN) and the corresponding Feature Headings were written as recommended for the examples given along with the provisional depth schedules published in the *Library science with a slant to documentation* and in the volume of papers of the DRTC Seminars. A specimen entry is given below:

MP85,9R1–9B3.8,03; a7:fR,51 Fountain Pen, Gold Nib, Platinum tipped, Nib, Tip, Homogeneity, Test, Microscopy. GEORGE (S) and McGILVER (J D). Microscopic examination of nib tips. (Canad stat. 11; 1956; 25–40).

2 Punched Card

21 FORMAT

The input entries for the Catalogue-on-Tape (=C—O—T) were prepared by punching each of the entries on a standard 80-column card. In deciding a format for punching, the following factors were taken into consideration:

1 The length of each document — that is, the number of characters in each entry;
2 Convenience of sorting the cards; and
3 The different digits used in the (CN) and other elements in the entry for the document.

22 CONTINUOUS PUNCHING

The number of characters used in the entry varied with each document. Therefore, assigning a fixed number of cards for each document resulted in a large number of cards in the catalogue. This would increase the operation time of the computer. Therefore, it was decided to punch continuously the characters, in each entry using continuation cards, wherever necessary.

221 *Division into Fields*

An entry essentially contained the following sections written as a single paragraph:

1 Class Number;
2 Feature Heading; and
3 Bibliographical description of the Host Document as shown in the example in Sec 13.

The sections were separated from each other into fields by a virgule (/). The digit is not used in the CC notational system. This division of an entry into fields was helpful in the Document Finding Phase of the job (*See* Paper R in this issue).

Annotation.— In a subsequent experiment in which the machine synthesised the Class Number on giving the Kernel Terms, it was found helpful to separate the different sections of the entry into paragraphs.

23 MODIFICATION OF ENTRY FORMAT

A few minor modifications of the entry format given in Sec 13 was adopted to save on the number of characters. The modified format of the example in Sec 13 is shown below:

MP85,9R1-9B3-8.03;A7:FR,51/ FOUNTAIN PEN. GOLD NIB. PLATINUM TIPPED, NIB. TIP, HOMOGENEITY, TEST. MICROSCOPY./GEORGE S. MCGILVER J D. MICROSCOPIC EXAMINATION OF NIB TIPS. (CANAD STAT. 11;56;25-40).

In the above example,

1 The conjunction "and" between the names of the authors is replaced by a 'Comma';

2 The Secondary Elements in the Name of the Author are not enclosed in circular brackets;

3 Only the last two digits of the 'year' are given in the Locus Section as all the documents in the collection are those published later than 1900; and

4 All the terms are written in all Capitals.

Any other detail, such as an indication of the language of the article, can be added to the entry.

3 Sorting

31 DOCUMENT NUMBER AND CONTINUATION CARD NUMBER

The characters in an entry were punched continuously in columns 1 to 76. Columns 77 and 78 were used for punching an alpha-numeric Document Number. Columns 79 and 80 were used for punching a Serial Number in the continuation cards for an entry. This document identification code was also helpful in the punching operation being assigned to different operators. For example, the Document Number 'A' may be assigned to an operator. He will number all the documents punched by him as A1, A2 ... AB, AC, etc. If A1 results in 3 continuation cards, they will be numbered 01, 02, 03 in columns 79 and 80, in the first, second, and third cards respectively of the set.

32 METHOD

After the cards are punched, they may be sorted either on a conventional Sorter or by the computer when the number of cards is large. In the present case a Sorter was used. The sorting was done using the Document Number and Serial Number, giving a set of cards arranged according to the Document Number and within a set of continuation cards by their Serial Number.

A Catalogue-on-Tape was created from these punched cards. A flow chart and a program were written for this purpose. (*See* Sec 5A to 5Z).

4 Storage Location Assignment

41 FOR ENTRY AS A WHOLE

For convenience, symbolic names were allotted to locations in the computer where the data were stored. In the experiment, the maximum number of characters in any one entry for a document forming the input to the system, did not exceed 360. Each computer word can hold 4 characters. Therefore, 101 words were reserved for storing each entry. This location was called BUF in the program. The successive storage locations were each named as BUF, BUF+1, BUF+2 and so on to BUF+100.

42 FOR SECTIONS OF AN ENTRY

The contents of columns 1 to 76 in a card were stored in a location called CARD A to CARD A+18. Similarly, the contents of columns 1 to 76 of a second card were stored in a location called CARD B to CARD B+18.

The contents of columns 77 and 78 in a card were stored in a location called DOC A. Similarly, the contents of columns 77 and 78 of a second card were stored in a location called DOC B.

The contents of columns 79 and 80 were stored in a location called SER A. Similarly, the contents of columns 79 and 80 of a second card were stored in a location called SER B.

An Accession Number (⇐ACN) assigned to each entry by the machine, was stored in BUF+100.

43 LAST CARD

The last card of a set of cards was identified by placing immediately after it a card in which an asterisk was punched in each of the first four columns.

5 Stage 1: Preparation of Master Catalogue-on-Tape

51 FLOW-CHART (*See* Fig 1 in Appendix 1)

The operations described in Sec 5A to 5Z are those performed by the machine. The letters m and n represent word position, and the letters M and N the character position. For convenience, the steps are indicated as Step C1.1, C1.2 etc.

5A SETTING THE VALUE OF ACCESSION NUMBER (C1.1)
The value of ACN was made equal to 0 (Zero).

5B MODIFYING WORD ADDRESS (C1.2)
The value of m and n were made equal to 0 (Zero) and 18 respectively.

5C FILLING UP WITH SPACE (C1.3)
The locations BUF to BUF + 99 were filled up with space.
Annotation.— This ensured that only the data read from the cards would be stored in this location, leaving the rest of it blank —that is, no printable character. In most of the computers, all the locations are filled with zeros in the beginning. A zero is a printable character. Therefore, it is necessary to start with space in BUF.

5D TRANSFER OF THE CONTENTS OF A CARD (C1.4)
The machine read a card. The contents of the columns 1 to 76 were transferred to the location CARD A (m) to CARD A(n); those of the columns 77 and 78 to the location DOC A; and those of the columns 79 and 80 to the location SER A.

5E CHECKING FOR LAST CARD (C1.5)
Examined whether the card just read in was the last card. If it was not, the program proceeded to Step C1.6. If it was the last card, it branched to Step C1.23.

5F TRANSFER OF CARD A TO BUF (C1.6)
The contents of CARD A was stored in BUF + m to BUF + n.

5G TRANSFER OF CONTENTS OF ANOTHER CARD (C1.7)
The machine read another card. The contents of columns 1 to 76 were transferred to the location CARD B(m) to CARD B(n); those of the columns 77 and 78 to the location DOC B; and those of the columns 79 and 80 to the location SER B.

5H CHECKING FOR LAST CARD (C1.8)
Examined whether the card just read in was the last card. If it was not, the program proceeded to Step C1.9. If it was the last card, it branched to Step C1.23.

5J CHECKING FOR CONTINUATION CARD (C1.9)
Compared the data in the location DOC A (= Contents of the columns 77 and 78 of the first card) with those in the location DOC B (= Contents of the columns of 77 and 78 of the second

card). If they were the same, then the second card would be a continuation of the first card. The program then branched to Step C1.17. If they were not the same, then the second card would be the first card of the entry for another document. In this case, the program proceeded to Step C1.10.

5K INCREASING THE VALUE OF THE ACCESSION NUMBER (C1.10)
The value of ACN was increased by 1.

5L TRANSFER OF ACCESSION NUMBER TO BUF (C1.11)
The ACN was stored in BUF + 100.

Annotation.— The original value of ACN was 0 (Zero) (*See* Step C1.1). The numbering of the entries on the tape should start from 1. Therefore, the value of ACN was increased by 1 and stored in BUF + 100 before transfer of the entry on to the tape.

5M TRANSFER OF BUF TO TAPE (C1.12)
The contents of BUF to BUF + 100 were transferred on to a magnetic tape.

5N CHECKING FOR PRINT-OUT (C1.13)
Examined whether switch 23, indicating the need for a print-out of the data transferred on the tape, was on. If it was on, the program proceeded to Step C1.14. If it was not on, it branched to Step C1.15.

5P PRINT-OUT OF BUF (C1.14)
The contents of BUF were printed out for use as a checklist.

5Q MODIFYING WORD ADDRESS (C1.15)
The value of m and n were made equal to 0 (Zero) and 18 respectively.

5R FILLING UP WITH SPACE (C1.16)
The locations BUF to BUF + 99 were filled up with space. The program then branched to Step C1.19.

5S CHECKING FOR SERIAL NUMBER (C1.17)
If at the Step C1.9 the contents of DOC A were equal to DOC B, then the second card would be a continuation of the first card. Therefore, the Serial Number of the second card would be higher than that of the first card by 1. This was checked up by adding 1 to SER A and comparing it with SER B. If SER A + 1 was equal to SER B, the program proceeded to Step C1.18. If they were not equal, the program branched to Step C1.22.

5T MODIFYING WORD ADDRESS (C1.18)
The value of m and n were each increased by 19.

5U TRANSFER OF CARD TO CARD A (C1.19)
The contents of the location CARD B were transferred to the location CARD A.

5V TRANSFER OF DOC B TO DOC A (C1.20)
The contents of the location DOC B were transferred to the location DOC A.

5W TRANSFER OF SER B TO SER A (C1.21)
The contents of the location SER B were transferred to location SER A.
The program then branched to Step C1.6 and the operations were repeated.

5X ERROR CONDITION (C1.22)
If at Step C1.17, SER A + 1 was not equal to SER B, it indicated an error condition. The machine typed out a statement such as "Error" on the console typewriter.

5Y REPETITION OF STEPS C1.10 TO C1.14 (C1.23 TO C1.27)
If either at Step C1.5 or Step C1.8 the last-card condition was sensed, the program repeated the Steps C1.10 to C1.14. In the Flow-chart these Steps are indicated as steps C1.23 to C1.27.
Annotation.— The last-card condition would not arise in the first iteration while creating the Catalogue-on-Tape.

5Z COMPLETION OF JOB (C1.28)
The program instructed the machine to re-wind the tape and stop the operations.

6 Stage 2: Amending the Catalogue-on-Tape

61 NEED FOR AMENDMENT
The need for amending a Catalogue-on-Tape may arise for one or more of the following reasons:
1 To delete an entry when the document which it represents is to be withdrawn from use;
2 To correct some detail in an entry, such as the (CN), the Feature Heading, and the Specification of Host Document; and
3 When a new entry is to be added to the catalogue.

62 SEQUENCE OF OPERATIONS
The deletions, if any, should be carried out first. The addition of new entries should be made following the last record.

The deletion and/or addition of entries will give rise to a new amended Catalogue-on-Tape. The sequence of the entries may change. But it does not matter in the selection stage, because a full-file search will be made. However, by programming, it is also possible to rearrange the entries in a desired sequence on another tape.

63 MASTER TAPE

The original Catalogue-on-Tape or C-O-T is called the Master Tape. The first amended tape would be the first generation tape (C-O-T.1). The Master Tape is recommended to be kept until at least a second generation tape (C-O-T.2) is created. After that, if the original is no more required, it may be over-written, that is, re-used.

64 DELETION CARD

The Accession Number for the entry to be deleted is punched on a card. Several such Accession Numbers may be punched on one and the same card. However, they should be fed into the machine in the same sequence as the occurrence of the corresponding entries in C-O-T.0.

Annotation.— 1 It is assumed that a print-out of C-O-T.0 will be available, so that the sequence of the entries is known.

2 The last-card (with an asterisk punched in the first four columns) is placed immediately as the last card in the set of deletion cards.

3 In the example described here, the Accession Number of the record to be deleted was punched in columns 1 to 8, giving one card for each entry to be deleted.

65 FLOW-CHART (*See* Fig 2 in Appendix 2)

The steps in the program are indicated in the following Flow-chart. For convenience, the steps are indicated as Step C2.1, C2.2 etc. These steps are described in Sec 6A to 6M.

6A READING-IN A DELETION CARD (C2.1)

A deletion card was read.

6B STORING THE ACCESSION NUMBER (C2.2)

The ACN was stored in a location called AMEND.

6C CHECKING FOR THE LAST CARD (C2.3)

Checked for the last card. If it was the last card, the program branched to Entry Point 2. If it was not the last card, it proceeded to Step C2.4.

6D READING-IN ENTRY FROM C–O–T.0 (C2.4)
An entry from the C–O–T.0 was transferred to BUF to BUF + 100.

6E CHECKING FOR END SENTINEL (C2.5)
Checked for the End Sentinel. If it was sensed, it indicated an error condition, if there was no continuation reel. It implied that the deletion cards were not in proper sequence. If no End Sentinel was sensed, the program proceeded to Step C2.6.

Annotation.— In all the programs, the magnetic tape house keeping routines of ICL were used. Whenever a 'Rewind Tape' instruction was given while creating an output tape, the program would write a special block of information as the last block. When the tape is used as an input tape, the End Sentinel would indicate to the machine whether it was the end of the file or whether a continuation file should be opened.

6F COMPARISON OF BUF AND AMEND (C2.6)
Compared whether BUF + 100 was equal to AMEND. If it was, the program branched to Step C2.8. If it was not, it proceeded to Step C2.7.

6G TRANSFER TO TAPE (C2.7)
The entry read in from C–O–T.0 was copied on to C–O–T.1. The program then re-entered at Step C2.4.

6H PRINT-OUT (C2.8)
If, at Step C2.6, BUF + 100 was equal to AMEND, a print-out of the BUF + 100 was made. (*See* Fig 3 in Appendix 3)

6J ENTRY POINT 2 (C2.9)
When a last-card condition was sensed at Step C2.3, it indicated that there were no more deletion cards to be read in. Therefore, the remaining entries on C–O–T.0 were copied on to C–O–T.1 till the End Sentinel was sensed.
At the beginning of this operation, Step C2.4 was repeated.

6K CHECKING FOR END SENTINEL (C2.10)
Checked for the End Sentinel. If it was sensed, the program proceeded to Step C.12. If it was not sensed, it proceeded to Step C2.11.

6L TRANSFER TO TAPE (C2.11)
The entry read in from C-O-T.0 was copied on to C-O-T.1. The program then re-entered at Step C2.9.

6M TRANSFER OF BUF (C2.12)

When the End Sentinel was sensed at Step C2.10. the contents of BUF + 100 were assigned to ACN. The program re-entered at Step C1.2.

Annotation.— The Entry Point 2 can also be used if there were only entries to be added — that is, no deletions to be made. In that case, the entries on C-O-T.0 would be copied on to C-O-T.1 till the End Sentinel is sensed.

7 Time Taken

The computer ICL 1903 used could read 300 cards per minute. The printing speed was 300 lines per minute. The card reading speed dropped to an average of 250 cards per minute when the checklist of the Catalogue-on-Tape was also printed. The average number of cards per document in the experiment was 3. Therefore, the approximate time to transfer the entries for 10,000 documents on to the tape and print the Catalogue-on-Tape would be 120 minutes.

8 Bibliographical References

1 Sec 15 IS:2381-1963. Recommendations for bibliographical reference. 1963. Indian Standards Institution.

2 Sec 12 NEELAMEGHAN (A), GOPINATH (M A), and DENTON (P H). Motor vehicle production: Depth classification: A demonstration. (Lib sc. 4; 1967; Paper H).

3 Sec 15 RANGANATHAN (S R). Classified catalogue code. Ed 5. 1964. Chap TD.

4 Sec 12 ——. Design of depth classification: Methodology. (Lib sc. 1; 1964; Paper A).