

39
A)

A Note-card System with Full-text Databases

Kwok-Yam Tam, Zy-Kaan Ding, Ching-Chun Hsieh

Computing Center, Academia Sinica

Nankang, Taipei, R.O.C.

I. Introduction

Full text databases can be applied in many aspects, for example, in the studies of laws and regulations, in the studies of historical documents and in query and auditing of monopoly articles [1] and etc. Queries are facilitated by query languages to retrieve the required information. Satisfactory query languages are available in most full text databases such as STAIRS (*), STATUS [2], BRS and etc. The retrieved text can be processed as interactive display on the screen or off-line output from print-out devices.

Another approach of application after text retrieval is proposed here; that is a note-card system to associate with full text databases. In the system, extraction of the retrieved text is available to make note-cards. Simple editing functions are provided as post-processing of the note-card, for example, to rearrange the extracted text or to add comments. Functions of

up-loading and down-loading of the note-card files to floppy diskettes are also provided to facilitate system storage and for keeping one's own card information. The same set of editing functions is available in the query process.

The reason for designing this note-card system is that the relevant information obtained through full text queries may be transferred to personal document processors for further processing or re-organization without bothering to reproduce the text.

II. Functions of Note-card System

1. Creation of Note-cards

This function produces a note-card whose contents are extracted by text tailoring provided in the full text database, with the extracted text, notes on the origin of the extracted text and on the path to access the text in the logical compilation of the full text database [3], the date of creation, the series number, classification and key word fields. In this way, effort and errors in input are reduced. The contents of the note-card can be changed through the editing function provided.

The classification and the key word fields of the electronic note-card system are equivalent to the holes for query in the physical note-card system. These two sets of fields require user-defined data.

2. Query of Note-cards

Queries can be made through the classification fields or the QBE (query by example) approach is available. After query, functions of duplication, deletion, print-out of the relevant note-cards are provided.

3. Management of Note-cards

This function is designed to attain a multi-user management environment in the note-card system. It facilitates users to keep their note-cards and to transfer and share data for reference. It has the following sub-functions.

i) Up-loading and Down-loading

The note-card files can be up-loaded from and down-loaded to floppy diskettes for storage and book-keeping. The files down-loaded are ASCII files. They can readily be ported to some PC softwares for example, LOTUS 1-2-3 for further processing.

ii) User Privilege Alteration

This is to manipulate the capabilities granted to duplicate note-card information and to protect the listing of one's directory information for privacy.

iii) Deletion of Note-cards

Users can delete their own note-cards.

iv) Directory Listing

This is to survey the current states of the note-card system which include the status and the number of note-cards of users.

v) Duplication of Note-cards

Users can duplicate their note-cards to specified users for reference.

4. Print-out of Note-cards

Options of picking a specified note-card to print and of dumping the whole note-card file are provided.

III. System Design

1. Prerequisite

The note-card system is developed under the environment of UNIX (**) System V in the AT&T 3B series computer. The programming language applied is C, involving a software library called 'curses'. The workstation control software is BX10 (**).
[4]

2. Interface with Full Text Database

The technique of shared memory is applied to pass information between the full text database and the note-card system for note-card creation. The format of the shared memory is as shown in fig.1. There are different ways of note-card creation according to different needs. The contents of the note-card include the origin and the logical path of the extracted text, the text itself and etc.

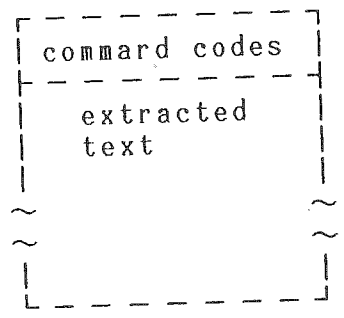


Fig.1

3. Data Structure

The interaction between the contents file and the classification file of the note-cards is depicted as in fig. 2. The classification field is actually an index in the linked list of classification to the card number in the contents file.

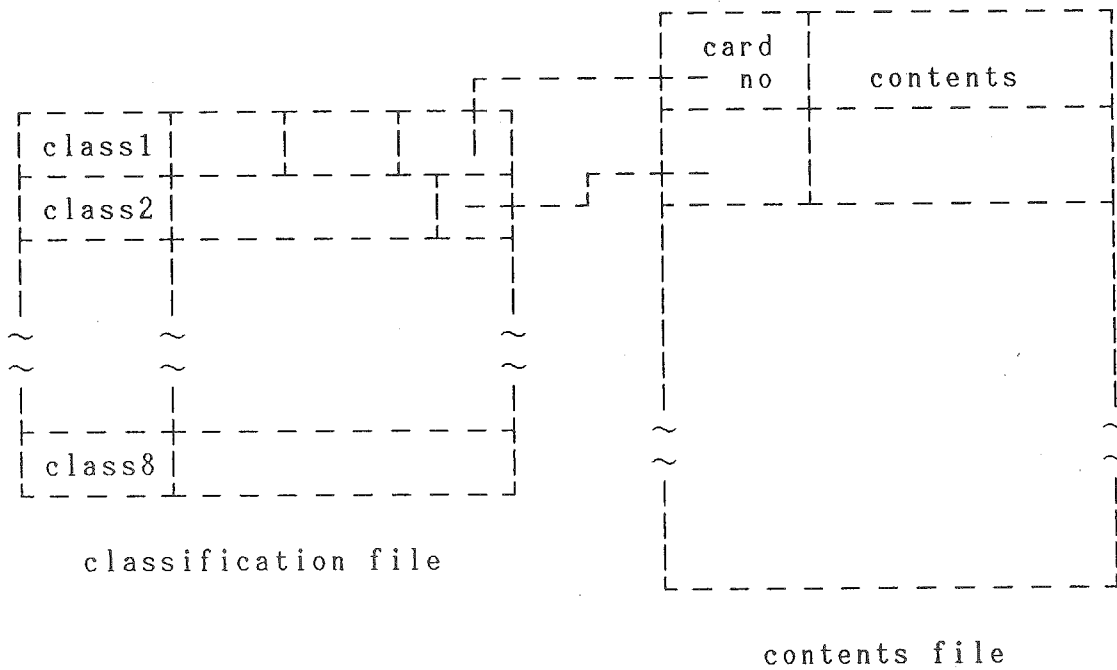


Fig.2

Different fields are distinguished by delimiters. The field of contents in the contents file is of variable length.

4. Program Structure

The program structure of the note-card system is depicted as in fig.3. Modules interact through flags setting in the approach of control coupling [5].

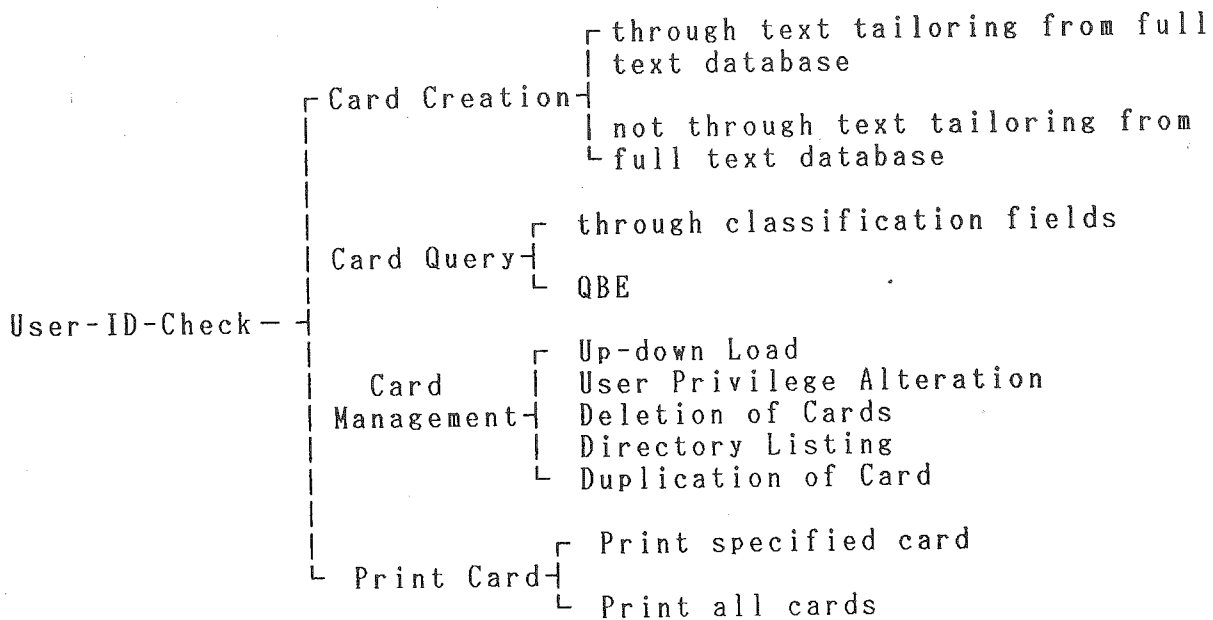


Fig.3

5. File Compaction

Segments exist in the note-card files after changes such as amendment, deletion and duplication of the note-cards. Consistency between the contents file and the classification file is assured after file compaction. To do file compaction is optional.

IV. Conclusion

The note-card system possesses the following characteristics :

1. The privacy of individual set of note-cards is preserved.
2. It is an independent processor.
3. It is a common utility that can associate with any full text database system.

In this system, a different approach is applied to make use of the full text database. That is to extract retrieved text by text tailoring then make it into an electronic note-card for further research.

The future development includes the following.

1. Apply full text search in the field of contents.
2. Apply multiple options and definitions in the classification field.
3. Involve both graphic and textual data so as to integrate into hypertext databases with multi-media functions.
4. Apply a user-directed menu driven approach in the human-machine interface.

V. References

- [1] L.A. Hollaar, "Specialized Merge Processor Networks for Combining Sorted Lists", ACM Tr. on DB System 3,3(Sept.); pp272-284.
- [2] Computer Power, "STATUS ----- Information Storage and Retrieval System", 1982.
- [3] C.C. Hsieh et al, "Full-text Database for Chinese History Document", ICCPCOL'88, Toronto, Canada.
- [4] AT&T, "UNIX System User's Manual System V", Bell Lab, 1983.
- [5] R.S. Pressman, "Software Engineering A Practitioner's Approach" McGraw-Hill, 1982.

- (*) STAIRS is registered mark of IBM.
- (**) UNIX is registered mark of AT&T.
- (***) BX10 is registered mark of Acer.

----- end -----