

General-Purpose File-Based Information System

Ivan Moustakerov

Institute of Information Technologies, 1113 Sofia

1. Introduction

It is frequently needed to collect and manipulate large volumes of information in current everyday practice. There are developed many technologies and tools for that and each has its specific advantages and disadvantages [1, 2, 3, 4]. The most frequently used for information retrieval systems is data base technology. It is developed and improved for years and satisfies all currently known practical needs. There exist many commercial implementations of that technology with proven quality (IBM's dBase, Oracle, etc.). The increasing of the computational power of the modern personal computers and the existence of the global computer networks (Internet) made possible database using in many lives areas. The advantages of database technology are quite known and will not be discussed here. It is interesting to discuss some of their essential disadvantages instead, because these are important for developing and exploitation of the information systems. On the first place it should be stated that databases are very complex software systems and that fact affects all stages of developing and exploitation [2, 4]. It needs specifically trained and highly qualified personnel. The process of developing is long and it is not unusual to get to the need of organizational information structure changes in order to use the full potential of the database capabilities. The exploitation and maintenance of the database also needs specialists and need of training exists even for the ordinary users. The database complexity and huge volumes of stored information lead to different practical risks - exploitation malfunctioning, security risk, etc. Last but not least in consideration are the price of database software itself and the costs for developing of the information system with that database. The developing stage is mandatory and critical for the end success of information system implementing. The question is if it is always necessary to use database technology for information systems? There are real life applications where functional requirements are only part of those of a database and where the described database disadvantages are unacceptable. There exist other information systems technologies [2] and the choice

is dependent of the specific application needs. For example the scientific workers need to search in a small number of journals on specific scientific area. It is also important to have the possibility to adjust the information system to the needs of each specific user. So, it would be very valuable to have a "real" personal information system - to be installed on his personal computer without any hardware or software limitations and to be used without any complex training. Similar requirements can be successfully met using file-based technology [2]. The main disadvantages of that technology are difficult achieving of similar to databases functional flexibility, existence of specific to each application file organization, possibility for data duplication and difficulties for data actualizing. Compared to that are its advantages as easiness of using, very good adjustment to the user requirements, exact functional correspondence and low cost.

Following these considerations a file-based information system general-purpose technology was developed and tested as a demo version of library search system.

2. Main functional requirements to the system

The first evident main requirement is storing of structured information. It is satisfied indirectly on the development stage by choosing of file types, file organization and their content [2]. Aiming to the universality of the application two types index-files, one type master-file and one type transition-file were chosen. The specific characteristic of the chosen file types is their simple logical organization and the fact they are text files and could easily be created by untrained users. That characteristic is important for users with requirements for the confidentiality as police service for example, where even software developers should not have access to the files content. Another important characteristic is chosen file structuring leading to small volume files, the number of which is increasing with the information volume increasing. That gives easiness of maintaining, actualizing and adjusting to work environment and user requirements as well.

List of the needed basic functions is as follows:

- Function SEARCH - it is the most important function for any information system. In the first demo version most practically needed search requirements are included - searching by keyword, by part of keyword, sequential "AND" combination of keywords, keyword with unknown symbols, replaced with "?". All other variations of searching will be included when specific requirement exists. All found records are shown on the screen with scrolling. There is special function for looking on the results.

- Function LOOK - it shows screen by screen the results of search function executing. The results are listed in an easy to understand format. All results are stored in temporary transition files that are available for looking or researching until the exiting of the information system. After exiting these files are lost so there exist a special function for permanent saving of the results.

- Function SAVE - it saves the results found after searching in a file with user defined name. The check for existing file with same name is performed and that name can be changed. The saved file is of "ASCII" or "TEXT" format and can be processed with any text editor if needed. The file format is similar to those shown on screen by LOOK function but also includes information for the number of the found records and the source where they were found. If needed time and date of searching can be added.

- Function PRINT - it serves documenting needs when working with system, i.e. all search results can be printed by means of the existing printer. That demo version uses OS high level printing capabilities with processing of the printer errors. If practice

needs low level OS printing functions will be used instead.

It was decided to test capabilities of the developed file-based information system technology with the above described set of basic functions but the functional expanding is open for any real life needs.

3. Software realization

The first demo version works in so called "text mode" and has minimal requirements to the working environment i.e. personal computer with any version of DOS. Borland Turbo C++ programming language was chosen. The attractiveness of the graphical user interface (so called "windows"-interface) was sacrificed for the removing of any requirements to the computer operating system and to concentrate to functional effectiveness of the system. The interface is realized as choice of the menu items by their position numbers, foolproof input data control and possibility for input data correction, hierarchical menus and menus level jumping and break at any level with direct jump to Main Menu.

The program code is compiled in "large" memory model with execution speed optimization. There is no requirement to the RAM memory amount and the external storage is used to fully exploit the developed file structure and to have unlimited information volume.

4. Algorithmic structure of the system demo version

Using "top-down" developing approach module structure of the first demo version of the system is as shown in Fig. 1. The following program modules are coded:

- Main control module (Main Menu).
- Search module (for SEARCH function).
- Module for look at the results of searching (for LOOK function).
- Module for printing of the results (for PRINT function).
- Module for saving the results in user "text" file with user chosen name (for SAVE function).
- Auxiliary modules (for integer data input with validity control and editing capabilities, symbol data input with validity control and editing capabilities, transition files deleting on exit of the system, check for existing file on results saving and other system modules).

For the demo version of the system creating and editing of all system files is done by any text editor following the developed file structure. This could be advantage because user can create all files by himself (confidentiality property) with his favorite text editor. Practical using of the system will show if some other editing modules are needed.

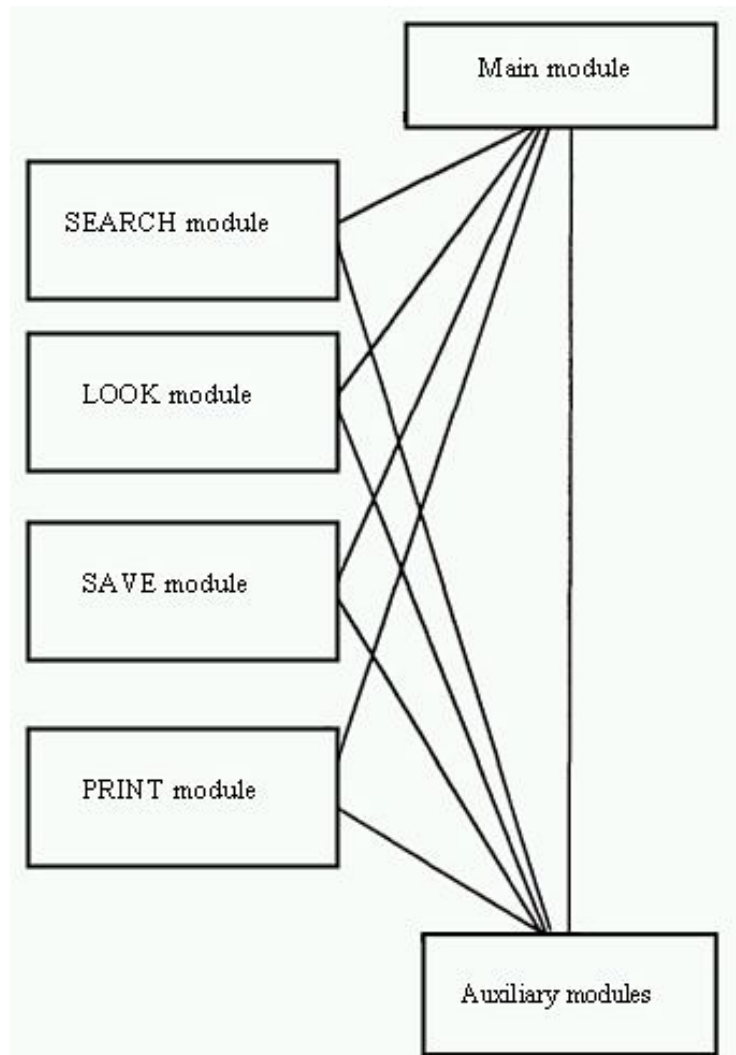


Fig. 1. Algorithmic block-diagram of the system demo version

5. Conclusions

The first demo version of the developed file-based information system has been tested and has proved its functionality. The customer requirements and the need for future expanding of the system options are investigated also. The system will be developed by adding a graphical user interface for Windows environment. Java technology will be used to make the system independent of the computer OS platform. That will expand the application field enormously (for PDA for example) and will add principally new functions as Internet connections and searching when needed.

6. References

1. Met ev, B., I. Moustakero v, V. Monov. User-oriented information system for bibliographic references. - IIT Working papers, Nov. 2000.
2. Curt Is, G., Business Information Systems - Analysis, Design and Practice. New York, Addison-Wesley Publ. Co., 1995.
3. G I o s h k o v, L. Information Technologies. Sofia, Technika, 1988.
4. O l l e, T. V., Suggestions of CODASIL for data bases control. Moscow, Finances and Statistics, 1981.

Файловая информационно-справочная система общим применением

Иван Мустакеров

Институт информационных технологии, 1113 София

(Резюме)

Обсуждается сохранением и манипулированием большого объема персонализированной текстовой информации для настольных компьютеров. Описывается технология для реализации простых в обслуживании файловые информационно-справочные системы, привязанные к требованиям конкретных потребителей. Описываются основные идеи при программной реализации демо-версией персональной информационной системы для библиографских справок, реализующая основных функций – поиск текста, показывание результатов поиска, сохранение и печат данных. Обсуждаются преимущества и недостаток файловых информационных систем и возможных областях их применения.