CARROLL, JOHN M.
--Evolution of computing as a
community resource

(2)

EVOLUTION OF COMPUTING AS A COMMUNITY RESOURCE

by

John M. (1) Carroll and F.M. Lim

Computer Science Department

University of Western Ontario

London, Canada

The research reported in this paper was performed under contract with the Canadian Department of Communications whose support is gratefully acknowledged.

ABSTRACT

The system is at once (a) a dynamic directory of human services, (b) an electronic bulletin board, and (c) a socio-political early-warning system. It permits storage and retrieval of information required by citizens and of the questions causing such retrieval. Both semi-static and variable information is stored. The semi-static information is accessed by an extremely fast and powerful on-line inquiry programme which affords a rich variety of access points. It is initially processed in batch-mode by a separate computer especially suited to such operations. The variable information exists as both annotations and detailed files. The annotations are created and stored in a file with forward and backward linkage; and automatic garbage collection and relocation. The detailed files take advantage of the text editing and search capabilities of the time-sharing system. Citizens' questions are collected in machine-sensible format; cross-tabulated; and made subject to a powerful on-line inquiry programme for detailed analysis.

INTRODUCTION

What can computers do for people?

Our first response to that challenge was that perhaps a computer could sort out the various human service provisions our society affords and put the individual needing some kind of help in contact with someone willing and able to help him.

In our city of about 250,000 people one agency, Information London, was doing that already. It consists of a group of ladies who respond to phone calls, letters or walk-in visits and call upon published material and interpersonal contracts for answers. For a year and a half now we have been trying to help them.

The university made available $20,000 worth of free computer time. We did our batch processing of files and production of directories and reports on their CDC Cyber 73/14 (ie CDC 6400) and used their PDP 10/50 (128 K of core) for on-line terminal access. The Federal Department of Communications granted $10,000 that went mostly for supplies and salaries of part-time research assistants. We contributed two terminals: a CDC-700 visual display terminal and a GE Terminet 300 hard-copy terminal - Control Data of Canada lent us another CDC-700. Terminals were linked to PDP 10 by telephone lines through acoustic couplers.

Various organizations have given us moral support and more importantly information and critical appraisal. These include:

Information London

Tourist and Convention Bureau of Greater London

London Public Library and Art Museum

London Free Press Publishing Co. and Station CFPL-TV

Coalition for Development

London Chamber of Commerce

Bell Canada

City of London, Office of the Mayor.

SCOPE OF THIS PAPER

This paper is a discussion of the problems encountered in structuring information in this kind of a community environment.

Essentially two kinds of information must be handled:

(a) That which is to be disseminated to the users.

(b) That which characterizes the users' requests.

The two categories interact dynamically. The continuing analysis of user requests identifies areas where new information must be gathered and guides the development of storage and retrieval strategies.

OVERALL INFORMATION STRUCTURE

There are three data files:

(a) The Master file contains 5,000 short records each consisting of four fields: name, address, phone number and classification code. The name field contains the full name of an organization furnishing some kind of human service, its common name, initials or acronym, if any, and an additional notation if the name is not descriptive of the organization's function. Classification is denoted by a 12-digit code which specifies its affiliation, specifies what human service it furnishes and provides unique identification.

(b) Any number of ephemeral files linked to a unique 12-digit code can be created. Such a file can contain narrative information to any desired length.

(c) A scratchpad file provides space for brief comments regarding organizations or their services. These comments are linked to the 12-digit codes designating the organizations commented upon.

THE MASTER FILE: A CATALOGUE OF HUMAN RESOURCE

We have in essence a relational file structure in which the 12-digit code functions as a common link.

The master file is created and updated in the batch mode. Only in this way can entries be added or deleted or names changed. Addresses and telephone numbers, can, however, be changed from on-line terminals.

Batch mode is also employed to produce either individual printouts for each of the 5,000 organizations or directory listing of the entire master file.

The individual printouts are mailed to the subject organizations in window envelopes when the master file is to be revalidated and updated. A letter of explanation, a pamphlet describing the 12-digit code system, and a reply envelope are sent along with the printout. Respondants are encouraged to check the accuracy of their listing and contribute additional information for the ephemeral or scratchpad files.

The directory listings can be produced in three formats.

1) Entries are listed by affiliation (municipal, provincial, federal, private, etc.), by service offered (basic income; food, clothing and shelter; health care, etc.), and alphabetically within these categories.

2) Entries are listed by service offered, affiliation, and alphabetically within these categories.

3) Entries are listed alphabetically according to a key-word-in-context permutation applied to the augmented name field of the record.

Directories in bound form are consulted by systems users not processing terminal access and provide alternative information resources where terminals exist.

ADDRESSING THE MASTER FILE

There are six modes of operation implemented at terminals: search, display, create, release, change and locate. See Figure 1.

The search and change modes pertain to the master file. The create, display, and release modes pertain to ephemeral and scratchpad files. The locate and change modes are privileged only to systems personnel.

The search mode enables the user to search the master file. The master file can be searched by alphabetic or numeric keys.

The alphabetic index is an inverted file. It consists of an entry table and 4,000 records. The entry table contains a

list of the index terms obtained by the keyword-in-context permutation of the master file name field and a pointer to a record. Each record contains a header telling how many records satisfy that key and a list of pointers to master file records in which the key term appears in the name field. See Figure 2. The alphabetic index entry table is addressed by binary search.

The numeric index is likewise an inverted file but in four parts. Each part corresponds to a permutation of the 12-digit code which brings the desired digit pair or single digit of the 12-digit classification code into the indexing slot. The first digit pair designates affiliation. The second and third digit pairs and the seventh digit designate with increasing specificity the human service offered. Each instance of a digit pair or single digit is one line in the entry table. Corresponding to each of the 12-digit codes is a pointer to the master file record possessing it. See Figure 3.

Search logic permits any desired combination of the operators: AND, OR, NOT, AND-NOT, OR-NOT with alphabetic or numeric keys ; NOT can only be the initial operator.

In searching the file the user is given a report of the number of hits scored in response to his initial logical request. He is then afforded the opportunity to enhance the precision of his request by adding more terms to it.

The master file is addressed by pointers from a list satisfying the logical search request.

The user can obtain an initial display of only the first line of each of the records so addressed: the 12-digit code and part of the name field N at a time. In the resulting display, each line is assigned an index 1 to N. The user can obtain the complete records one at a time by entering the appropriate index numbers.

Thus four techniques:1) audit of hits, 2) provision for increasing the specificity of search, 3) abbreviated display of search results, and 4) the facility to ask for only those retrieved records actually desired help prevent the user from being overwhelmed by the undesired product of imprecise search requests.

The combination of numeric and alphabetic search facilities and the provision for five logical operations constitutes a powerful search facility.

One could, for example, request a listing of all Roman Catholic churches by specifying

1st digit : 06 (an association)

2nd digit : 09 (community affairs)

3rd digit : 02 (churches)

AND Roman * OR RC.

Perhaps one would then increase the specificity of search by adding

AND-NOT Greek * OR Anglican *

The star operator permits right truncation of search terms.

The syntax in the search mode is:

<NUMERIC-KEY> AND <NUMERIC-KEY> AND

-NOT- -OR- -OR-

-AND-NOT- -AND-NOT-

-OR-NOT- -OR-NOT-

<NUMERIC-KEY> OR <ALPHA-KEY|<ALPHA-KEY>*...

-AND-

-AND-NOT-

-OR-NOT-

THE EPHEMERAL FILE: AN ELECTRONIC BULLETIN BOARD

Both the ephemeral and scratchpad files are accessed by a key table. See Figure 4. The key table is headed by a counter telling the number of keys sorted and the number of keys in an associated overflow table if any. Like the alphabetic and numeric indices, the key table is stored using the computing system's random-access file mechanism. The key table contains the 12-digit codes of organizations for which either ephemeral or scratchpad files exist.

Attached to each code is a file name if an ephemeral file exists; a password, a status indicator and a pointer if a scratchpad notation exists.

The status indicator pertains to ephemeral files. It can contain the following indicators:

1) indicates that provision for a file exists but the file has no data in it;
2) means that the file exists and has data in it;
3) means the file has been released by implementing the user release option.

The password corresponds to a password assigned to users who are privileged to create ephemeral or scratchpad entries regarding the organization designated by the 12-digit key. If a user wishes to create or release an ephemeral or scratchpad file, he can do so only if he possesses the record password corresponding to the organization he wishes to comment upon.

Should the user forget his password, systems personnel can retrieve it for him using the privileged locate mode of operation to override the security provisions of the system.

Ephemeral files are created using the computer system's serial file access mechanism and by employing macro commands in the computer system's text editor language.

For example: the local Tourist and Convention Bureau

has used their ephemeral file to store an index to 20 major points of interest. Entries in this index consist of the name of the attraction and its 12-digit code.

A user retrieves the master file record of the Tourist and Convention Bureau using the search mode and notes its 12-digit code.

He then uses the display mode to retrieve the ephemeral file pertaining to that 12-digit code.

The index provided by the Bureau then enables him to get detailed information on whatever attraction interests him by using the display mode in conjunction with the appropriate 12-digit code.

Ephemeral files can be lengthy. The ones describing tourist attractions will contain descriptions, hours of accommodations, admission fees, and directions to it. These files are paged so that the user can look just at each page heading and either have it displayed or skip to the next heading.

Success of the system as a community information medium depends upon each privileged user keeping his own ephemeral files up to date. In addition to the tourist bureau, the public library and art museum has ephemeral files pertaining to current activities at its several branches; the local newspaper has ephemeral files dealing with forthcoming events of a social

or civic nature; Information London maintains an ephemeral file specifying the scope of each local child-care centre; and a coalition of local churches keeps up files dealing with social services extended by its members.

The very name of the ephemeral files connotes that the information they contain is subject to change over time in contrast to information in the master file which can be presumed to remain invariant between file updates. Even here, however, systems personnel utilizing the privileged change option can enter from terminals information regarding changes in address and telephone number.

The intent of the ephemeral files is, therefore, to realize a dynamic electronic bulletin board for the community at large.

THE SCRATCHPAD FILE: ELECTRONIC GRAFFITTI

The scratchpad file, on the other hand, provides a convenient vehicle for entering brief comments regarding an organization offering some human service.

Scratchpad notations are contained within a common file. See Figure 4. The pointer stored in the key table tells the location of the header of the variable length scratchpad record. The header contains a flag telling whether the location is free

or not, a count of the number of lines following at this location, and a back pointer to the key table. The flag is reset when the release option is exercised. In this way the system performs dynamic garbage collection.

If, when locating a scratchpad record, it is determined that not enough space exists at the current location to contain the information entered, the system searchesfor another location that is free and can hold it, transfers the record to the new location, resets the flag in the old location, uses the back pointer to return to the key table, and updates the pointer stored there so that it now points to the new location. The scratch-pad can therefore be described as a double linked list. The programme counts the lines required by the new entry after it is typed by the user. The relocation operationis transparent to the user.

One use of the scratchpad is to store information collected by the local Chamber of Commerce regarding the number of consumer complaints received about an agency or a business over a period of time and the number of these complaints re-solved to the consumer's satisfaction.

ANALYSIS OF USER REQUESTS

The principal user of this system has been Information London. This is a voluntary agency that responds to citizen

inquiries regarding the local availability of various human services.

Information London uses the search and display facilities of the system as well as other information in responding to citizen queries. It handles about 1,500 queries a month.

Each query is categorized according to 24 characteristics which include the age, location, and economic status of the inquirer, the service required and the success or lack of success various agencies have had in satisfying his request.

The personnel at Information London record this information on mark-sense cards. Each month these cards are processed by computer. The characteristics of the queries are cross tabulated to produce a monthly report to the board of directors of Information London, many of whom are active in various other human service agencies.

THE ON-LINE INQUIRER: THEY ARE JUDGED BY THEIR DEEDS

In addition, such month's activity is added to a data-bank that can be addressed from terminals by an on-line inquiry programme.

The syntax of this inquiry programme is such that the

user can define a class of relations. The same programme is used to produce cross tabulated summaries and to answer on-line inquiries. The syntax is shown below. The ID option is used for on-line inquiry; the CR option is used for cross tabulation.

```
 GET   ID   (relation):(relation-attribute-OR-relation-attribute)
       CR                                  -AND

WITH   (relation-GT-attribute-OR-relation-GT-attribute)
                -GE-          -AND-         -GE-
                -LT-                        -LT-
                -LE-                        -LE-
                -EQ-                        -EQ-
                -NE-                        -NE-

 -AND-  (relation ....)
 -OR-
```

This provides a powerful logical tool which enables the user to construct dynamically a multidimensional index to human service agencies based upon actual operating experience.

Our generalized cross-tabulation and on-line inquirer programmes have had additional applications.

We helped the Action League for Physically Handicapped Adults (ALPHA) analyze the results of a survey they conducted of 900 handicapped persons in our area. One result of their work was to have signs placed on public washrooms to indicate whether or not the doorway would accomodate a wheelchair. Alpha is now campaigning to have lower steps on public busses, ramped entrances to public buildings, and a few seven-place taxis made available to help handicapped persons get around. The statistics

we helped them develop will play a significant role in this work. We found for example, that 193 handicapped adults felt they could become self-supporting if such aids to mobility were available. This would represent a net annual saving to the province of half a million dollars in disability payments.

We helped a local artist; an associate of the Royal Ontario Museum, classify some 2,000 examples of native rock paintings. As a result of his studies, he was able to propose a tentative hypothesis regarding the age of the paintings, linking them with migrations of Canada's native peoples in the 17th, 18th and 19th centuries. These findings have served to stimulate interest in young native people in their cultural heritage.

We have developed abbreviated programme packages for tabulation, cross-tabulation, and on-line inquiry that will be made available to citizens' groups who need to analyze the results of surveys or other data sets in preparing presentations on various subjects.

RESULTS

We have in this way developed a system that utilizes time-shared computing resources to marshall a large amount of information freely available to the citizen faced with some

important problems. He is able to learn:

Who can help him and where are they located? (Master File)

What is their current hours of operation and conditions of service? (Ephemeral File)

Are they reliable? (Scratchpad)

Have they been able to help others with similar problems? (On-line Inquirer).

PLANS FOR FUTURE DEVELOPMENT

Future development of this system depends upon

(a) educating institutional users to make effective use of the dynamic file creation and updating facilities available and

(b) making the search and display facilities more accessible to individual users regardless of their level of sophistication.

The first problem is a training task familiar to operators of remotely accessed computer information systems but exacerbated by the fact that this system is not restricted to information that can be contained within any easily delimited data set and the operational milieu inasmuch as the institutional users are not bound in any formal contractual scheme and

frequently must themselves rely on untrained volunteer help.

Finding a solution to the second problem is basic to any plan aimed at making computing resources beneficial to citizens at large. One can visualize addressing the data bank initially by prestructured inquiry trees responsive to graphical input, or by text analysis of free-form inquiries. However, its ultimate solution will probably entail radical reconfiguration of terminal input/output mechanisms.

TOWARD A COMMUNITY INFORMATION SYSTEM

The principal accomplishments of this effort have been to devise computer programmes capable of handling variegated and volatile kinds of data with sufficient flexibility such that useful output can be obtained even though the eventual format may not always have been evident in the planning stage and to open channels for acquisition and dissemination of information of value to the individual citizen.

It is already evident that questions frequently go quickly beyond the competence of local agencies and that to be totally effective any community information system must interface effectively with the provincial and federal governments. The Provincial Departments of Consumer and Corporate Affairs and of Community and Social Services on one hand; Information Canada,

on the other, as well as other federal instrumentalities and departments will have to provide input.

Furthermore, it is necessary but not at all sufficient merely to open channels of communication. These channels must be kept open. The system is only as good as the quality and timeliness of the information fed into it. Part-time and volunteer help, however well intentioned or highly motivated, cannot maintain a continuing and reliable flow of information.

We have found that the more information we put into the system, the more users we attract. And the more users we attract the more information is needed.

We can see where we should set out to capture statistical information currently published annually in the municipal guide book.

We should also capture city council decisions as they are taken. Perhaps at some future time, we can capture information directly from council minutes.

Another useful input would be an inventory of available housing together with a collation of landlord tenent complaints and their resolutions.

Furthermore, names of agencies and businesses are insufficient, we need to add information on ownership, management,

and employment.

This research was sponsored in part by the Federal Department of Communications.

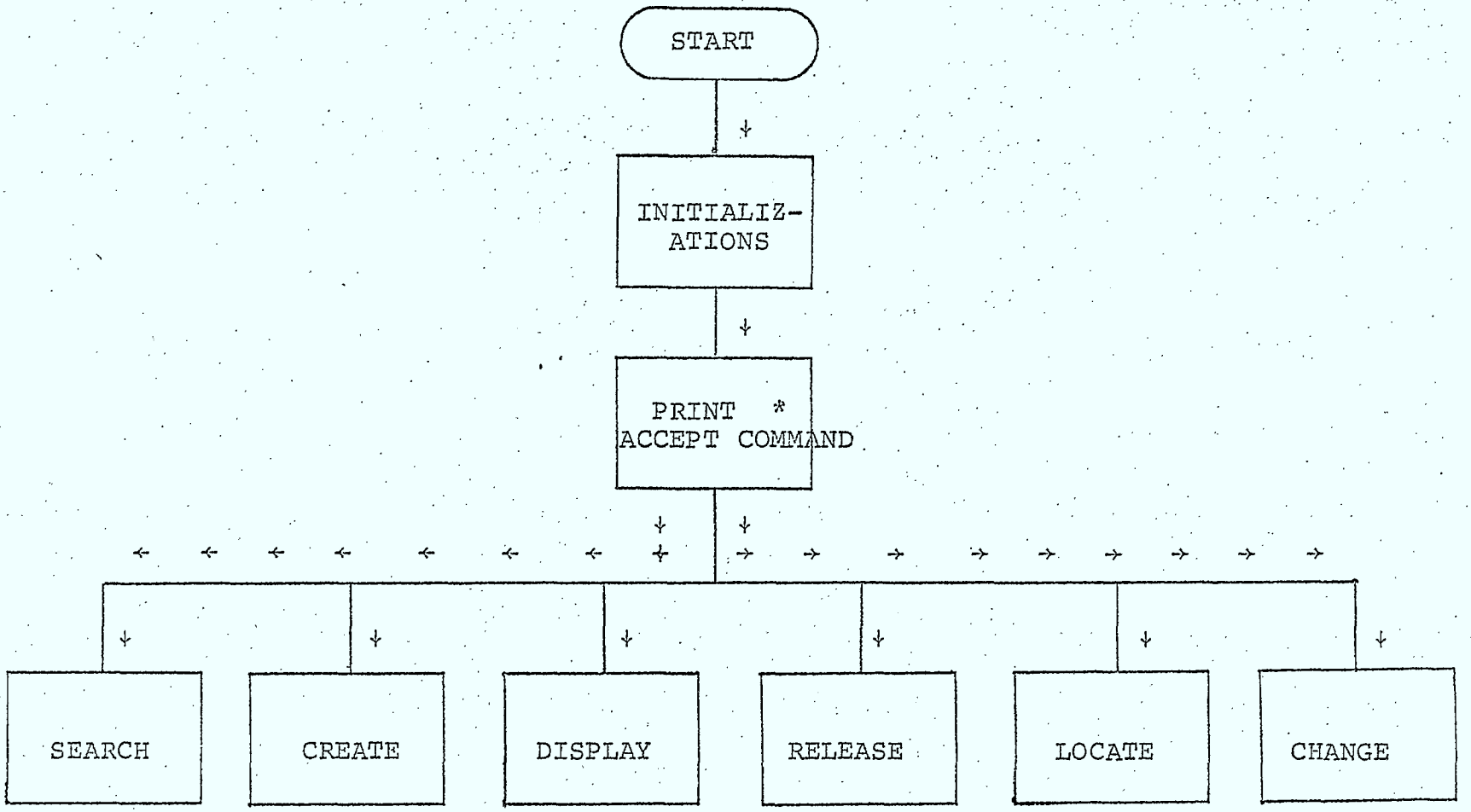
LEGENDS

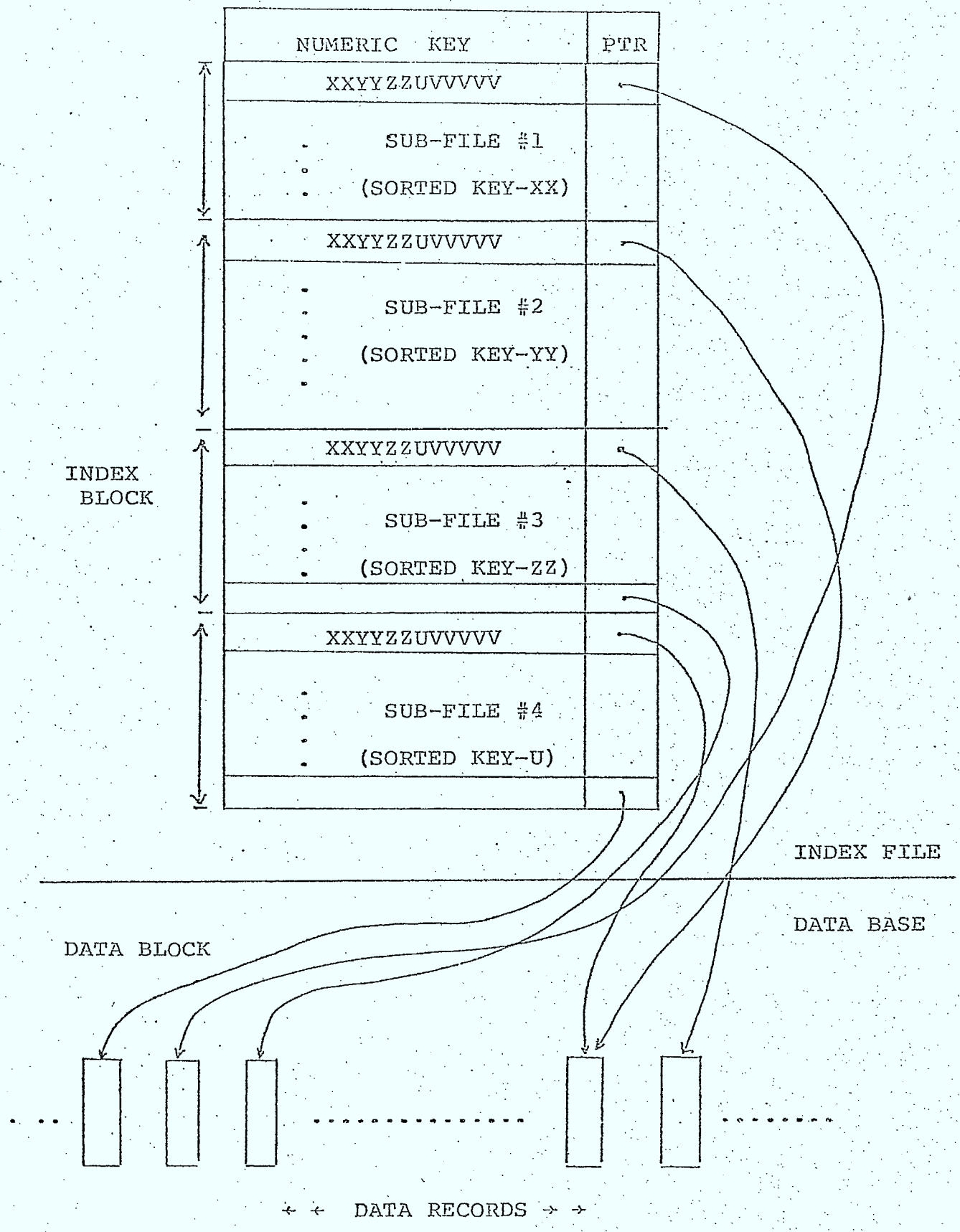
Figure 1. Modes of programming system operation.

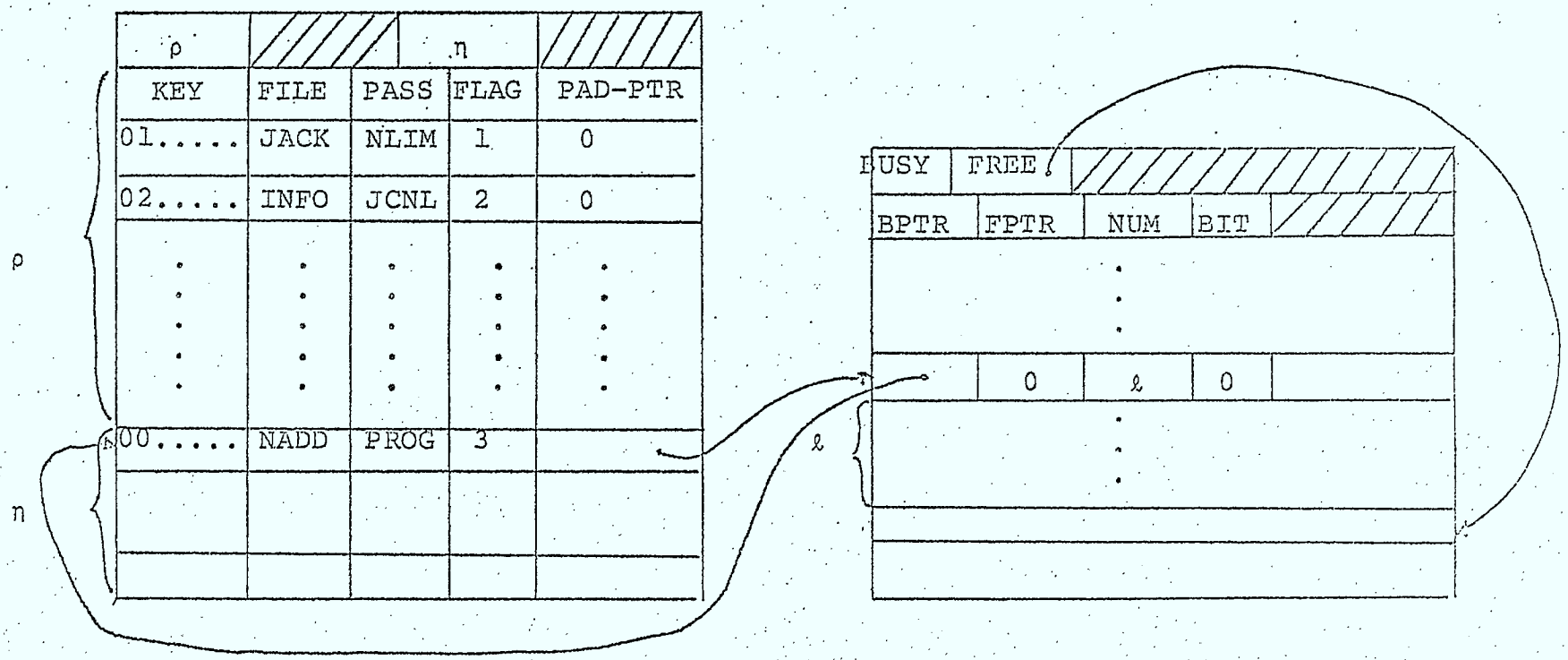
Figure 2. Structure of the alphabetical index.

Figure 3. Structure of the numeric index.

Figure 4. Structure of the index to the ephemeral and scratchpad files.

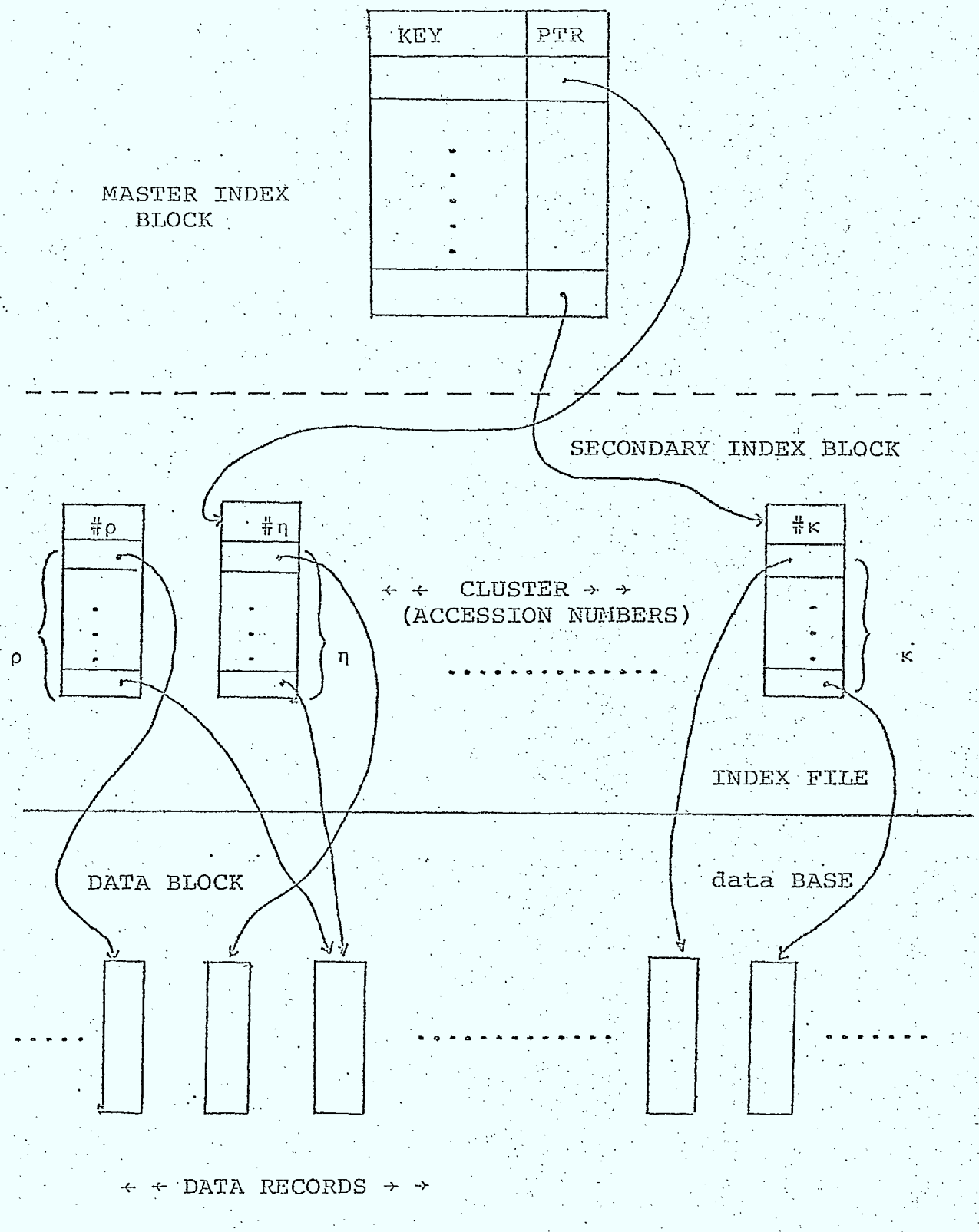

START
INITIALIZ-
ATIONS
PRINT *
ACCEPT COMMAND
SEARCH
CREATE
DISPLAY
RELEASE
LOCATE
CHANGE


NUMERIC KEY
PTR
XXYYZZUVVVVV
SUB-FILE #1
(SORTED KEY-XX)
XXYYZZUVVVVV
SUB-FILE #2
(SORTED KEY-YY)
XXYYZZUVVVVV
SUB-FILE #3
(SORTED KEY-ZZ)
XXYYZZUVVVVV
SUB-FILE #4
(SORTED KEY-U)
INDEX BLOCK
INDEX FILE
DATA BASE
DATA BLOCK
DATA RECORDS



ρ := NUMBER OF ENTRIES SORTED

η := NUMBER OF ENTRIES UN-SORTED


KEY
PTR
MASTER INDEX BLOCK
SECONDARY INDEX BLOCK
#ρ
#η
#κ
ρ
η
κ
← ← CLUSTER → →
(ACCESSION NUMBERS)
INDEX FILE
DATA BLOCK
data BASE
← ← DATA RECORDS → →