Interactive Information Processing, Retrieval, and Transfer

J.C.R. Licklider
Massachusetts Institute of Technology
Cambridge, Massachusetts, USA

ACKNOWLEDGEMENT

This paper was prepared under the aegis of Project MAC, supported by the Advanced Research Projects Agency through the Office of Naval Research under Contract Nonr-4102 (01). In addition to some of the work of Project MAC, the paper describes some of the work of Project TIP and Project INTREX. Project TIP is supported by the National Science Foundation under Research Grant GN-589. Project INTREX is supported in part by a grant from the Carnegie Corporation, in part by Contract NSF-C472 from the National Science Foundation, and in part by the Council on Library Resources, Inc.

INTERACTIVE INFORMATION PROCESSING

By far the greatest part of the experience in interacting directly with computers comes not from experiments in information retrieval or transfer but from use of computers in preparing computer programs and in solving scientific and engineering problems. The earliest digital computers were programmed on line, and there has always been a bit of on-line-programming and "debugging" (elimination of program errors), but not until the advent of multi-access computing, based on the technique of "time sharing", was it possible for large numbers of people to work as a matter of course, day after day, at computer consoles. Now there are several hundred experienced console users at M.I.T. and many times that number elsewhere.

The most widely used interactive computer systems are two almost identical systems using IBM 7094 computers and a supervisory program called the "Compatible Time-Sharing System". Connected through a telephone switchboard to the 7094 computers are about 200 consoles. Most of the consoles are merely typewriters, I shall assume that your console is a typewriter. What you do first depends upon the kind of work you have before you. The most trivial thing to do --- and therefore a good introductory task --- is to write a memorandum. The computer will make it easy for you to correct typing mistakes and to effect editorial corrections, and then type out a "clean copy" of your memo.

You call the writing and editing program typset by typing "typset", and then you type the name you want to give the file you are going to create --- for example, "jsmith".

The computer then types "W1720.2" (which means "Wait a moment." and "It is now 2/10 of a second past 17:20 o'clock.") and then "R000.2+000.1" (which means "Ready" and "You have thus far used 2/10 second of processor time and 1/10 second of drum-core transfer time.") and then "Input" (which means "I am ready for input from you.") You start to type the memorandum:

```
To:   J. R. Smitt#h
      .space
      Subjeck:
      @Subject:  Plans for Improvement of planning Strategy
      .space
        Next Tuesday is the last day for ###to submit your ideas
```

The character (#) ordinarily means to erase a character. The character (@) ordinarily means to erase a line. The control word ".space" means to skip a line. Your memorandum therefor stands as:

```
To:   J. R. Smith

      Subject:  Plans for Improvement of planning Strategy

        Next Tuesday is the last day to submit your ideas
```

You notice that you forgot to give the date and that you need to capitalize the "p" in "planning". To go into "edit mode", you press the carriage return key twice. The computer thereupon types "Edit". You are supposed, of course, to know the control words (or characters), only a few of which appear in this example. You type: "i Date: June 15, 1968" (the "i" meaning "insert"), press the carriage-return key, and then, for format control, type ".space". Then, realizing that you want to center the date, you type "t" (for "go to the top of the file") and ".center" (for "center the next line"). In order to capitalize the "p" in "planning", you type "l plan" (for "locate the character string 'plan'") and "c /plan/Plan" (which changes "planning" to "Planning") and press the carriage-return key twice to go back to the input mode and complete your memo.

And when you have completed the typing --- and have corrected all your mistakes --- you file the memo by typing "file jsmith". The computer types "W1735.1" and then "R011.3 + 008.1" You type "runoff jsmith". The computer waits for you to put a fresh sheet of paper into the typewriter. You press the carriage-return key. The computer types the memo perfectly at 15 characters per second.

Such are the mechanics. They are very convenient if you want to prepare a long paper and don't type well.

Programs

One of the main aims of Project MAC has been to see what kind of a community would grow up around multi-access computers with rich and growing software resources. The system now has over a million words of public programs and over 6 million words of private programs. Let me illustrate with a few examples: A Mathematical Assistance Program makes transformations, and solves equations for you. It handles algebra, trigonometry, differential equations, Fourier and Laplace transforms, and so on. It plots graphs. You do not have to know all about it to use it: it asks you questions until it "understands" your problem.

Another program solves even quite complex problems in symbolic integration.

ADMINS facilitates the preparation, maintenance, and use of data bases.

TEACH teaches computer programming.

OPS is a large system of programs for interactive, incremental simulation and modelling. It provides a language in which you can define objects or entities and specify their properties and the relations. It lets you set into mo ion the situation thus described and make its behaviour unfold. It records the history and prepares summaries. It lets you intervene at any time to modify anything you like.

The On-Line Community

The foregoing paragraphs described a few of the hundreds of programs of the accumulated software resource. By using their programs, each user of CTSS can take advantage in his own work of pertinent efforts of his predecessors and his colleagues. An accumulative process is beginning to operate also in the domain of data.

Through computer-facilitated human interactions, a new kind of research community is arising at M.I.T. It is of course only in an early, formative stage, but there is little doubt that something significant is happening.

As the local on-line community has emerged from concept into actuality, the idea of a broader, geographically distributed community has taken form in the minds of several people. This idea involves inter-connecting several multi-access computer systems and combining their communities of users into a super-community. When geographically distributed computers and information networks come into being, their impact upon the process of information transfer may be great.

Some Conclusions Based on Project MAC's Experience

During its five years of operation, Project MAC has explored more fields than I can summarize, but let me, nevertheless, attempt to state some conclusions pertinent to information storage and retrieval:

(1) The computer turns information into a dynamic, living thing.

(2) Everything one does in an active informational environment is "complexity limited".

(3) Man-computer interaction is the most hopeful approach to the mastery of informational complexity.

(4) Even with the help of on-line interaction, it will take cooperative on-line teamwork to achieve significant solutions to the "big" problems of science, technology, industries, cities, nations, and alliances.

(5) The basic thing in the user's concept of an interactive information system is the "name space" of the filing (i.e., memory or storage) system.

(6) Although the term "time sharing" has achieved wide currency, the sharing of processor time is not fundamentally important. Much more important are memory sharing and communication. The aim of multi-access design should not be to make each user think he has a computer all to himself; it should be to immerse each user in a cooperative, interactive, computer-based community.

(7) The importance of controlled access to files can hardly be overstated.

(8) The importance of fast interaction makes itself felt when a problem gets complex.

(9) In a community with many interests, the "general-purposeness" of the general-purpose multi-access computer system has real meaning and significance. The system must lend itself to a great variety of applications and serve as ready host to diverse subsystems. Generality and open-endedness cost something, of course, but Project MAC's experience indicates they are well worth it.

(10) Reliable operation is vital and --- since the reliability will not be perfect --- effective "back-up" arrangements and recovery procedures are vital, also. Before they will invest their main intellectual capital in, or entrust it to, a multi-access computer system, people have to be confident that the system will be available when they want it and that it will not lose their valuable programs and data.

Interactive Information Retrieval

Project TIP uses the facilities of the 7094-CTSS multi-access systems. The main TIP data base is a growing collection of bibliographic data, presently from almost 100,000 journal papers in the field of physics. The TIP programs are programs for processing the data in ways formulated by the user during his interaction with the system. Using TIP, you type in lower case and the computer types back to you in all capitals. To explain what some of TIP's abbreviations mean, I shall insert comments in parentheses.

tip (You type "tip" to evoke the TIP program.)
W1019.5 (Wait. It is 10:19 and a half.)
TYPE YOUR REQUESTS.
search annals of physics v.26 to v.28
find title pion not author boyling j.b.(Find all articles with titles containing "pion" except those by J.B. Boyling, whose work you already know.)
output print title a i and l (One-letter abbreviations are adequate for the TIP words such as "author", "identification", and "location". You could just as well have typed " o p t a i l" to instruct TIP to type as output the specified information about the items found.)
go (Go to work, TIP.)

ANNALS OF PHYSICS
VOLUME 26
VOLUME 27
J384 V027 P0079
DEUTERON PHOTODISINTEGRATION AND N-P
CAPTURE BELOW PION PRODUCTION THRESHOLD
PARTOVI F.
CAMBRIDGE, MASSACHUSETTS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LABORATORY FOR NUCLEAR SCIENCE AND
PHYSICS DEPARTMENT

VOLUME 28
J384 V028 P0034
ANALYSIS OF THE PHOTOPRODUCTION OF POSITIVE PIONS
HOHLER G.
SCHMIDT W.
GERMANY
TECHNISCHE HOCHSCHULE KARLSRUHE
INSTITUT THEORETISCHE KERNPHYSIK

(No article meeting the specification was found in Volume 26. One was found in volume 27. One was found in volume 28. "J384" stands for "ANNALS OF PHYSICS". If you ask for "pion you will find "pions", also --- but not if you ask for "pion*".)

The examples exercised only a few of the TIP commands, but it may have conveyed a notion of how one works with TIP. The TIP commands may be combined in many different patterns. Users develop ingenious strategies for filtering out irrelevant articles without losing the ones they want.

While some TIP programs make searches and print lists, other TIP programs take notes on how the system is used. The data thus collected are periodically analyzed, and modifications and adjustments are continually made. TIP has developed through a process of guided evolution.

A bibliography, two review articles, and a catalog of the books in the Student Center Library have been prepared with the aid of TIP, and it is now being used to prepare a catalog of the journal and periodical holdings of the M.I.T. libraries.

Most of the work done thus far with TIP has been hampered by the slow pace of typewriter output. We are looking forward eagerly to cathode-ray displays.

Interactive Information Transfer

The purpose of Project INTREX is to conduct experiments that will clarify design objectives, methods, and techniques for information-transfer systems of about 1975. Emphasis is placed on the word "experiments".

Experiments have been planned in four main areas:

1. bibliographic access,
2. physical access,
3. fact retrieval, and
4. network integration.

Thus far the project has concentrated on the first and second.

Bibliographic Access

The purpose of "bibliographic access", of course, is to take the user from a nebulous idea of what he wants to the accession numbers (or equivalent identifiers) of the documents that will satisfy his requirements. Most of the INTREX effort towards that end is centered upon a computer-based "augmented catalog" that will approximately contain 50 "fields" of information about each of approximately 10,000 documents in materials science and engineering: journal articles, theses, and reports as well as books. With so many kinds of information it should be possible to determine which kinds are helpful enough to warrant inclusion in an operational system. Much of the work of selection of the 10,000 documents is being done by the research people and will serve in the planned experiments. Consoles will be located in the Materials Science and Engineering Center and the Engineering Library, and the experiments will be conducted within the context of actual use.

Physical Access

Bibliographic access must of course lead directly to physical access. Limitations of the present technology make digital storage and processing of a library-sized corpus uneconomic. The course being followed by Project INTREX is therefore to hold the substantive documents themselves in a non-digital microform storage system associated with the computer, and to use the computer to execute their delivery to the user. Images of the pages of the 10,000 documents are being made in microfiche, and a computer-controlled subsystem for picking out and scanning selected pages is being constructed. Plans call for experimental investigation of such interrelated factors as the speed, the form, the resolution, and the cost of physical access. By varying the parameters and making measurements of preference and performance under conditions of actual use, optimal engineering compromises will be approached and design objectives for operational systems will be formulated.

Advanced Experiments in a Library Context

Looking beyond the experiments with the 10,000-item collection in materials science and engineering, Project INTREX is conducting design studies that postulate a corpus of a million documents. At the same time, the M.I.T. Engineering Library is being reconstructed in such a way as to provide for simultaneous operations in conventional and computer-based modes. Card catalogs, book stacks, reading tables, microform equipment, and computer consoles will be brought together in an arrangement designed for advanced experiments in an operational library setting.

Synthesis and Prospect

Throughout MAC, TIP, and INTREX, and indeed throughout M.I.T., there is a feeling that a great and fundamental change is taking place in the way men relate to information. The force behind the change is the computer, of course, but it is not the same computer we have known these last 20 years. It is the computer cast in the new role of the moldable and retentive yet dynamic medium --- the medium within which one can create and preserve the most complex and subtle patterns and through which he can make those patterns operate (as programs) upon other patterns (data) derived from nature or the works of other men. In that role, the computer will change the very nature of libraries and information systems.