TUTORIAL ON

ELECTRONIC DATA PROCESSING CONCEPTS

Instructor Edition.

Edited by
Paul J. Fasana
Columbia University Libraries
New York, N. Y.
and
Russell Shank
The Smithsonian Institution
Washington, D. C.

Based on materials prepared and presented at the 1967 ASIS Annual Convention

Principal Tutor: Bruce Stewart

Partially supported by a grant from the National Science Fourndation (GN-657)

Tutorial Subcommittee
Conference Program Committee
American Society for Information Science
(formerly American Documentation Institute)
Annual Convention (1967)
New York, 1968

TUTORIAL ON ELECTRONIC DATA PROCESSING CONCEPTS

Instructor Edition.

Edited by
Paul J. Fasana
Columbia University Libraries
New York, N. Y.
and
Russell Shank
The Smithsonian Institution
Washington, D. C.

Based on materials prepared and presented
at the 1967 ASIS Annual Convention

Principal Tutor: Bruce Stewart

Partially supported by a grant from the
National Science Fourndation (GN-657)

Tutorial Subcommittee
Conference Program Committee
American Society for Information Science
(formerly American Documentation Institute)
Annual Convention (1967)
New York, 1968

PREFACE

A prime responsibility of a professional society is to foster continuing education activities covering new developments in topics of importance to the work of its members. This is particularly true in rapidly expanding and highly complex technologies such as those in the field of information science.

With this view, the 1967 Conference Planning Committee of the American Documentation Institute (now the American Society for Information Science) chaired by Paul Fasana of the Columbia University Libraries, established a Tutorial Subcommittee to organize training sessions for presentation at the Conference. The Subcommittee, under the direction of Russell Shank, then Associate Professor at the Columbia University School of Library Service, agreed to develop three workshop tutorials for the following areas: elements of information systems; electronic data processing concepts; and generalized programming languages and systems.

The tutorials on these topics ran concurrently. Each began with a general session in which the tutorial leader gave an overview of the topic to be covered. The participants were then formed into smaller workshop or seminar groups for detailed instruction by a team of tutors. Each tutorial lasted the entire day. The general sessions were limited to about 100 people; seminar groups were limited to about 25 people. In the seminar groups each of the tutors either covered the entire topic simultaneously, or presented a part of the information to be covered, repeating their presentations as groups were rotated among them.

In planning the tutorials it was apparent that syllabi or workbooks were needed to assure that the basic information to be presented was uniform and organized for the instructors of each of the groups. It was assumed that if syllabi were carefully prepared they could be made generally available and be useful in similar courses at other national and local meetings of information science groups.

Three manuals, covering the three different topics, were prepared and used experimentally at the Conference. Initially, it was hoped that each manual would contain a comprehensive outline of the topics to be presented, a display of the illustrations and visual material used in the lectures, glossaries, bibliographies, and problems. It was further assumed that the sessions would be more meaningful if an instructor's version and a student's outline (with sufficient space for notes) were prepared. The instructor's edition would have enough detailed information to allow other instructors to present the course.

The variation among the approaches to the three topics treated in the tutorials made it difficult to attain this objective of uniformity in style of presentation, at least on the first attempt. All three manuals have been extensively revised for publication. This material will undoubtedly be improved through refinements as the tutorials are presented elsewhere in the future.

This package contains both an instructor's and a student's version of the syllabus for each topic. Undoubtedly other instructors will wish to make modifications of these manuals to suit local needs and instructor's talents. These manuals are offered, therefore, primarily as examples for

for those who might be planning similar tutorials. The student edition may be produced in quantity locally as required.

Very briefly, the scope of each of the three sessions of the 1967 tutorials was as follows:

Tutorial I. Elements of Information Systems.
Paul L. St. Pierre, Principal Tutor

An introductory course for those with no previous experience or formal training in systems analysis. Objective will be to familiarize participants with the techniques (file analysis, record analysis, flow-charting, costing) and the terminology of systems analysis. Those who complete this tutorial should have a knowledge of what systems analysis is, what function it serves and how it can be applied.

Tutorial II. Electronic Data Processing Concepts.
Bruce Stewart, Principal Tutor

For those with little or no experience with EDP equipment. Emphasis will be on a functional description of various types of equipment. Participants will be given an understanding of how such equipment works. They will not be trained to operate particular machines.

Tutorial III. Generalized Programming Languages and Systems.
Thomas K. Burgess, Principal Tutor

For those with considerable systems analysis and programming experience. Presentation of a comparative analysis and review of the various programming languages currently available, especially as they apply to information storage and retrieval. The relative merits of different program languages for use in textual analysis, file structure, file manipulation, and similar topics will be stressed. Program languages to be covered include FORTRAN, COBOL, PL 1, SNOBOL, and ALGOL.

Russell Shank
Washington, D.C.
July, 1968

TABLE OF CONTENTS

TABLE OF FIGURES

History & Concepts

Introduction

During the past forty-five years, more has happened than in all previously recorded history. Compressing trends of history and the civilization of man will give some perspective on what has occurred:

Stone Age	500,000 years
Bronze Age	50,000 years
Iron Age	5,000 years
Industrial Age	500 years
Atomic Age	50 years
Space Age	5 to 7 years.

A great portion of our recent technological advances have been dependent upon the computer.

The First Twenty Years

The computer revolution has come in waves, each wave opening new areas of application. The result has been an increased demand for computers.

One can trace five major waves of computer history. The period 1949 to 1960 saw the development of vacuum tube scientific and commercial computers. During that time, small computers, like the

IBM 650, took over payroll and accounting chores and some order processing, inventory control, and production scheduling.

The first wave, during the years 1944 to 1949, is marked by the invention of the stored program computer. Examples of this are the Mark I at Harvard, which was an electro-mechanical machine weighing several tons; the electronic ENIAC which was developed at the University of Pennsylvania. This early work lead to UNIVAC; the first UNIVAC was retired to the Smithsonian Institution in 1963, 12 1/2 years after its birth.

The years 1958 - 60 are of importance because of the appearance of transistorized computers. These were smaller, more pliable, and potentially cheaper. At the same time, computer applications became much broader, extending to the use of process control computers in running petroleum, chemical, steel and electric generating plants. At this time, automatic computer programming emerged as a major advance.

1960 - 64 may be called the era of proliferation. At this time automatic programming became economic, and progress in simulation led to widespread use of computers for product and process design. Medium and small sized computers replaced punch card systems with solid state computers (chiefly the IBM 1401).

Time-sharing came into use in the years 1964 - 1967.

Major Computer Manufacturers are

IBM	526 million earnings in 1966
Honeywell	Minneapolis
UNIVAC	St. Paul
RCA	
Burroughs	Detroit
NCR	Dayton
CDC	Minneapolis
GE	Phoenix
SDS	Santa Monica

US built computers - over 41,000 1966

30,785 figures

Over $10 billion.

In 1966 10,903 systems were shipped.

Spectrum of Computing

"Computer tree"

Abacus 400 B.C.

Pascal's Adding Machine 1642.

Liebnitz Multiplier 1671.

Jacquard Coom 1745.

Electromagnet (Wm. Sturgeon, England) 1825.

Telegraph - Morse.

Pulse Code for Machines - Boudot.

George Boole - University of Cork 1847 - Lobic in mathematical forms.

Vacuum tube (Fleming, English) 1904.

ENIAC (J.W. Mauchly El J.P. Eckert) University of Pennsylvania, 1946.

Begins 1946 ENIAC Electronic Numerical Integrator and Computer.

It was built at the University of Pennsylvania by J.W. Mauchly and J.P. Eckert for the Army and was used to calculate ballistic tables for the Ordinance Department.

ENIAC filled a space 30 x 50 feet and had 18,000 vacuum tubes, 130,000 watts of power (major cooling problem) and could perform 5,000 additions per second, 350 multiplications per second.

First fully electronic machine.

Others followed:

EDVAC - University of Pennsylvania.

SEAC - National Bureau of Studies.

Whirlwind I - MIT.

UNIVAC I

IBM 650.

Generations of Technological Development.

End of 1950's, transistors invaded the computer market.

Discuss 2nd Generation - Microseconds.

Discuss 3rd Generation - Nanoseconds.

Difficult to define 3rd Generation; consequently it is difficult to tell where 2nd Generation leaves off and 3rd Generation takes over.

Future.

4th Generation characterized by:

Terminals.

Display Capabilities.

Multiprogramming.

Multiprocessing.

Time-sharing.

Total Information Systems.

10,000,000 calculations per second.

Chess - Example of limitation even in future.

Estimated	10^{120} possible continuations of the game.
At 10,000,000 calculations per second	10^{113} seconds or 10^{103} centuries.

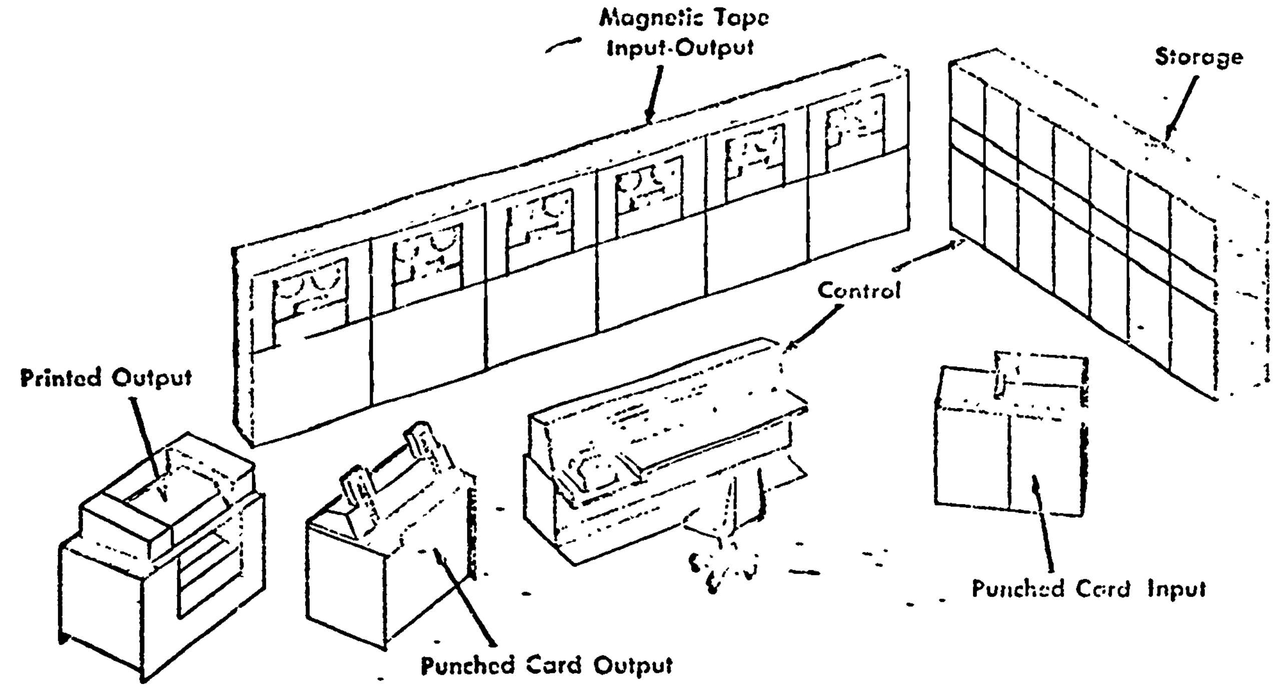
Hueristic programming - encourages computer to learn. (May be key to future.)

Basic Concept.

Computers do not "think".

Computers are tools, must be applied.

Computers extend, augment man's capabilities.


Magnetic Tape
Input-Output
Storage
Control
Printed Output
Punched Card Output
Punched Card Input

Computer Components and Organization

General

This is to acquaint the student with the basic organization of the digital computer and some of the more fundamental concepts inherent in the operation and use of digital computers.

Computer Organization

A digital computer consists of 5 logical components. Any particular computer may have one or many physical devices in each of the five categories.

Input is the method by which data is given to the computer. Input may be instructions controlling the computer or data to be operated on.

Storage is the device(s) which stores data read by the computer. We are concerned here primarily with the main storage of the machine as opposed to auxilliary or mass storage devices. In the majority of computers today this is core storage, characterized by high speed and high cost. Core is used primarily for storing instructions, controlling the operation of the computer and relatively small amounts of data the machine is manipulating.

Arithmetic. A digital computer solves problems by performing arithmetic operations on data represented in a numerical form. Generally speaking this is the binary number system. The arithmetic unit is, therefore, composed of the circuitry to perform such operations as addition, subtraction, multiplication, division, and shifting.

Output is the method the computer uses to communicate with the user.

Control Unit. It would seem with input, output, storage, and arithmetic units the computer has all that is necessary to operate. However, before the machine can function it must interpret the commands given by the user and/or operator to cause the various units to operate. The component (circuitry) which performs this function is the control unit.
An additional function performed by the control unit provides the computer with the ability to make decisions and take alternate courses of action based on the result of these. This is one of several capabilities without which the computer would have very limited use. The sophistication of the decision making process varies among computers, but no matter how limited this capability it allows the machine to assume certain human-like attributes.

Computer Logic.

Before discussing the way in which these components work together and the problem-solving mechanism, it would be useful to cover the concept of the stored program and answer the question. "What is a program?"

The Program

A program is a sequence of instructions that guide the computer in its operation. A computer is a "perfect" idiot. Thus, a program must describe an operation in the most basic steps. As an example if you were to be programmed to write in your notes, you would:

> Lift arm; move to pencil; down to pencil; close fingers; lift pencil; ...etc. (Note that we forgot to "open fingers" and assumed they were open. This is the same problem facing computer programmers.)

The Concept of a Stored Program

One of the more important concepts that brought computers to the level of usage they enjoy today is the concept of the stored program. This means that the instructions that control the operation of the computer are stored in the main memory (storage) of the computer itself. The earliest computers were controlled by a wired panel which was actually part of the circuitry.

A computer can perform only two basic operations. It can add (subtract, multiply, and divide by adding), taking advantage of electronic speed, and compare (equal, unequal). Operations are controlled entirely by the stored program, which was initially prepared by a human programmer. Every possible sequence of events must be provided for, because a computer can only follow its stored instructions. If it continues to run, but executes stored instructions at the wrong time, the result is trash output.

Levels of Programming

When the first stored program computers were built it was necessary for the individual who wanted to solve a problem to write instructions to the computer in the machine's "natural" language or numbers.

Machine Language

This requires that the programmer use a set of numeric codes to tell the machine to perform its operations. It requires the programmer to keep track of the numeric addresses of the storage locations where instructions and data are kept.

Example: [60 1735 7431]

Assembly Language

The use of machine language rapidly became an impossibly complicated task for the new profession of programming. The subsequent development of assembly languages allowed the programmer to use mnemonic abbreviations to instruct the computer and symbolic addresses to reference data and instructions.

Example: [S WHTX,GRSP]

Compiler Language

The development of assembly language was a significant step forward, but could still only translate one assembly language instruction into one machine language instruction and was still "machine dependent." Compiler language allowed the programmer to state his problem in terms of the problem rather than being bound by the design of the machine.

Example: [SUBTRACT WITHHOLDING-TAX FROM GROSS-PAY GIVING NET.]

It is important to realize that an assembler and a compiler are both programs which were written to convert what the programmer writes into a set of instructions that the machine can interpret. The use of current compiler languages has greatly simplified computer programming in comparison with its earlier complexity. But even today, there is a tremendous manpower shortage: 100,000 computer related jobs and 50,000 open positions.

Computer Function

Now that we have covered the basic components of the computer and the concept of the stored program, we shall illustrate how they are combined to solve a problem.

Component Interrelation

The control unit communicates with each of the components of the computer. The communication lines from the other components to the control unit allow the control unit to monitor their operation. For example, if a tape unit has a malfunction, the control unit would be aware of it. In many computers, there are communication lines between individual components other than the control unit. This capability allows the machine to do such things as simultaneous read, write, and compute.

Problem Solution

The functioning of computer components can best be illustrated by following the computer through the execution of a simple program. The problem presented is the solution of the equation D=A+B-C.

(INSTR. NOTE: Explain the execution of all steps necessary to solve the problem)

Storage Devices

Characteristics of Storage

Data Representation. Characters, whether numbers, letters, or special characters, are represented in the computer by a binary scheme. It is binary because of the bi-state condition of computer storage and circuits. The circuit is on-off; a magnetized spot is present or absent.

Addressability. Each item of data must be uniquely identifiable by location and individually retrievable. Explain concept of destructive read-in, non-destructive read-out.

Types of Storage

Internal Storage is the integral physical part of the processor, directly controlled and automatically accessible. Examples: magnetic core and magnetic drum.

Secondary Storage is not an integral part of the processor, but directly connected and controlled. Examples: magnetic disk and magnetic tape.

External Storage is divorced from the processor, but holds data in the form required for processing. Examples: magnetic tapes, paper tapes, and punched cards.

Device Types

Magnetic Core is a tiny, doughnut-shaped ring of ferrite capable of retaining either of two magnetic states.

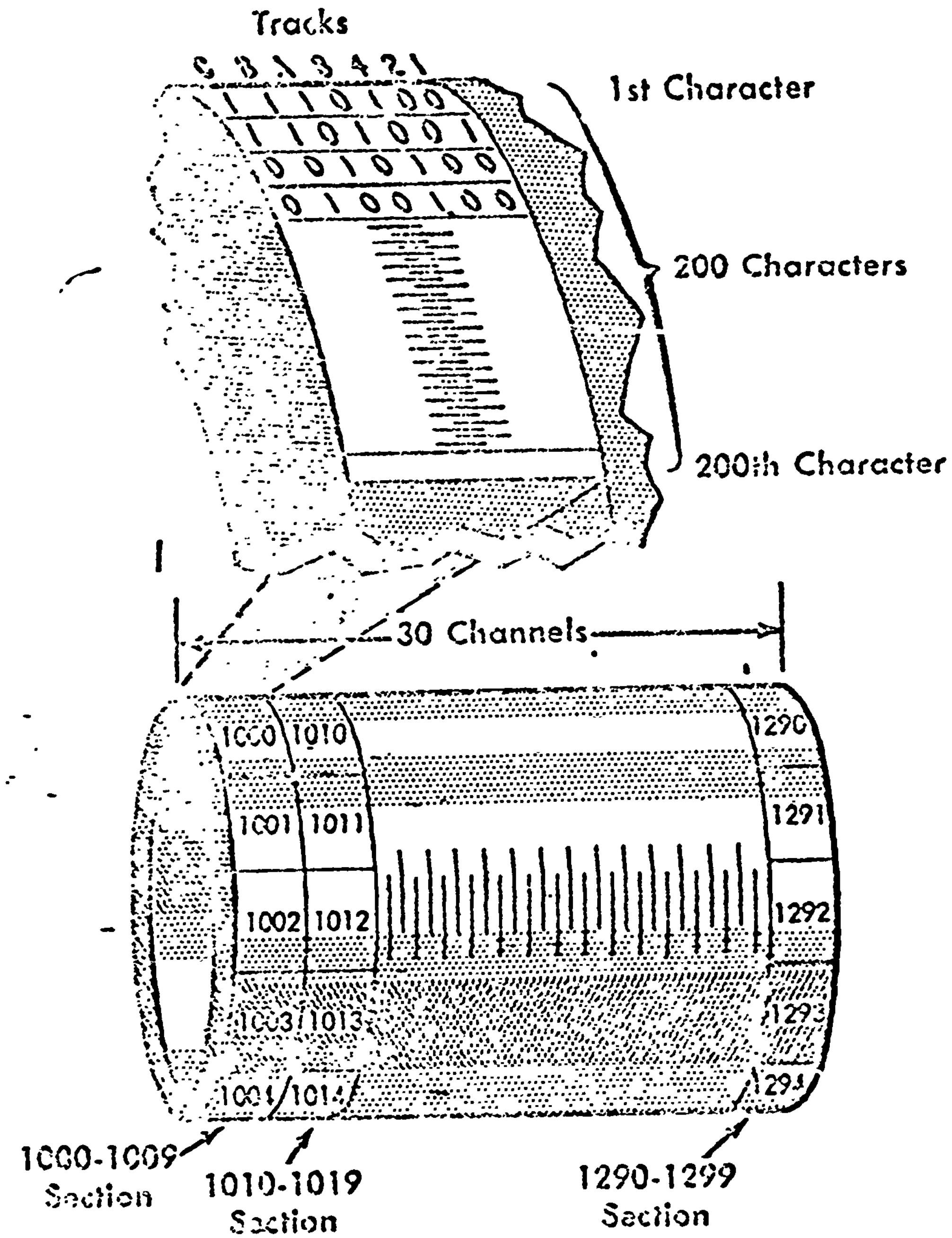
(Explain operation.)

Access times range from several nano seconds to several micro seconds. Magnetic-core storage is both erasable and non-volatile. The cores are individually inexpensive but the electronic circuits which they require make them among the most expensive. Wires as small as one-thousanth of an inch in diameter are used in twist or grids to sense cores.

Magnetic Drum

The drum is actually a metal cylinder coated with magnetic material which rotates at high speed. Heads for reading and writing from the drum are mounted in the drum housing. "Bands", or groups of recording tracks, are set up around the drum for storing data. (There are several bands per inch in the physical drum.) Each band around the drum is divided into many unit areas, each of which stores a bit, or magnetized spot. The bit is written in or read from a particular unit area as it rotates under the read-write head. There is no movement of the read-head. Multiple read heads greatly increase the rate of input and output. Nominal capacities for magnetic drums will run approximately 4 million characters with

Magnetic Drum



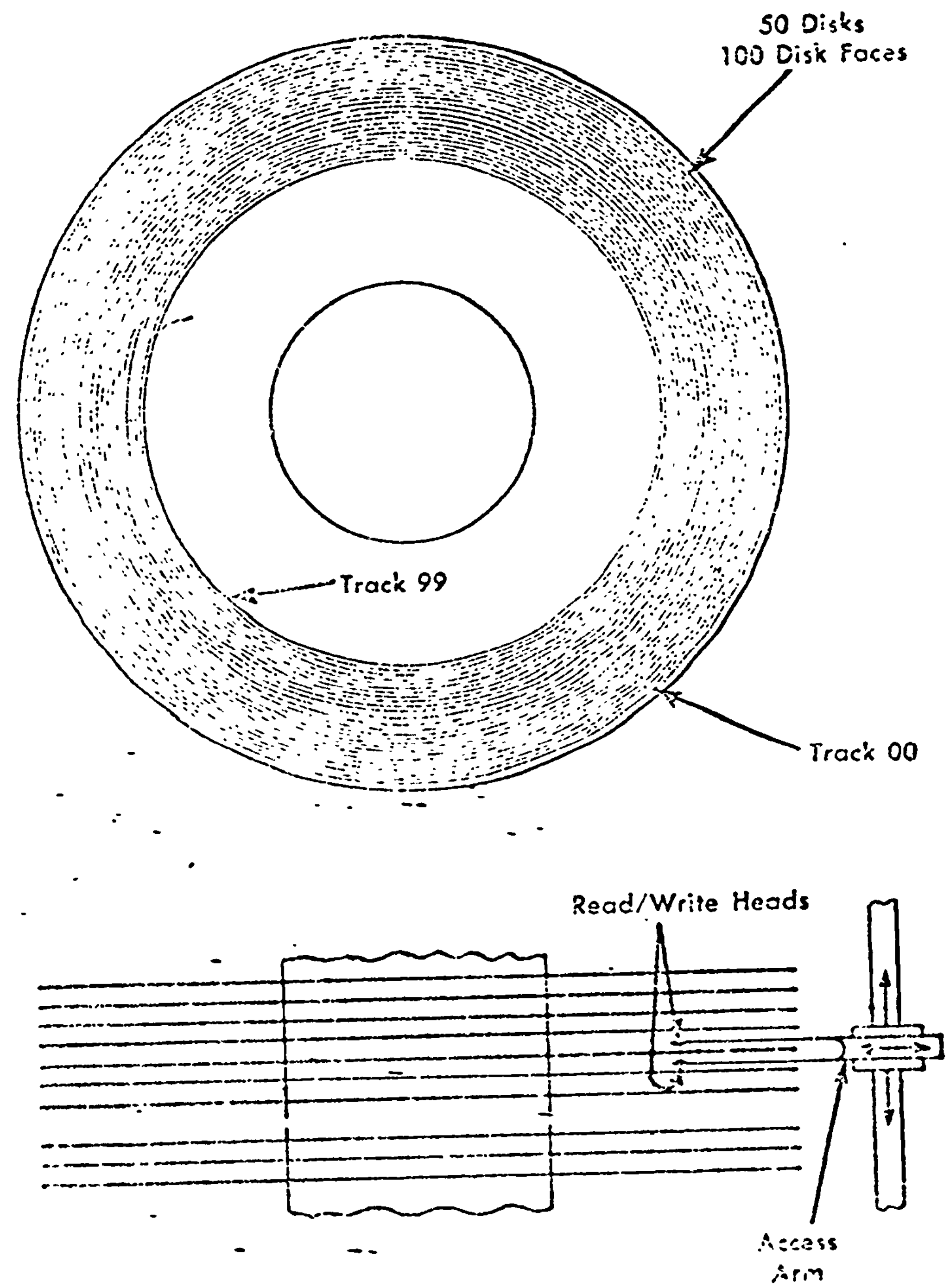
access times in several milliseconds. Data can be transferred in or out at rates exceeding 1 million characters per second. The cost of drum storage is considerably more than that of disk or high speed tape drives. The fastest drums cannot be installed on most medium and small scale computers because their circuity is not fast enough to accept data and move it to main storage as fast as the drum can transfer it.

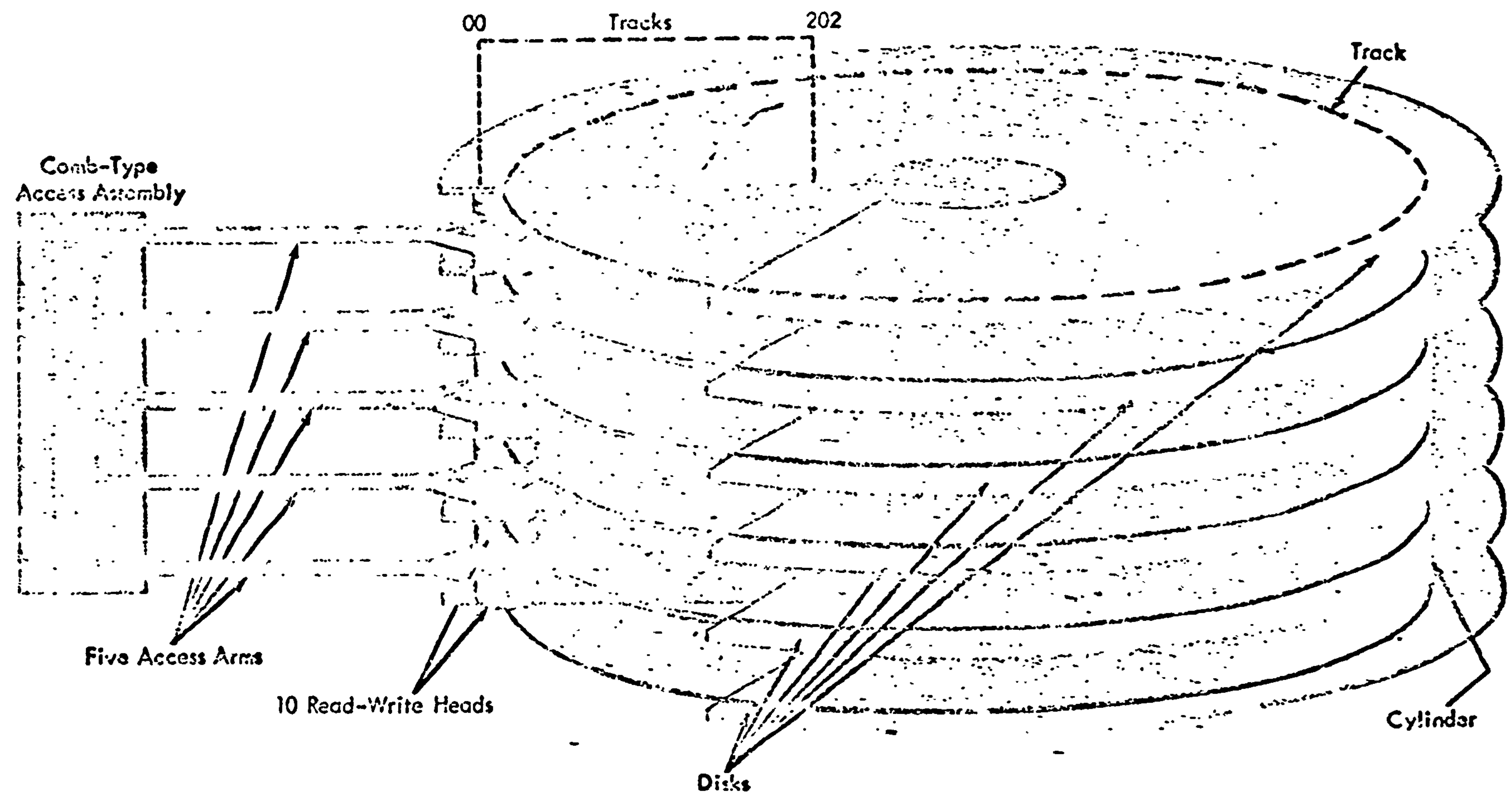
Magnetic Disk

This is a metal disk with magnetic coating on both sides. The disks are mounted on vertical rotating shafts, with circular data tracks on both sides. Read/write heads are mounted on movable arms which slide in and out to access a particular track. Bits are represented by small magnetized spots on the surface of each track. The maximum access time to find an item on a particular track is the time required for one revolution of the disk plus the movement of the access mechanism to the proper track (on the order of 100-200 milliseconds.)

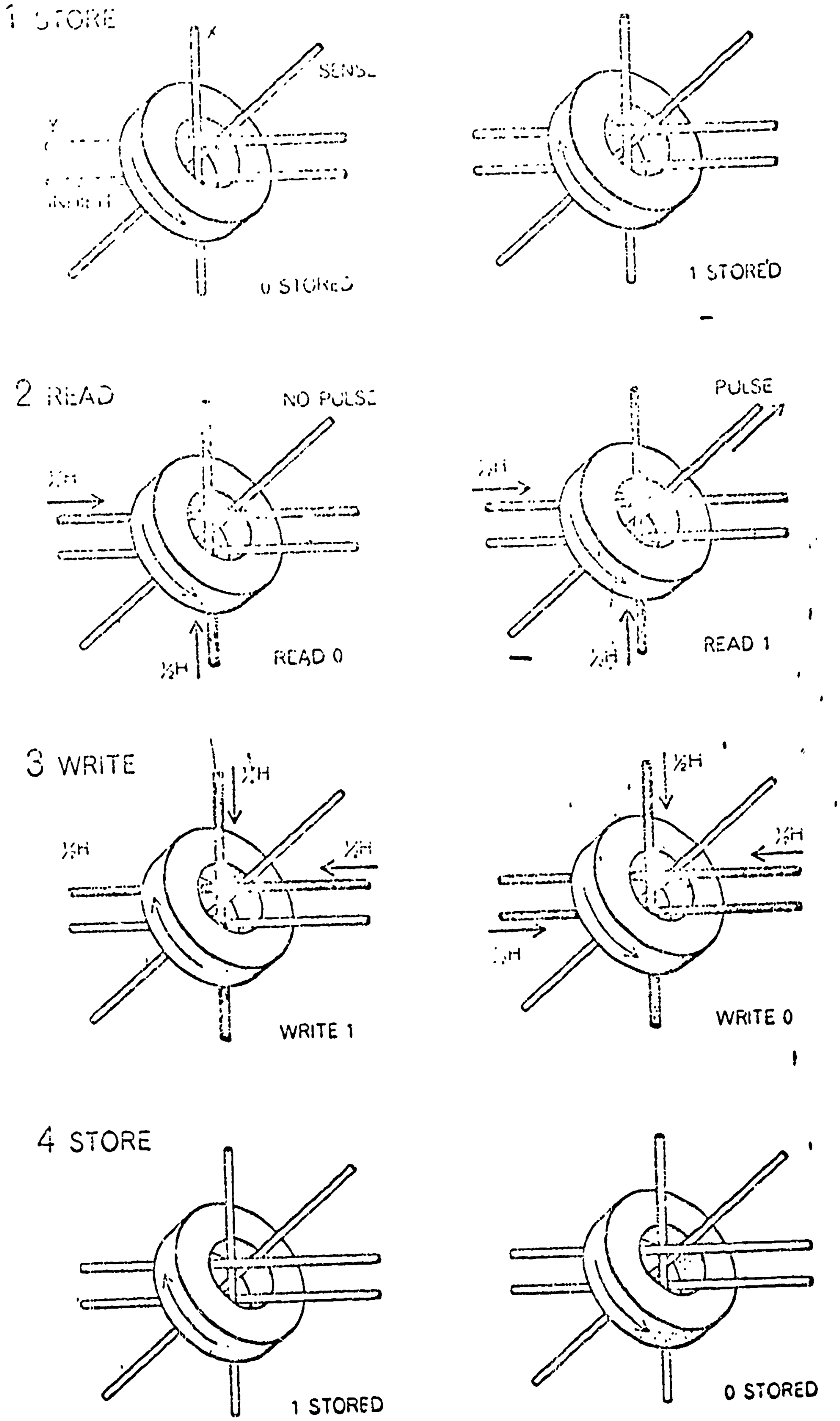
Illustrate cylinder concept - since the access mechanism has all read/write heads located at the same position on each track vertically, access time is saved if the logical cylinder is used for data storage.

Magnetic Disks





00
Tracks
202
Track
Comb-Type
Access Assembly
Five Access Arms
10 Read-Write Heads
Disks
Cylinder


1 STORE
0 STORED
1 STORED
2 READ
NO PULSE
PULSE
½H
READ 0
READ 1
3 WRITE
WRITE 1
WRITE 0
4 STORE
1 STORED
0 STORED

There are two widely used forms of disk storage. One form is that of small removable disk packs holding from 7 to 29 million characters. The second uses large disk units which are more or less permanent components of the computer (as large as 255 million characters). Transfer rates vary from 75,000 to 175,000 characters per second. The cost of disk storage units is considerably less than drum or core.

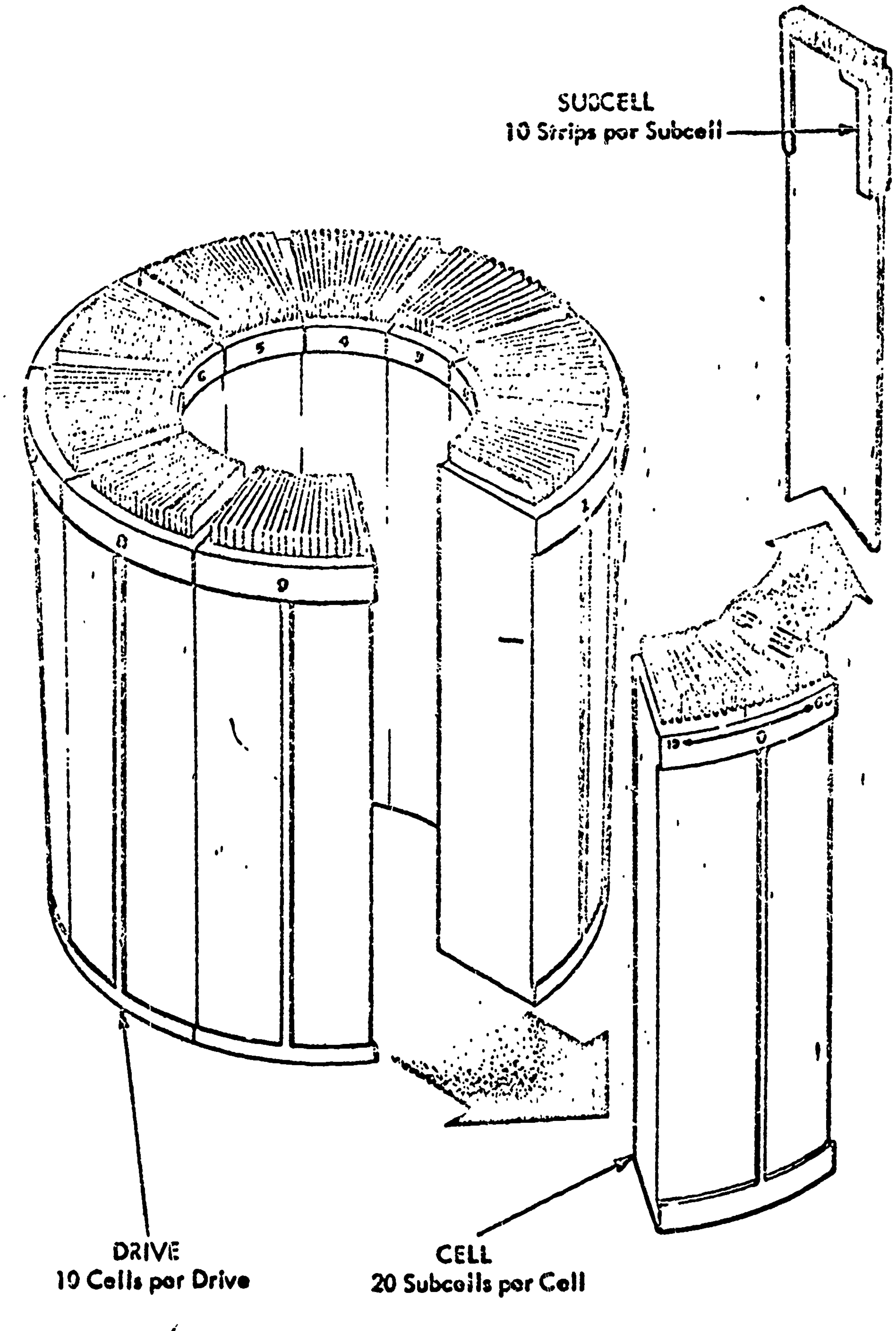
Data Cells

The data cell offers great promise in applications requiring large data bases that need be accessed in reasonable but not at extremely fast rates. The data cell has a capacity of 400 million characters. The drive consists of ten data cells, each with a capacity of 40 million characters. Each data cell contains 20 subcells comprised of 10 magnetic tape strips. Transfer rate is 55 thousand characters per second with access time between 175 to 600 milliseconds.

Once the address of a data item is determined, the cell containing the strip is rotated under the reading drum. Once the strip is on the drum it is read in much the same manner as a magnetic drum with the exception that there is only one read/write head for five tracks. There are a total of 100 tracks per strip with 20 read/write heads.

Magnetic Tape

Magnetic tape consists of a mylar or acetate ribbon coated on


SUBCELL
10 Strips per Subcell
DRIVE
10 Cells per Drive
CELL
20 Subcells per Cell

one side with some magnetic material. It is usually 1/2 inch wide (similar to the tape used on home recorders). Each 2400 ft. roll can store 15 million characters. The effective storage capacity however, is less because of recording requirements. Magnetic Tape is extremely durable under ordinary operating conditions. Some tapes have been read satisfactorily 50 thousand times. It is also easy to recopy tapes if they are in danger of wearing out. It's the most widely used form of bulk storage, cost per tape reel of only approximately $70.

External Storage

External storage facilities hold large amounts of data in a form required for processing at low cost. Three major media are:

Magnetic tape.

Punched cards.

Punched paper tape.

A major economic consideration in external or inactive storage is the cost of recreating data vs. cost of storage.

Storage Organization

With the advent of random access capability for large storage capacities, various methods of selecting a particular data item out of millions resident in a store have been developed.

Sequential organization and access

This can probably be best illustrated by an axample, such as student information on magnetic tape in social security number order.

It would mean that all the information for a particular student would be together on a tape. The student with lowest social security number would be first, the next lowest would be second, etc. until the last student having the highest social security number. To locate an individual student on the tape, the computer would read the first number of the student for which it was searching. If the number on the tape was lower, the machine would read the next number. It would continue this procedure until it either found the correct number or it found a higher number on the tape, which would indicate the number for which it was searching was not present. It is essential to realize, however, to read the numbers from the tape, the computer must also read all the information that goes with the number. For many tasks (inventory, payroll), sequential organization is preferred; however, in many applications it is extremely inefficient. Magnetic tape, because of its physical attribute, must be sequential; random access devices may also be organized sequentially if desired.

Random Organization and Access

There are numerous methods for organizing and accessing non-sequential (random) data sets. To provide a basic understanding of the concepts and capabilities of this type organization only one method will be presented. Direct organization is built around a "key" (identifier of some data item, so a key may be an employee number, part number, etc.). By numerically manipulating this key and very carefully organizing the file, the absolute address of a data item can be calculated from the key.

Indexed Sequential Organization and Access

This is a reasonably successful attempt to gain the advantages of both sequential and non-sequential organization and access. With it, data is organized sequentially with separate areas of the device set aside for indices which allow for reasonable access times to a single entry within the data set. For example: consider an inventory system where inventory information is sequentially organized with two indices to help find an individual part number. The first index would be a cylinder index that would contain the location on a cylinder where the track index can be found; it also would contain the highest part number to be found on an individual cylinder. By searching this index sequentially, the computer could find which cylinder contained the data. The machine would then proceed to the track index which would contain the highest part number on each track and the home address of the data track. By searching this track index the computer could go the track containing desired data item.

Input/Output

General

Data processing is characterized by large volumes of input and output. Computer processing speeds are so fast that the computer usually waits on the relatively slow input and/or output media. The biggest bottleneck is, usually, conversion of data to machine-readable form.

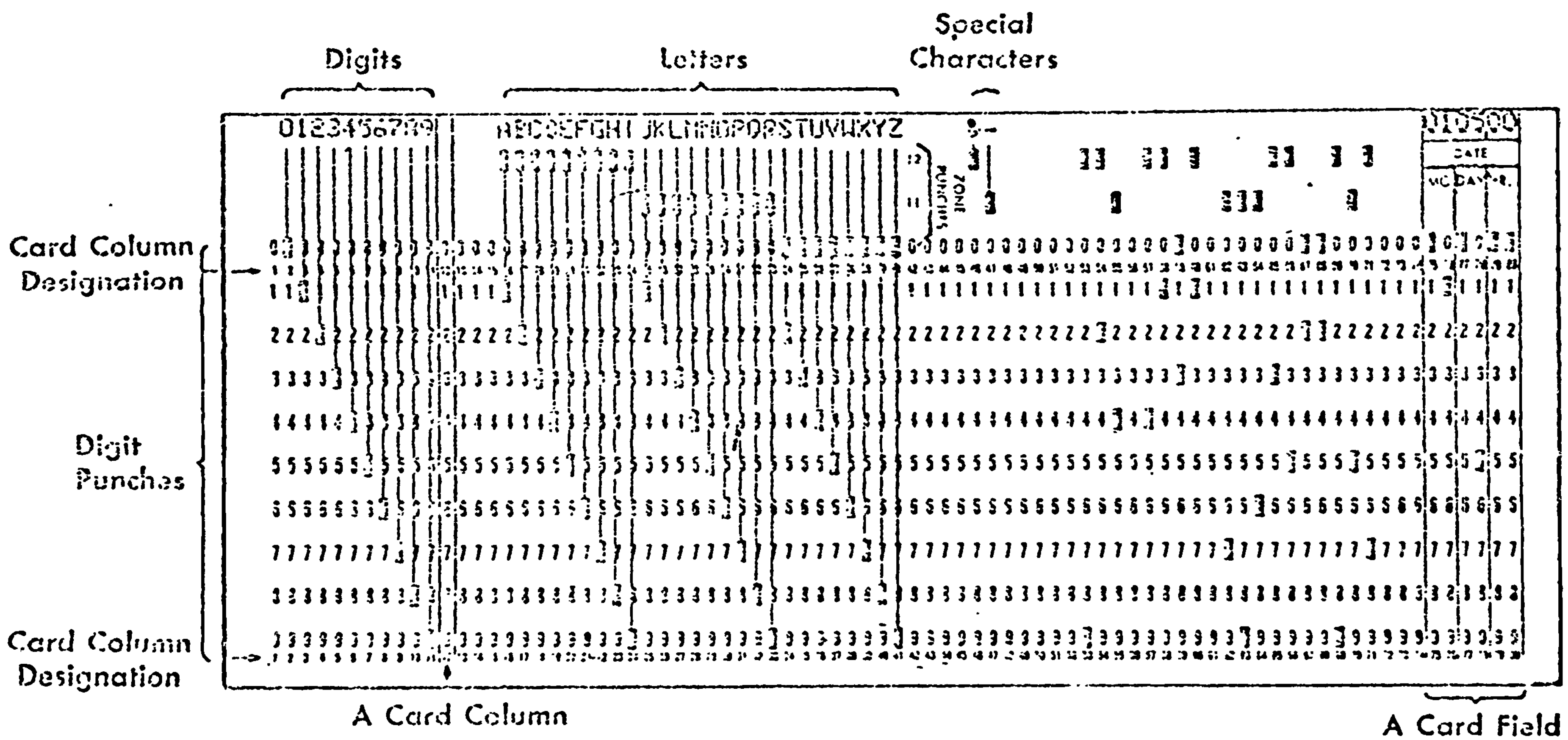
Input Devices

Punched Card

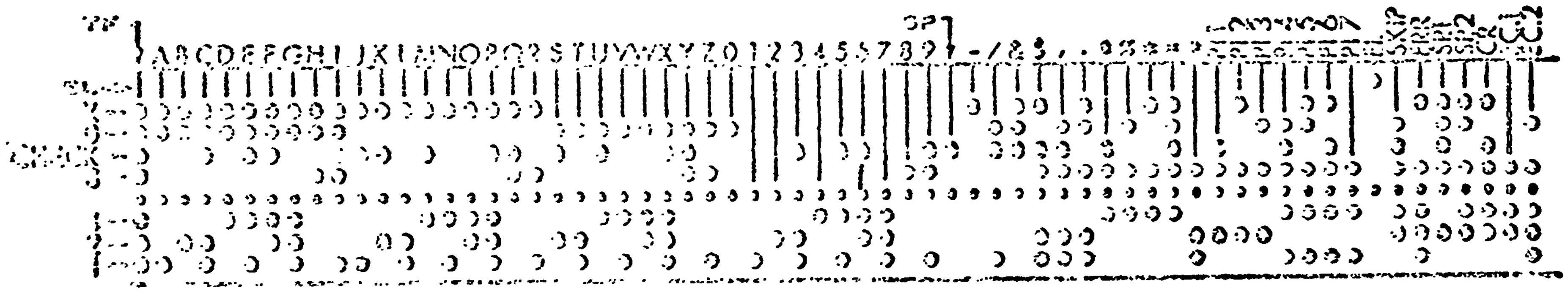
Explain card characteristics from example prepared on keypunch. Cards have an input reading rate of 800 to 1000 cards per minute, or 80,000 characters per minute, which is extremely slow. This is due to the physical limitation of mechanically moving card through reader.

Punched Paper Tape

Explain characteristics from example prepared on typewriters which punch tape as they type document. Computer input reading rates vary from 20 to 500 characters per second; or approximately 30,000 characters per minute. Again, the physical limitation is still that of mechanically moving paper tape.


Digits
Letters
Special Characters
0123456789
ABCDEFGHIJKLMNOPQRSTUVWXYZ
ZONE PUNCHES
DATE
MO. DAY YR.
Card Column Designation
Digit Punches
Card Column Designation
A Card Column
A Card Field

Punched Paper Tape



Magnetic Tape

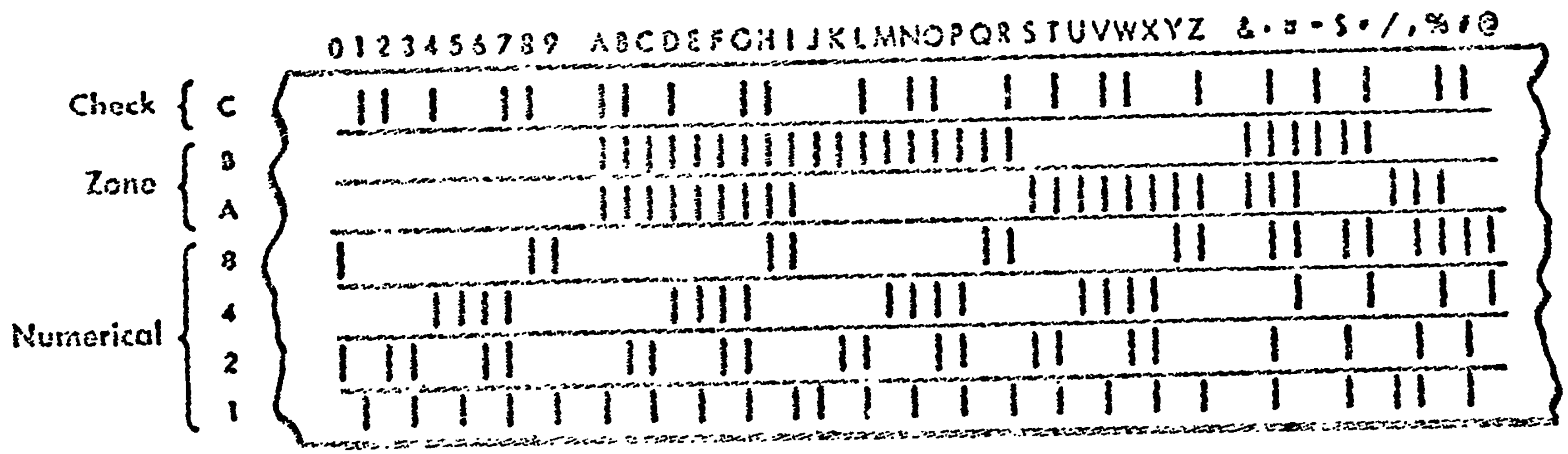
Explain data representation from example. Magnetic tape is currently the most widely used form of input/output media as well as the most inexpensive form of bulk storage. It is especially useful where large volumes of data must be transferred. Reading speed for magnetic tape is 75 to 112 inches per second. Explain read/write heads--start/stop motion. Transfer rates are 30,000 to 200,000 characters per second. Explain density (e.g., 500 bpi=10-80 column cards per inch). Although tape reels are inexpensive and store well, magnetic tape transports are extremely expensive devices.

Character Readers

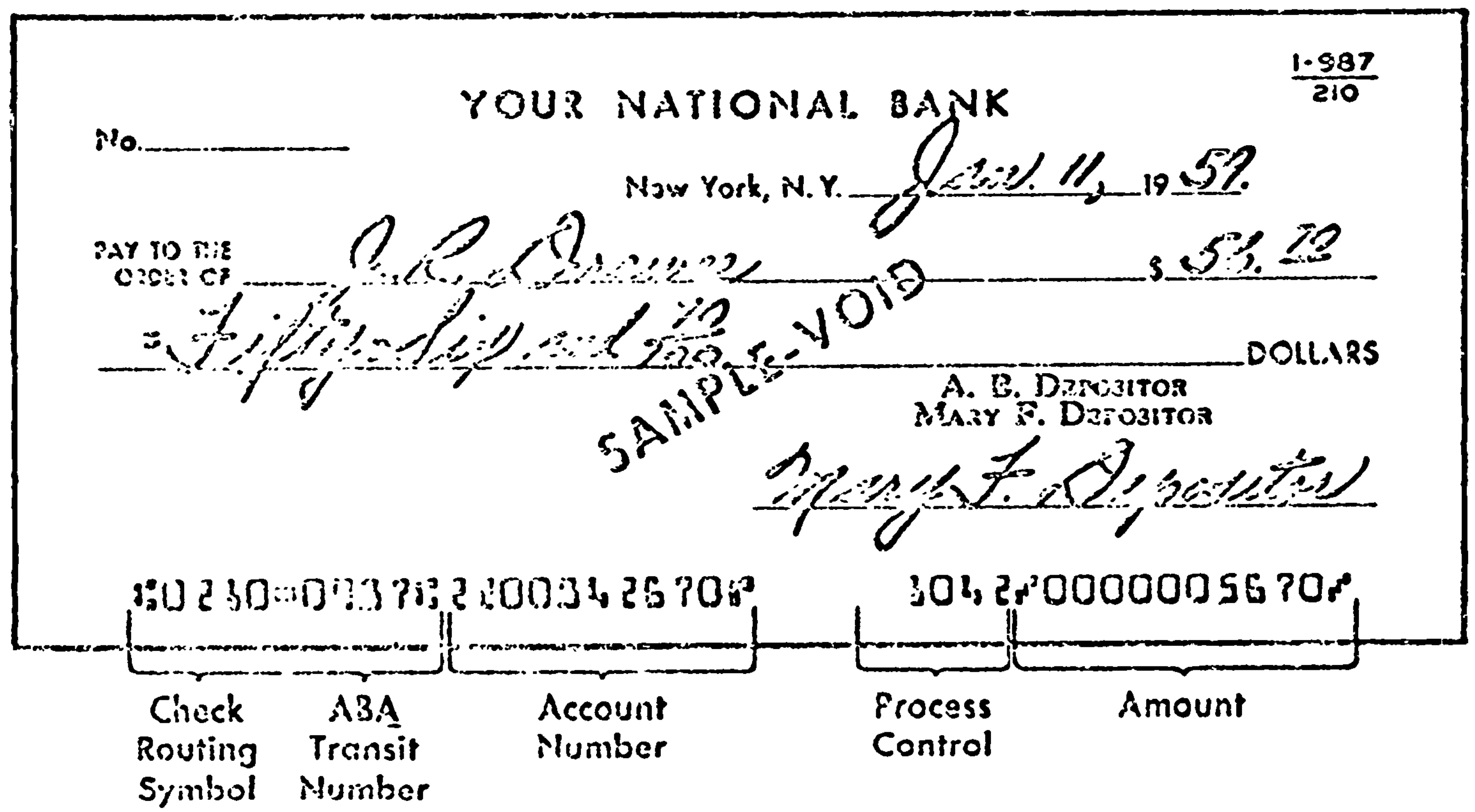
Two types are in use today.

Magnetic ink character readers. Data is printed with magnetizable ink and passed through a permanent magnetic. Used primarily for special purposes such as encoding checks, ticket stubs for eventual use of source document and input, etc. Input reading rates on special purpose MICR vary from low to medium speeds. The biggest advantage comes in being able to use the source document as input to the computer.

Magnetic Tape



Magnetic Ink Characters



Optical Character Recognition

Scanners actually "recognize" characters printed on paper by "seeing" light and dark. Current models can accept 8 1/2 x 11 pages and multiple type fonts. Optical scanning is the real hope in solving the problem of converting printed information to machine-readable form.

Output Devices

We will consider primarily those devices which create output in a form for use by humans, excluding those which merely output results onto magnetic tape, etc. Communication between the computer and the human is the most critical aspect of man-machine interface. A basic tenet has emerged: the computer is man's tool; therefore the output system must be designed to facilitate the human and not the machine. Example: early computers punched card input and output; currently computer controlled graphical and visual displays are being developed.

Mechanical Printers

Used for most computer outputs. Generally mechanical printers can produce capital letters, numbers, and a few punctuation symbols only. Early models used the mechanical bar printing methods of old adding machines and were able to achieve a maximum speed of 100-200 lines per minute. Modern high-speed printers are designed on the principle of a rotating chain or train of characters (briefly explain the operation of a chain printer). Speeds of 600-1100 lines per minute, 132 characters per line are possible. Another type of currently available printer is a high speed electrostatic printer with rated speeds of thousands of lines per minute. This printer operates with thermoplastic "toner" which forms an image, which is transferred to paper and fixed by heat.

Visual Display

There are new kinds of output equipment which enable computers to display images and graphs as well as characters on a TV like tube. These devices have changed the whole concept of computer output. The speed of visual display devices is a real advantage. Permanent output copy can be provided by special display units which incorporate some means of copying the face of the display screen. The basic hardware for visual output is a cathode-ray-tube (CRT). Given a set of coordinates by the computer, a dot is flashed on the face of the tube. Complete pictures, including lines, circles, and letters are made up out of thousands of individual dots in this way. Many display devices contain character generating hardware that will cause a letter or number to appear on the display screen. It requires extremely complex computer programs to drive a visual display device. An example of the sophistication which is possible is a perspective view of solid objects as computed by the program. The object can be rotated on the screen or even made transparent if necessary. At present, such use of the computer is limited since the expense involved in equipment and programming is enormous.

Input/Output Considerations

There are two criteria of a given input/output media which are important.

Speed: A large scale computer must have very high input/output speed or much of its internal processing speed will be wasted.

Volume: Processing large amounts of data requires that medium to high speed media be used to enable processing.

Scientific vs. Commercial Data Processing:

Scientific data processing is characterized by low volume input/output with complex internal processing.

Commercial Data Processing is characterized by high volume input/output with relatively little internal calculation.

Special Techniques

The Concept of Multiplexing. (A fuller explanation is given in the following section on Time Sharing.)

Keyboard input becomes feasible even though limited by manual typing speeds, because a large number of keyboards can be connected to a single computer and all work seemingly simultaneously. (Cite examples of Airlines reservations systems or bank teller systems.)

Method of Operation: A computer is so fast that it can scan all terminals and accept input from each as necessary, giving the impression to each terminal that it is always connected.

Buffer Concept

A computer can accept and record data in storage faster than input units can supply it. Buffer storage is used to help compensate for this difference in operating speeds. Data is read into buffer area while the central processing unit continues with other operations until the read-in operation is completed. It is then transfered from buffer storage into main internal storage at electronic speeds, without being limited by the speed of a card reader. Buffering has the advantage of minimizing interruptions of the computer while data is being read in allowing more processing to be done per unit of time.

Time-Sharing

General

With the advent of time sharing electronic data processing has entered a new phase. Computer systems, which are described as being "on-line" and "real-time", are sprouting up in many business and industrial environments.

On-line means that input/output equipment is connected directly to the computer and functions as part of the computer's processing unit. Two-way communication is implied in such operation. An example is a teletypewriter connected to an operating computer by a communication channel.

Real-time operation means that the computer receives data, processes it, and returns results in sufficient time to control the functioning of the environment. ("Response time" is a critical factor in any definition of real-time and differentiates a real-time system for any other type of operation.)
The magnitude and complexity of modern civilization gives rise to many complex situations which are best handled by real-time computer systems. Air traffic control, banking, or automobile routing are just a few examples in which response time is a critical factor.

Time-sharing and teleprocessing (or remote use of computers through common carrier communication lines) are both essential to real-time system operation.

Organization of Time-Sharing Systems

Time-sharing is a result of the tremendous speed of large-scale computers. In a batch system, the computer works on one problem-program at a time resulting in long waiting times for users. Operating internally at tens of thousands of operations per second a computer spends a great deal of time, as much as 40%, literally waiting on its relatively slow input/output devices.

Time-sharing means that the computer actually processes many programs at once. In reality the computer processes one program for a set period of time (or time slice), then "jumps" to the next program in line, taking up where it left off in its last cycle. If the program being executed is delayed for any reason, such as output of data, its time slice is automatically terminated and the machine cycles to the next program.

Crucial in this operation is resource sharing. The computer must spend a great deal of time in "housekeeping". Each program's input/output must be kept separate. Data stored in memory by one program must be protected and not erased by the next program in cycle.

A major problem in the development of economical time-sharing systems was to develop a system in which the "overhead" was acceptably low. Early systems on early computers often failed because to much time was spent managing themselves so that no productive work was done.

Psychology of Time-Sharing

The interaction of the user and the processor is the most striking characteristic of time-sharing operations.

Human response time is in the range of 2 to 5 seconds while computer operations are in the range of macroseconds and milliseconds. As a consequence, many users can time-share a computer simultaneously, and each will be unaware of any other user. Immediate response also enables the user to conduct a continous dialogue with the machine allowing the computer to become, in a sense, an intellectual assistant to the creative effort, rather than a passive tool.

Transmission of Computer Data

The burden of increasing use of teleprocessing systems falls squarely on communication common carriers. Loading of circuits is result of:

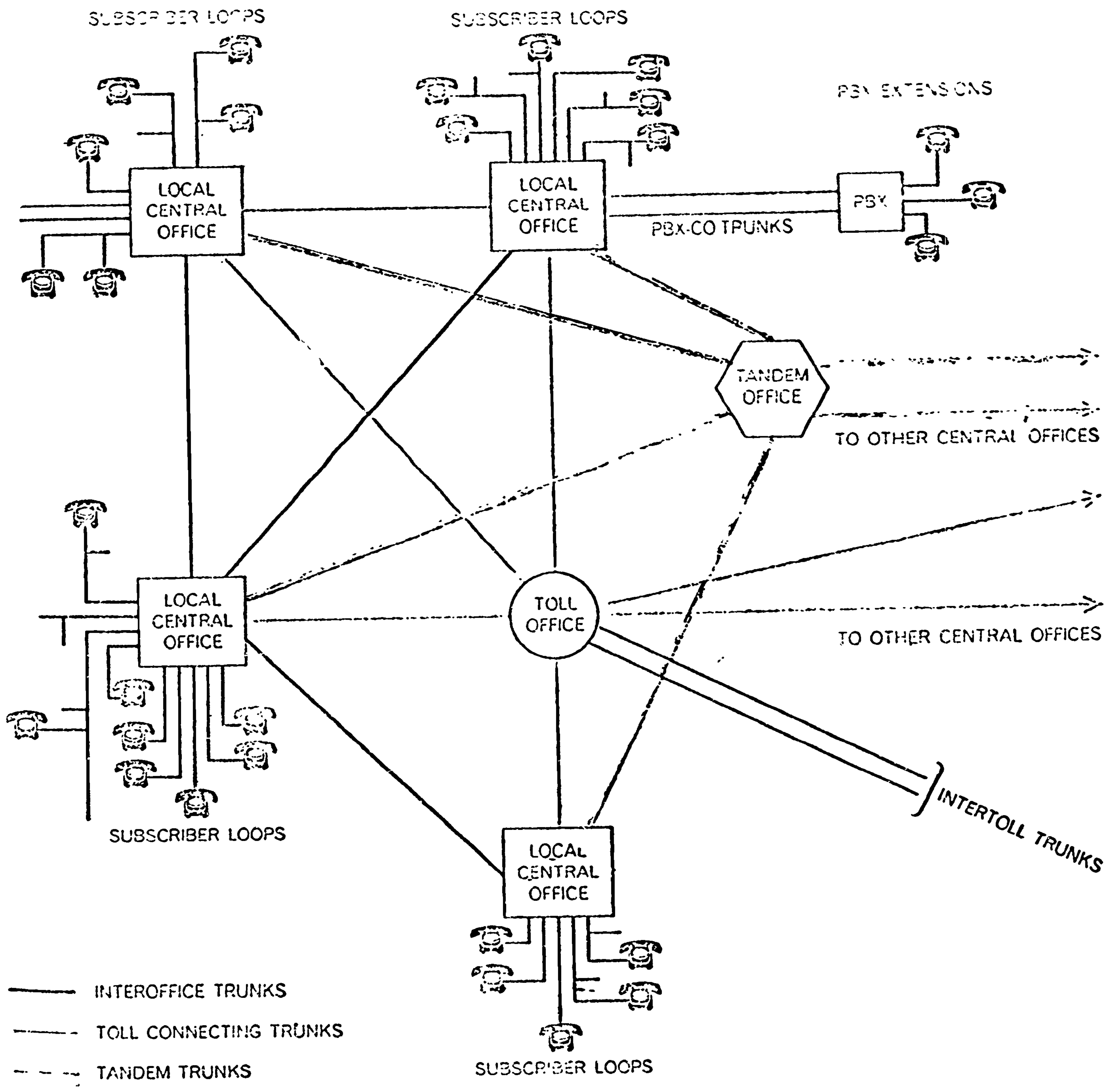
a) Long length of compute "calls" - existing circuits are designed with a capacity for a fraction of all telephones at once.

b) A high-quality of transmission is needed; the human ear does not mind if some components of a signal arrive slightly earlier or later, but for computer each bit must be exact.

c) Exact control of amplitudes and static is required; a message switching system which is part of a common-carrier circuit means making and breaking many connections during the course of transmission.

Telephone switching circuits - briefly explain telephone switching circuits.

The entire system is a vast computer switching system; for example, a coast to coast call: no human can determine the exact circuit route used. The job could not be handled by human operators given today's volume.


SUBSCRIBER LOOPS
SUBSCRIBER LOOPS
PBX EXTENSIONS
LOCAL CENTRAL OFFICE
LOCAL CENTRAL OFFICE
PBX
PBX-CO TRUNKS
TANDEM OFFICE
TO OTHER CENTRAL OFFICES
LOCAL CENTRAL OFFICE
TOLL OFFICE
TO OTHER CENTRAL OFFICES
SUBSCRIBER LOOPS
INTERTOLL TRUNKS
LOCAL CENTRAL OFFICE
INTEROFFICE TRUNKS
TOLL CONNECTING TRUNKS
TANDEM TRUNKS
SUBSCRIBER LOOPS

Another problem is the regulation of new uses of common carrier lines. When transmitting digitally coded data, graphs, or other computer output, how much a part of the computer system are the transmitting and switching circuits? Existing regulatory bodies are currently faced with problems of anti-trust regulations as well as the regulation of common carrier transmission. The question to be decided is what relation does the computer manufacturer have with the communication industry? This poses a threat to progress in this field. The future will be governed by economic questions as well as by needs of computer users. "Information" transmitted to a computer and back from a time-sharing station may be considered data and not truly information. It is, therefore, economical to transmit only as long as communication channels are cheaper than the computing machinery needed to generate the data.

Use of Time-Sharing, Teleprocessing Systems

Basically there are three broad functional classes.

Inquiry Systems are basically simple processing systems which will also answer questions of limited variety. Example: savings bank installation. The only justification for a time-sharing system is that tellers at the windows must be able to get a depositor's account balance on demand. All other processing, such as interest calculation, updating, etc., can be performed satisfactorily by sequential processing methods.

Dispatching Systems

This type of system has active rather than passive functions. Example: inventory control system. It features large random access storage. Since a variety of functions may be performed at random times, it must be more sophisticated than an inquiry system.

Most dispatching systems are used for some type of inventory control, though "inventory" may be interceptor aircraft or available shipping space. Also, most existing military command and control systems fall into this class. Two examples are the Semi-Automatic Ground Environment (SAGE) air-defense system, and the control system for our manned space flights.

Decision-Making Systems are similar to dispatching systems; the main difference is the method used. A dispatching system uses pre-established decision rules automatically, while a decision-making system finds an optimal answer to every demand. It generally requires a great amount of computing power in relation to other types of systems. Examples of such systems are allocation of resources in telephone company, determining production schedules in a manufactoring plant, controlling operation of a chemical plant. Massive savings are possible for nearly any organization which has the problem of applying complex resources to a multitude of demands.

The Computer Utility

All in all, mass memories combined with time-sharing and teleprocessing systems are becoming more and more like a community library. It is easy to envision an entire business organization making and executing all of its major decisions with the aid of time-shared computing system. Existing systems of computer aided instruction are evidence of the available capability.

In the future, computer utilities will play an increasingly larger role in human affairs. The coupling of such computer systems and human users will create new services, new institutions, and a new environment. Because such a system, with massive data banks, would bind members of a community more closely together, the problems will probably be ethical ones. Our current problems with wiretapping and other invasions of privacy suggest the seriousness of the problem of such a system which may hold in its memory detailed information on individuals and organizations. How will access be controlled? Who will regulate its use? What safeguards can be devised to prevent its misuse?

Problems of the future will be social as well as technical ones.

SELECTED READINGS

BOOKS

It is no longer true, as it once was, that anything published about computers was necessarily already out-of-date. The intent of this bibliography is not to cover the field but to present a few references which are generally clear and meaningful as well as being readable.

1. Fowler, John M. The Computer in Physics Instruction. Irvine, Calif.: Commission on College Physics, 1966.
   The Fowler monograph, though aimed at a single field, is extremely helpful in that it explains general features of computers which will enlighten the reader interested in their instructional uses.

2. Simon, Herbert A. The Shape of Automation. New York: Harper & Row, 1965.
   One of the finest discussions of the broad implication of computers and automation. It explores, primarily in the context of administration, what computers can be expected to do and their role in decision-making.

3. Smith, Paul T. How to Live with Your Computer. New York, American Management Association, 1965.
   Interesting to and appropriate for administrators.

4. Lickliker, J.C.R. Libraries of the Future. Cambridge, Mass.: The M.I.T. Press, 1965.
   May read like science fiction but it bases itself firmly in fact. It gives a mind-stretching picture of the future of the research library as the intellectual heart of the university.

5. Bernstein, Jeremy. The Analytical Engine. New York: Random House, 1964.
   Lucid and engaging introduction to computers--their history, their operations, their implications.

6. Desmonde, William H. Computers and Their Uses. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1964.

7. Berkeley, Edmund C. The Computer Revolution. New York: Doubleday & Co., 1962.

This author talks of computers as electronic brains. Since debate on this particular subject is continuing, one wishes he would avoid statements that computers are machines that think. However, this book contains a good discussion of workings and uses.

8. Dunlop, John C., Editor. Automation and Technological Change. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962.

A number of well-known figures in American industry and education have contributed articles that make up the collection in which the various aspects of automation in our society are discussed. The contributions were designed to eliminate some of the exaggeration, suspicion, and mystery that have surrounded the subject.

9. Englebardt, Stanley L. Computers. New York: Pyramid Publications, Inc., 1962.

Well written with emphasis upon the users of the computer. All examples are chosen well, but there is no attempt at explaining what the computer is.

10. Greenberger, Martin, ed. Computers and the World of the Future. Cambridge, Mass.: The M.I.T. Press, 1962.

A number of interesting papers appear concerning many aspects of computing.

11. Haldacy, D.S., Jr. Computers--The Machine We Think With. New York: Harper & Row, 1962.

12. Vorwald, Alan and Clark, Frank. Computers! From Sand Table to Electronic Brain. New York: McGraw-Hill Book Co., 1961.

Starting with the first finger counter, this book traces the history of computers, explains number systems, and finally shows the application of electricity to these systems in detail. For practical science experience, there are instructions, with many diagrams and drawings, for building "computers" from an ancient decimal abacus to an electrically powered binary counter.

13. Buckingham, Walter. Automation--Impact on Business and People. New York: New American Library of World Literature, Inc., 1961.

Dr. Buckingham, an economist, discusses the impact of automation, the role of computers as part of the control system in much automation, and how automation will affect all people in our society. His approach is clear, well paced, and realistic.

14. Booth, Andrew D. Automation and Computing. New York: Macmillan Co., 1959.
This is a survey of electronic computers and automation written for the non-specialist or newcomer to the field.

15. Fahnestock, James D. Computers and How They Work. New York: A.S. Barnes & Co., 1959.
This is a comprehensive, relatively easy-to-understand book describing the principles behind, and the manufacture, operation, and maintenance of computers. Building from a basic foundation of mathematics and electricity, the author explains how some complex machines work.

16. Chapin, Ned. An Introduction to Automatic Computers. Princeton, N.J.: D. Van Nostrand Co., 1957.
An excellent, thorough discussion of the ins and outs of computers. Each chapter is followed by a list of suggested readings.

17. Berkeley, Edmund C., and Wainwright, Lawrence. Computers: Their Operation and Applications. New York: Reinhold Publishing Corp., 1965.
Developments in both the techniques and equipment of automatic computing are described, illustrated, and compared. The book includes information on the function of computers, how they work and what they can do.

18. Gorn, Saul and Manheimer, Wallace. The Electronic Brain and What It can Do. Chicago: Science Research Associates, 1956.
An interesting treatment of computers. Although the title is misleading, the authors present some of the fundamental ideas about digital computers, how scientists simplify and code difficult problems for computers, and who uses computers, as well as a view of implications.

ARTICLES

Many articles on computers have appeared during recent years in various magazines. Because of the publication's reputation, clear presentation, and availability to students, selections were made only from the Scientific American. No annotations are offered, as most titles are indicative of the content of the accompanying articles. The first twelve articles from the September 1966 Scientific American are available in reprint in book form as Scientific American, A Scientific American Book of Information (New York: Scientific American, 1966.)

1. John McCarthy. "Information," 215: 64-73; September 1966.

2. David C. Evans. "Computer Logic and Memory," 215: 74-85; September, 1966.

3. Ivan E. Sutherland. "Computer Inputs and Outputs," 215: 86-111, September 1966.

4. Christopher Strachey. "System Analysis and Programming," 215: 112-127; September 1966.

5. R.M. Fano and F.J. Corbato. "Time-Sharing on Computers," 215: 128-143; September 1966.

6. John R. Pierce. "The Transmission of Computer Data," 215: 144-159; September 1966.

7. Anthony G. Oettinger. "The Uses of Computer in Science", 215: 160-175; September 1966.

8. Steven Anson Coons. "The Uses of Computers in Technology," 215: 176-191; September 1966.

9. Martin Greenberger. "The Uses of Computers in Organizations," 215: 192-205; September 1966.

10. Patrick Suppes. "The Uses of Computers in Education, 215: 206-223; September 1966.

11. Ben-Ami Lipetz. "Information Storage and Retrieval," 215: 224-245; September 1966.

12. Marvin L. Minsky. "Artificial Intelligence," 215: 246-...

13. Don L. Bunker. "Computer Experiments in Chemistry," 211: 100-108; July 1964.

14. "Science and the Citizen--The Information Explosion," 209: 72-73; November 1963.

15. C.L. Stong. "The Amateur Scientist--A simple analogue computer that simulates Pavlov's dogs," 208: 159-166; June 1963.

16. Victor H. Yngve. "Computer Programs for Translation," 207: 68-87; June 1962.

17. Martin Gardner. "Mathematical Gamer--How to build a game-learning machine and then teach it to play and to win." 206: 138-144; March 1962.

18. W.B. Ittner and C.J. Kraus. "Supersonducting Computers," 205: 124-37; July 1961.

PREFACE

A prime responsibility of a professional society is to foster continuing education activities covering new developments in topics of importance to the work of its members. This is particularly true in rapidly expanding and highly complex technologies such as those in the field of information science.

With this view, the 1967 Conference Planning Committee of the American Documentation Institute (now the American Society for Information Science) chaired by Paul Fasana of the Columbia University Libraries, established a Tutorial Subcommittee to organize training sessions for presentation at the Conference. The Subcommittee, under the direction of Russell Shank, then Associate Professor at the Columbia University School of Library Service, agreed to develop three workshop tutorials for the following areas: elements of information systems; electronic data processing concepts; and generalized programming languages and systems.

The tutorials on these topics ran concurrently. Each began with a general session in which the tutorial leader gave an overview of the topic to be covered. The participants were then formed into smaller workshop or seminar groups for detailed instruction by a team of tutors. Each tutorial lasted the entire day. The general sessions were limited to about 100 people; seminar groups were limited to about 25 people. In the seminar groups each of the tutors either covered the entire topic simultaneously, or presented a part of the information to be covered, repeating their presentations as groups were rotated among them.