ENGINEERING
AND SCIENCE
May 1965, Vol. XXVIII, No. 8

Caltech's Developing Computer System

by G. D. McCann

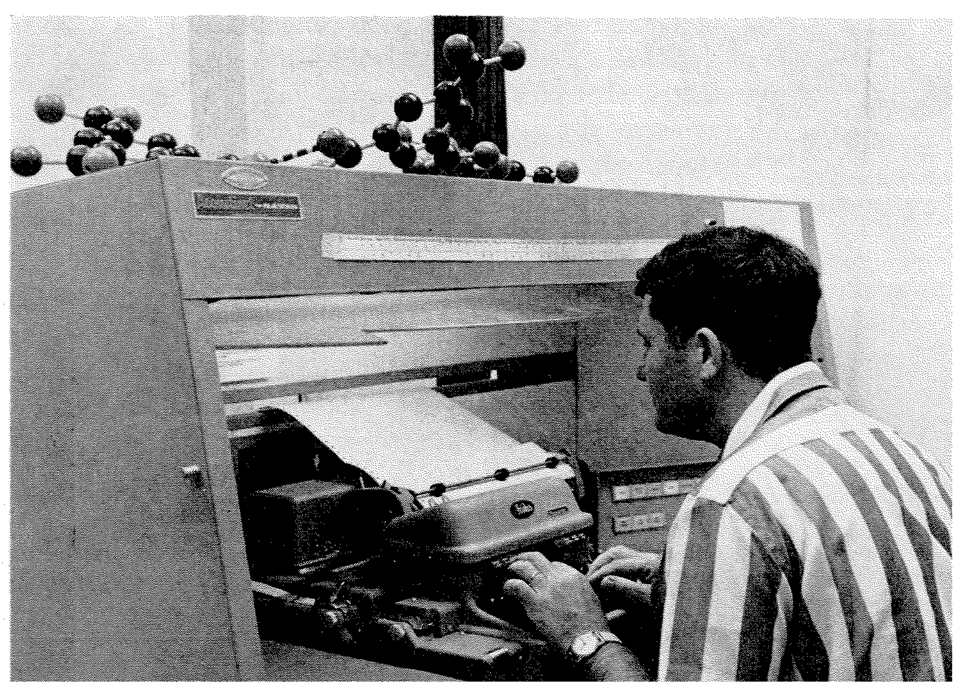
We define as "intelligent thought" all of man's efforts to derive orderly concepts or laws from observations of himself and the world around him, by focusing attention successively on small segments of the total body of information, and attempting to enlarge the field of his concepts by correlating these studies.

In certain areas of scientific investigation, such efforts have resulted in the development of formal mathematics, with their concise algebraic languages. In many fields, however – and most notably in the life and behavioral sciences – man has been forced to grope through complex masses of information without any suitable way of reducing his observations to orderly concepts.

From World War II technology has come a tool which adds a new dimension to this quest for intelligent understanding – the general purpose digital computer. This computer has two principal components. One component is a data processor which, with a fixed number of relatively simple algebraic algorithms, can simulate any intelligent information-processing function by the proper serial combination of these algorithms. The other component is a complex hierarchy of memory structures capable of storing vast amounts of information, which can be recalled and reprocessed in an infinite variety of ways. Because the rules or programs for this serial series of data-processing can also be stored in memory and processed, it is possible for the system to modify its own programs and thus simulate learning – an essential property of true intelligence.

Digital computers were first applied effectively only to formal mathematics – arithmetic, algebra, calculus, and stochastic processes. However, during the past ten years significant progress has been made in their use to simulate non-formal information processing – such as the modeling of game playing, problem solving, and theorem proving. These efforts have resulted in the development of important new concepts of information storage and retrieval and of information manipulation. Such computer systems can be particularly useful in our studies of actual living nervous systems.

The basic concepts emerging from the formalization of modern information-processing tech-



Richard Marsh, senior research fellow in chemistry, communicates with the computer from a remote users console in the chemistry laboratories. A crystallographer, Marsh uses the computer in studies to determine molecular structure by means of x-ray diffraction. There are four remote users consoles on campus now; next fall there will be 25; and within two years there may be as many as 100.

Stewart Smith, associate professor of geophysics, feeds data into a remote console connecting the computer and the Seismological Laboratory. The computer aids in determining the magnitude and location of earthquakes.

niques reach directly into the fundamental meaning of information and the precise mathematical structures of languages as related to both syntax and semantics. The precise description of language structures, and the simulation of a wide variety of languages on digital computers, provides material for research and instruction in the information sciences. Expansion of the concepts of the formal algebraic languages of mathematics is the role of modern applied computer mathematics. These concepts are also producing more precise and sophisticated methods for understanding and simulating natural or conversational languages on computers. Such methods lead directly to the simulation of creative thought processes as a mathematical basis for psycholinguistics and other areas of experimental psychology.

A new form of language structure suitable for describing and understanding experimental data in all fields of science is also emerging in the computer research at Caltech. This type of language, which also possesses true syntactic structure, is being applied to research on sight-sensory nervous systems.

Suitable computer systems now offer the possibility of completely new techniques for experimental research. This promises to be particularly important in all areas of living nervous system research, where it is being applied extensively at Caltech in work ranging from molecular studies, through neural network analysis, to studies of thought processes at the natural-language cognitive-process level.

Computers can play two roles in this activity. They permit completely new methods of modeling theories, even in areas where formal mathematics has not been applicable. They also play a perhaps even more important role by greatly extending the complexity of experimental research by interacting directly in a complex experimental environment, performing such tasks as (1) controlling complex experimental functions, (2) recording and documenting complex measurements, (3) providing rapid preliminary analyses of results so as to facilitate the progress of the experiment, and (4) performing sophisticated analyses for conceptual interpretations.

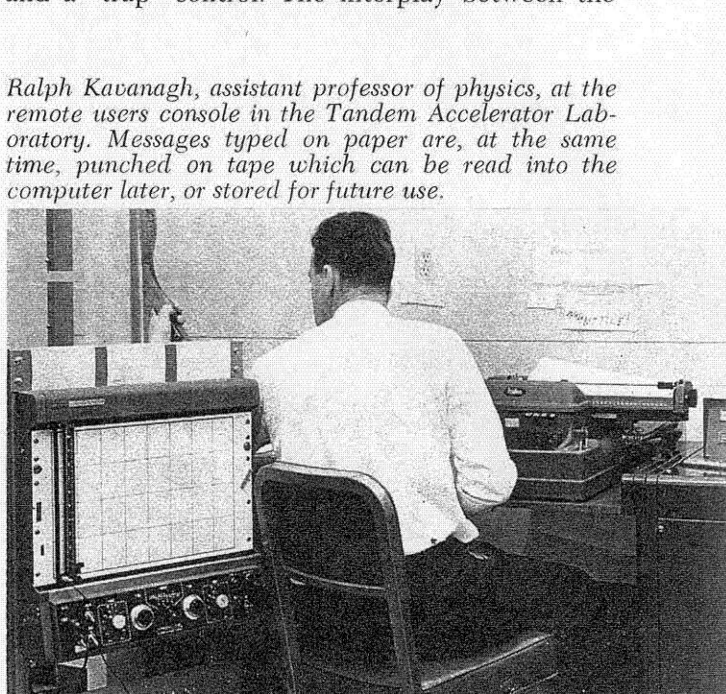
When the computer system that is now installed in Caltech's Willis H. Booth Computing Center was under development, the emphasis in computer usage was on the features of high data-processing speed and large fast-access memory. Of equal importance, however, is the enhancement of communication between the user and the computer, in terms of rapid simultaneous communication for several users, or rapid "turn-around" times for a large number of diversified uses.

To produce new methods of communication and uses, the Computing Center staff developed the "shared-file" system concept, illustrated in the diagram on the opposite page. This is actually one integrated system of three computers and a variety of communication-control and input-output devices – some at remote locations.

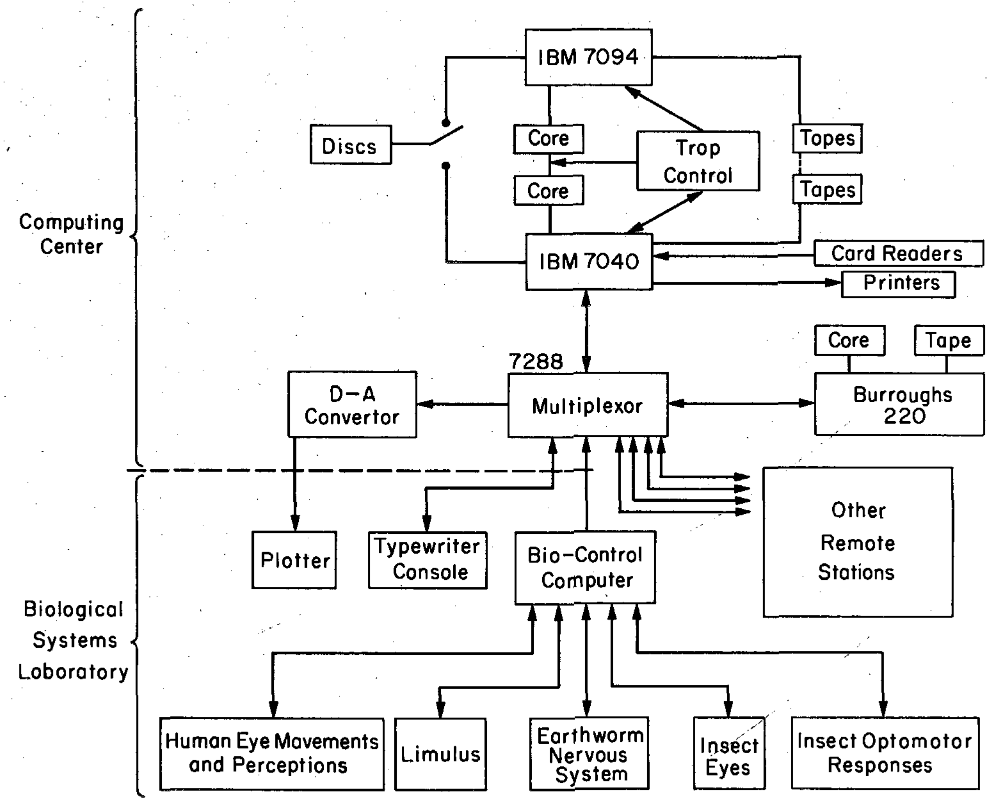
The three main computers are the IBM 7094, the IBM 7040, and the Burroughs 220. Simultaneous communication with the system effected through the 7288 multipelxor, a special-purpose core-memory to hold messages temporarily, or to keep data long enough for it to be accepted by the IBM 7040 as imput, or to be received from the 7040 by the various output devices.

The IBM 7040 and the IBM 7094 work together as a system through three memory interconnections and a "trap" control. The interplay between the

Ralph Kavanagh, assistant professor of physics, at the remote users console in the Tandem Accelerator Laboratory. Messages typed on paper are, at the same time, punched on tape which can be read into the computer later, or stored for future use.



PRESENT COMPUTER SYSTEM



two computers is achieved by four interconnections; the computers are interconnected through 25 million words of tape memory, 18.5 million words of disc memory, 64,000 words of core memory, and an instantaneous trap control.

In a typical application, the IBM 7040 will set up the proper programs and data for each new problem in the disc memory, periodically interrupting the IBM 7094 to send it (via the core memory) an updated disc map and priority listing. As the IBM 7094 finishes each job, it stores the output in the disc memory and interrupts the IBM 7040 with instructions for the removal of the output data.

It is important to note that this system concept was developed largely through a cooperative effort with major departments of the Institute interested in computer applications. In the original planning, three major modes of system usage were envisioned:

1. Collection and analysis of direct experimental data by major research facilities.
2. Rapid production computing using pre-developed programs stored in the main disc memory.
3. New program development through the use of compilers. These are programs that translate problem descriptions from some form of formal algebraic language into the machine language of the computers.

Some months after the system had been placed in operation, a fourth mode of interaction between people and the system was implemented. This is called QUIKTRAN and was originally developed by the IBM Corporation.

This mode of operation enables the human user and the computer to work together on the problem of creating a useful program. A remote typewriter console is the medium of communication, and many such stations can be used simultaneously. In its simplest application the remote typewriter can be used as a desk calculator. By typing in a statement asking for a function (such as a cube root or cosine function), together with the number to be used, one obtains the desired function. In its more sophisticated mode of operation, QUIKTRAN permits one to type in any series of algebraic statements that might form a given problem program.

The computer checks each statement for format consistency, looks for omission or illogical statements, and immediately types back its comments. The user can then ask the computer to carry out the statements, either singly or in groups, with certain trial data to test it for correctness, and can quickly correct any errors found by the computer in the process. Solutions can then be run off. Higher-order "conversational" modes are now under development by the Computing Center staff.

The progress which the Institute has made in the application of the present computer system has induced the IBM Corporation to make a completely new and much more advanced system available to the Institute. This system, scheduled for installation in October 1965, is illustrated below. For convenience, the diagram is divided into four sections. Section A, the central computer system, represents the expansion of the shared file shown on p. 9. Here the IBM 7040-7094 system is replaced by a 360/50 and probably a 360/70. The shared file becomes a high-speed bulk core with a random access time of 8 microseconds instead of the 7 milliseconds of the present disc memory.

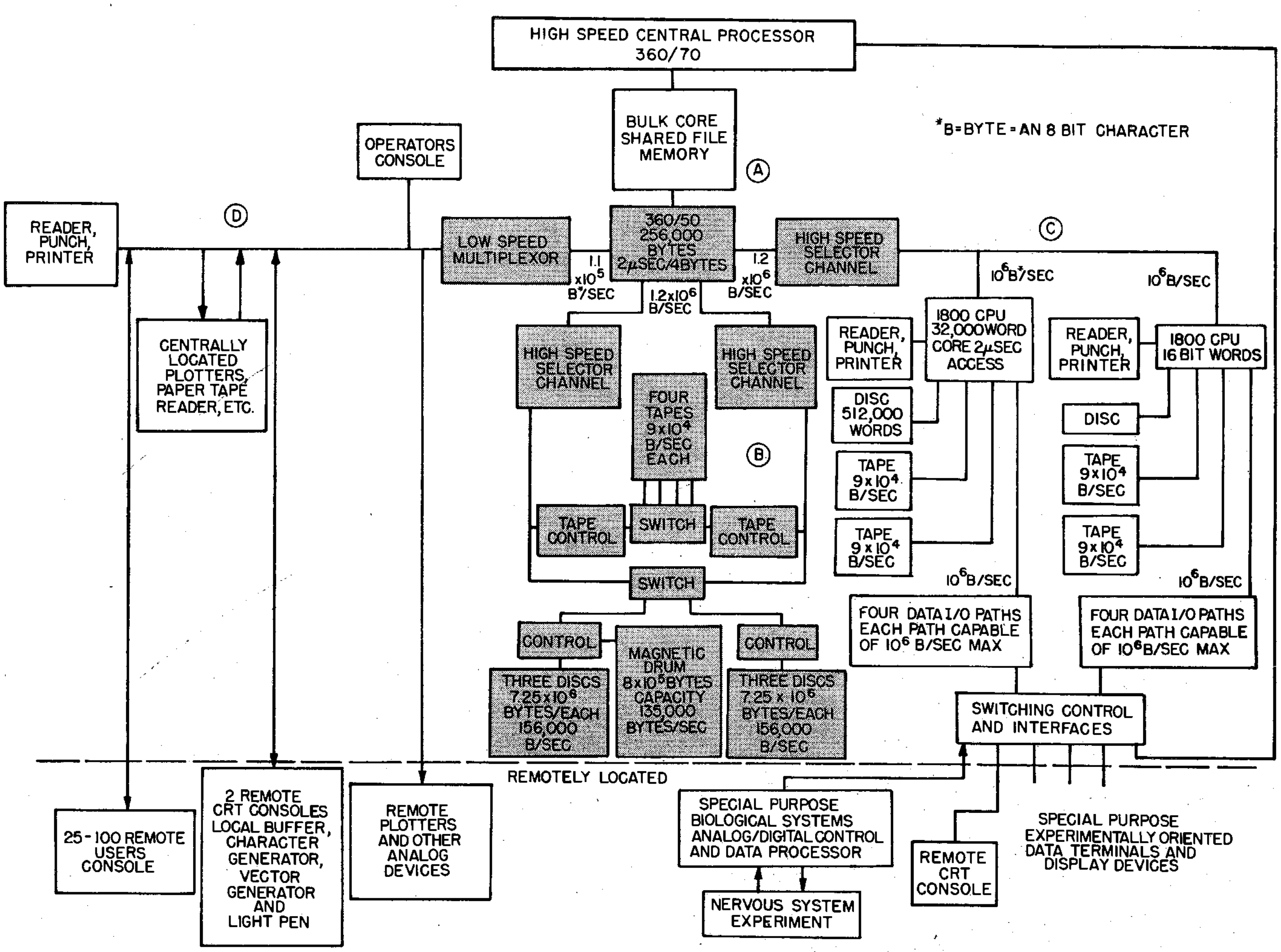
Section B (the shaded areas) functions as (1) the primary bulk library storage, and (2) the expanded form of algebraic conversational modes.

Section C shows a series of two 180 digital computers designed for experiment data collection and "on-line" preprocessing of experimental data.

Section D shows the large number of remote typewriter consoles and cathode ray display devices which are being provided for a variety of uses, including formal course instruction.

In the present system, the QUIKTRAN mode must be operated independently of other system modes. It is presently being run one hour per day. In the new system, the new form of algebraic conversational mode will actually be interflexed with the other modes in which the system can be used. Furthermore, there will be a generalized "translator" in the system which will enable a user to specify any consistent algebraic language to the computer. Once this has been done, the computer, having immediately learned this new language, works with the user, helping him to write and execute his program in his own language.

The very large bulk core memory of the new system also permits the development of a higher order of conversational modes or methods by which the computer and humans can work together on high-order thought processes. One form of these modes makes use of such natural conversational languages as English, and is suitable for the simulation of cognitive processes and research in linguistics and psychology.



Caltech's new computing system, to be installed in October 1965.