The Seeds of Artificial Intelligence

U.S. DEPARTMENT OF HEALTH,
EDUCATION, and WELFARE
Public Health Service
National Institutes of Health

SUMEX-AIM

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402

The Seeds of Artificial Intelligence

SUMEX-AIM

A publication of
Division of Research Resources
National Institutes of Health
Bethesda, Maryland 20205

Prepared by
Research Resources Information
Center
1776 East Jefferson Street
Rockville, Maryland 20852
under contract N01-RR-9-2114

March 1980

U.S. DEPARTMENT OF HEALTH,
EDUCATION, and WELFARE
Public Health Service
National Institutes of Health

NIH Publication No. 80-2071

This publication was written by
Gregory Freiherr

Acknowledgements

The efforts of many people made this publication possible. Special thanks go to Mr. Thomas Rindfleisch; Drs. Joshua Lederberg, Herbert Simon, Edward Feigenbaum, Bruce Buchanan, Paul Armand, William Baker, Jack Myers, and Harry Pople; and Mr. Edward Post.

Foreword

In the past century, science has not only changed our conceptions about the world, it has changed itself. Driven by an explosion of information, specialties in science have sprung up, inevitably giving rise to subspecialties. But staying abreast of new knowledge, even in narrowly specialized areas, is becoming increasingly difficult. One way to manage the continuing flood of new information may be to create entities of intelligence.

The proposed tool is the intelligent machine, a device that mimics the expert's reasoning power and can retain in retrievable form much of the knowledge currently available to experts in a given specialty. Most systems of this type are still immature. But some are already moving into the real world and others will make the transition within the next few years. As these activities become more formalized, a new branch of applied science will arise. Most likely it will be called knowledge engineering.

What systems will be available? Who will they help? How will they work?

Many answers are contained in existing books and articles. But technical publications suffer from the defect of their virtues. They are too detailed, too exhaustive and, most important, too focused on single areas of rapidly expanding disciplines. To understand this new branch of computer science, called artificial intelligence (AI), it is necessary to understand the foundation, the broader base, on which it rests.

This publication will present a general view of AI, the concepts from which it evolved, its current abilities, and its promise for research. The focus is on a community of projects that use the SUMEX-AIM (Stanford University Medical Experimental Computer for Artificial Intelligence in Medicine) network.

SUMEX-AIM is a nationally shared computing resource devoted entirely to designing AI applications for the biomedical sciences. It is funded by the NIH Division of Research Resources, Biotechnology Resources Program. Although SUMEX-AIM does not include all AI projects directed toward medicine and related research in this country, many of the programs now using AI techniques for medical decision-making were developed using this facility.

Table of Contents

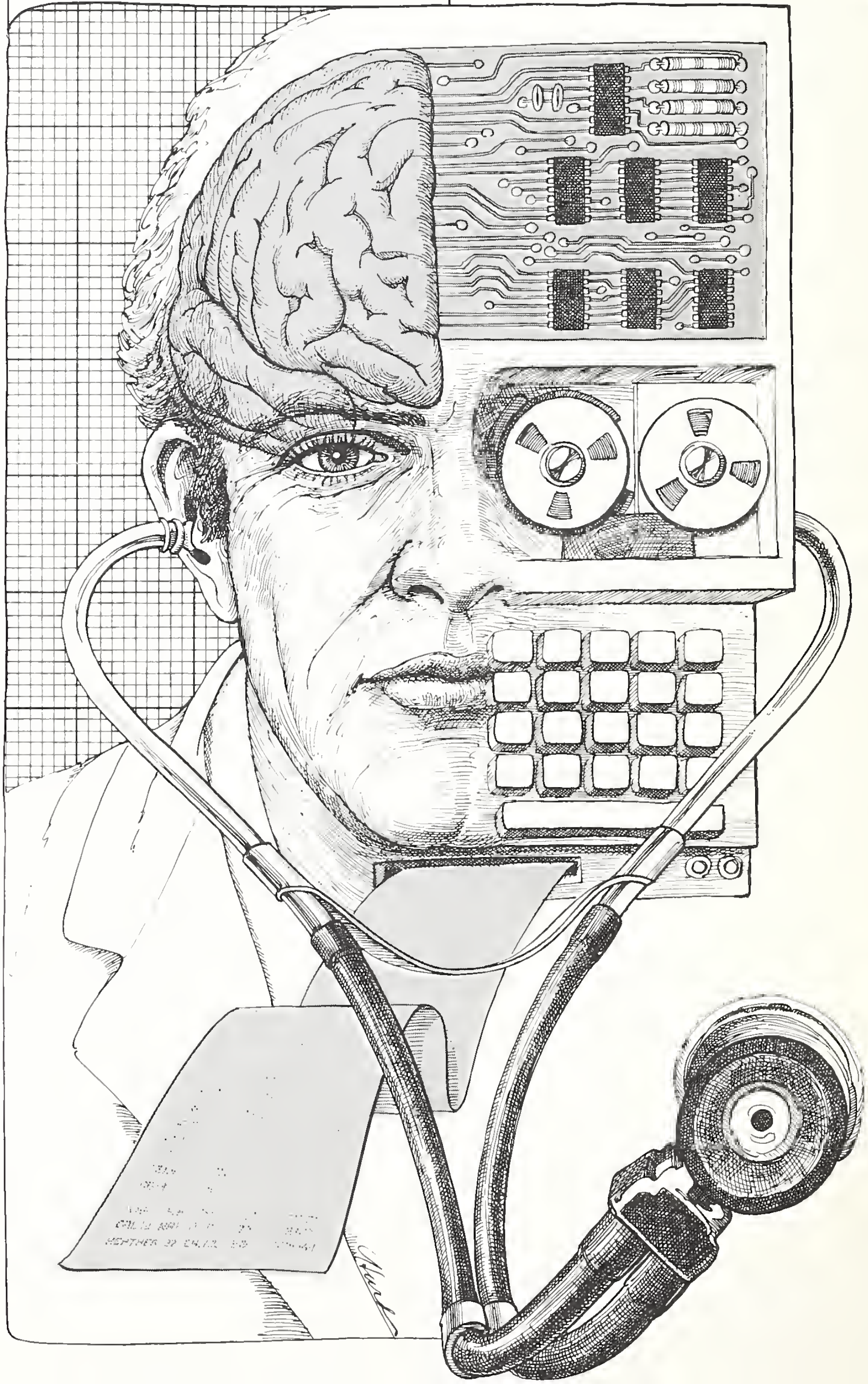
Introduction

Artificial Intelligence— What's in a Name?

For centuries, philosophers and linguists have grappled with the question of defining intelligence. Most have approached the issue by describing the function of intelligence, or the way it appears in behavior. An exact definition for this term is elusive.

As might be expected, machine intelligence is equally, if not more, difficult to define. According to Dr. Margaret Boden in her book *Artificial Intelligence and Natural Man,* computers are only research tools, machines programmed to do things that would require intelligence if done by people. Dr. Marvin L. Minsky, artificial intelligence (AI) researcher at the Massachusetts Institute of Technology and advisor for SUMEX-AIM, agrees. He says artificial intelligence is the science of making machines do things that people need intelligence to do.

Others take a somewhat differ-



ent view. Dr. Edward Feigenbaum, principal investigator of SUMEX-AIM, says the field is not primarily oriented toward technology, but toward investigating the nature of intelligence as information processing, whether the intelligence is expressed by man or machine.

One point of emphasis in current AI research is to design computer programs that capture the knowledge and reasoning processes of highly intelligent specialists. The practical goal of such work is to make specialized expertise more generally accessible. To do so, researchers are attempting to understand how experts go about acquiring and using knowledge. Principles of how knowledge accrues and how it is retrieved in logical sequence are extracted. They are then programmed into the computer.

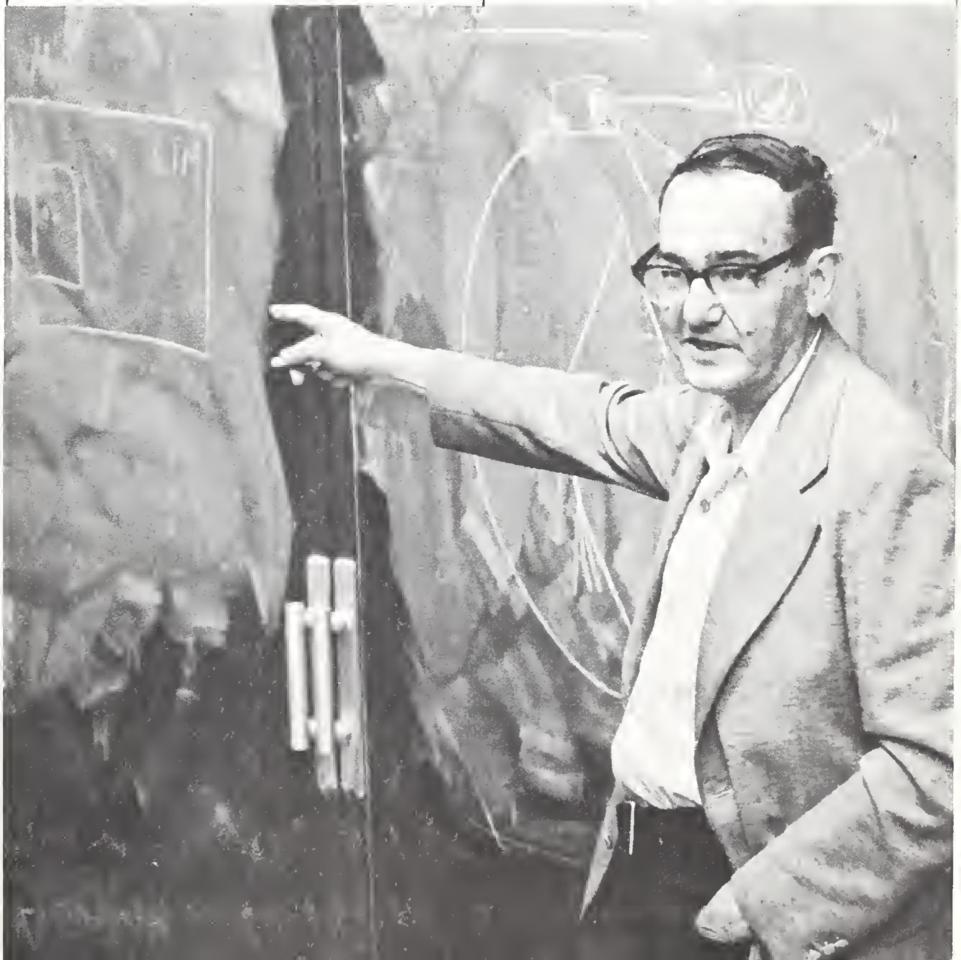
Within the SUMEX-AIM system, the reasoning processes of physicians, chemists, and other biomedical scientists are being analyzed. At present, the ability of most programs is limited and much less flexible than the corresponding human intellect. In specialized areas of medical diagnosis and chemical structure analysis, some programs can already rival human capabilities. Still, many people are skeptical of the computer's potential.

Nobel Prize winner Dr. Herbert A. Simon, psychologist-computer scientist at Carnegie-Mellon University and SUMEX-AIM advisor, is convinced that this potential is generally underrated. He says human behavior is based on a complex but definite set of laws. If these laws are discovered and reduced to computer software, Dr. Simon believes machine intelligence comparable to man's will become a certainty in specific areas of expertise.

To capture these higher level functions, AI researchers are developing a new approach. It is called symbolic computation, a set of methods by which abstractions can be expressed and managed in the computer to solve non-mathematical problems. They emphasize manipulations of symbolic rather than numeric information, and they use largely informal or heuristic decision-making rules gained from real-world experience.

When used in AI, heuristics focus the program's attention on those parts of the problem that are most critical and those parts of the knowledge base that are most relevant. The result is that these programs pursue a line of reasoning, rather than a sequence of arithmetic steps.

Use of complex symbolic structures is necessary when constructing computer applications for domains that cannot be well-formulated in mathematical terms —either because they are not fully understood, as in medical diagnosis, or because the underlying concepts are intrinsically non-numerical. "Seldom are there equations, in the mathematical sense, that relate measurements of body parameters to the diagnosis of disease," says Mr. Thomas Rindfleisch, director of the SUMEX computing facility. "Rather, the process of diagnosis is characterized by a set of strategies having to do with rules of experience and judgmental knowledge. These rules govern the interpretation of observations and guide decisions

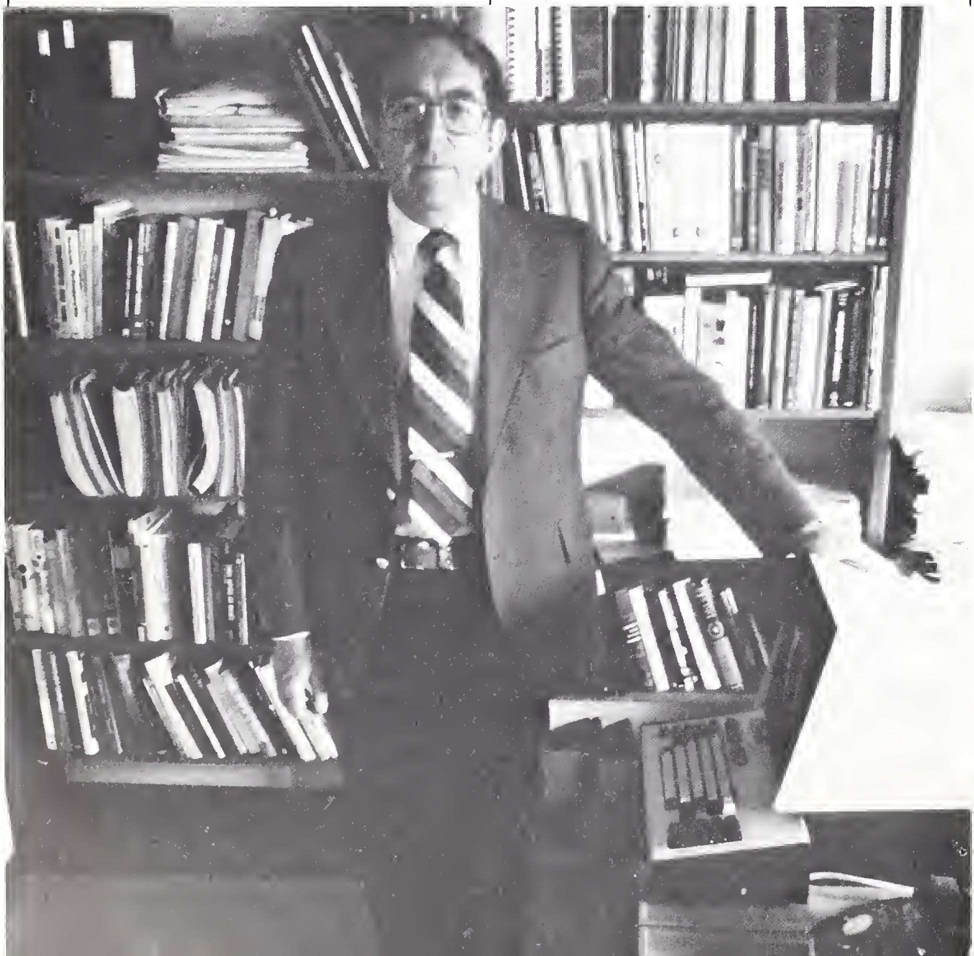


Dr. Herbert A. Simon, SUMEX-AIM advisor: sorting out the recipe of intelligence.

about what other information is needed to determine the disease process involved."

For example, INTERNIST, a diagnostic computer program in the SUMEX-AIM network, is focused on the broadest of medical specialties—internal medicine. It analyzes patient cases by mimicking the expert's reasoning process. "The method used by physicians to arrive at diagnoses requires complex information processing which bears little resemblance to the statistical manipulations of most computer-based systems," says Dr. Jack D. Myers, coprincipal investigator of the project at the University of Pittsburgh. "As a result, the focus of research in this field of medical applications has shifted during the past few years from models of statistical inference to those using the heuristics of artificial intelligence."

"In final form, INTERNIST will amplify intelligence," Dr. Feigenbaum says. It will supply expert advice to the general practitioner and physician's assistant, accelerating and improving their work. "An equally important outcome of research such as this at SUMEX-AIM is eliciting, organizing, and polishing a body of knowledge that rarely sees the light of day," he says. "It is the knowledge that underlies the expertise of practice, the knowledge that is normally transmitted by a kind of osmosis process from master to apprentice. That knowledge will now be codified, taught, used, and critiqued." In essence then, a key goal of artificial intelligence research in the SUMEX-AIM community is to capture in computer programs the knowledge and problem-solving abilities of experts. After studying this process in many specialized areas of expertise, Mr. Rindfleisch says, it may ultimately be possible to capture in computer programs something of the process of creativity and discovery itself. Programs then would possess the ability to detect patterns that establish order from chaos, to draw connections between seemingly unrelated ideas, and to establish the principles for solutions to new classes of problems.



—Photo by Andy Williams

Dr. Edward Feigenbaum, principal investigator of SUMEX-AIM: "The laws of expertise will be taught, used, and critiqued."

History of Computing

Abacus to ENIAC—and Beyond

In the millennium before Christ, amid the great cities and conquests of Greece and Rome, dreamers and theorists were laying the groundwork for today's thinking machines. Like seed crystals in a supersaturated solution, these visionaries drew from nature, assembling conclusions from observations about the universe. Their efforts brought important advances in mathematics, astronomy, and medicine.

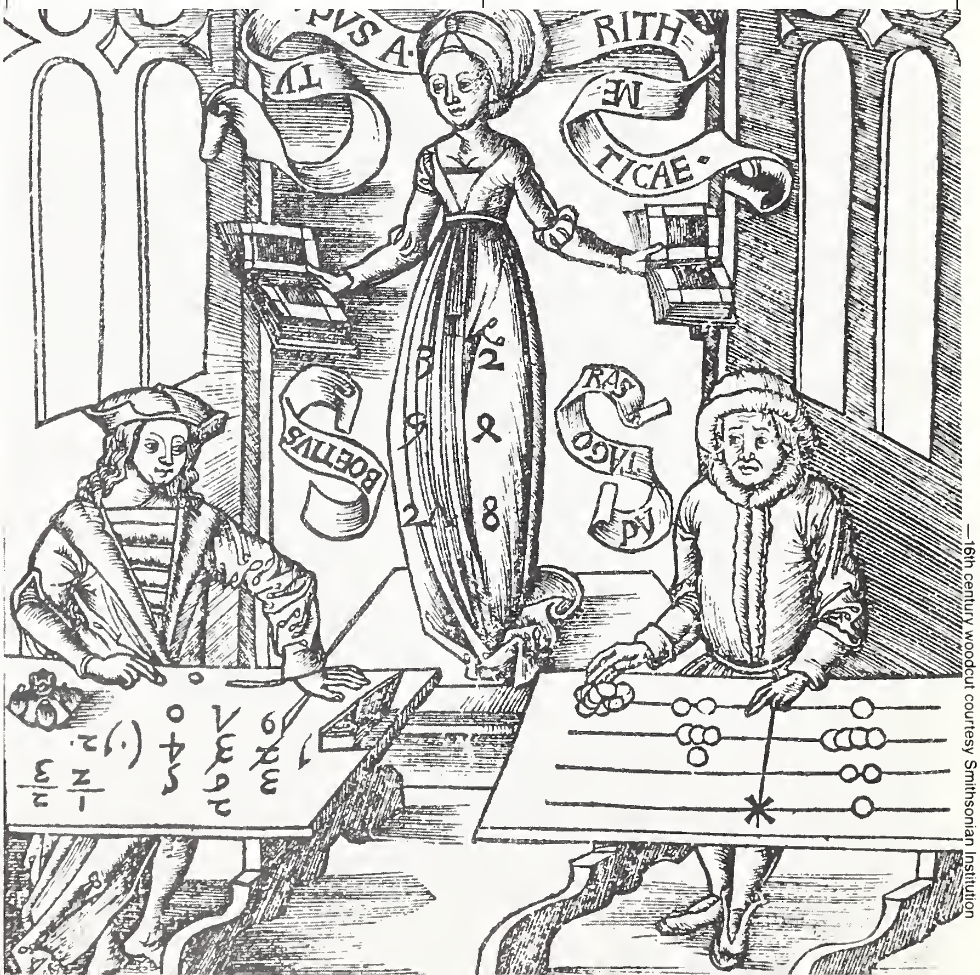
Much of the early work in formulating the laws of mathematics may appear to have little connection with the computer of today. But each step forward in this elaborate science was indispensable to the ultimate arrival of the computer.

Pythagoras, a 5th century B.C. philosopher known as the founder of Greek mathematics, was the harbinger. He first described the "mystical significance of numbers" and established the relationship between musical harmony and mathematics.

Perceiving in the skies a regularity similar to that of music, Pythagoras studied movements of the heavenly bodies, or as he called it, "the music of the spheres." He became the first to realize the importance of geometric shape, which governs all nature from crystalline rock to the human body. In so doing, Pythagoras set the direction of mathematical thought for centuries to come.

Mathematics was soon regarded as *exact.* It became the cornerstone of all science. For centuries scholars believed that its logic was infallible. But in the 19th century the first inklings of doubt surfaced. Two mathematicians, one in Hungary and the other in Russia, established irrefutably that it was impossible to prove Euclid's postulate of parallels, which states that no more than one line parallel to a given straight line can pass through a given point. Alternate theories sprang up, threatening to scatter the focus of science. Mathematics, the mainstay of scientific certainty, was suddenly uncertain. If there are two or more geometries, which is right?

After much thought and deliberation, Jules Henri Poincaré, a 19th century philosopher and mathematician, found the solution. He



Boethius (left) and Pythagoras: a fanciful battle between arithmetic calculation and the abacus.

—16th century woodcut courtesy Smithsonian Institution answered simply that the question is meaningless. Poincaré, described as one of the most eminent scientific thinkers of his generation, asked, "Is the meter more true than the foot? Are Cartesian coordinates false and polar coordinates correct?" One geometry cannot be more true than another, just more convenient, he concluded.

Through his philosophy, Poincaré provided the flexibility necessary for science to advance from an age of scientific complacency. Few realized the significance of Poincaré's study of mathematical truth. Even fewer guessed that, in 2 decades, absolutes of classical science such as space, time, and substance would become *approximates*, and the most respected astronomer would agree that, if man could look deep enough into space, he would see the back of his head.

But the human mind is capable of much more than just abstraction. Driven by the social pressures of war, business competition, and ego, labor-saving machines were developed. The first mechanical aid to calculation was the abacus. The Phoenician word ABAK, the name of a flat slab covered in sand in which figures could be drawn, provided the root for the English word. During Greek and Roman times, the primitive abacus was a flat wooden board with counters. It developed into the now familiar arrangement of beads threaded on wires or laid in grooves.

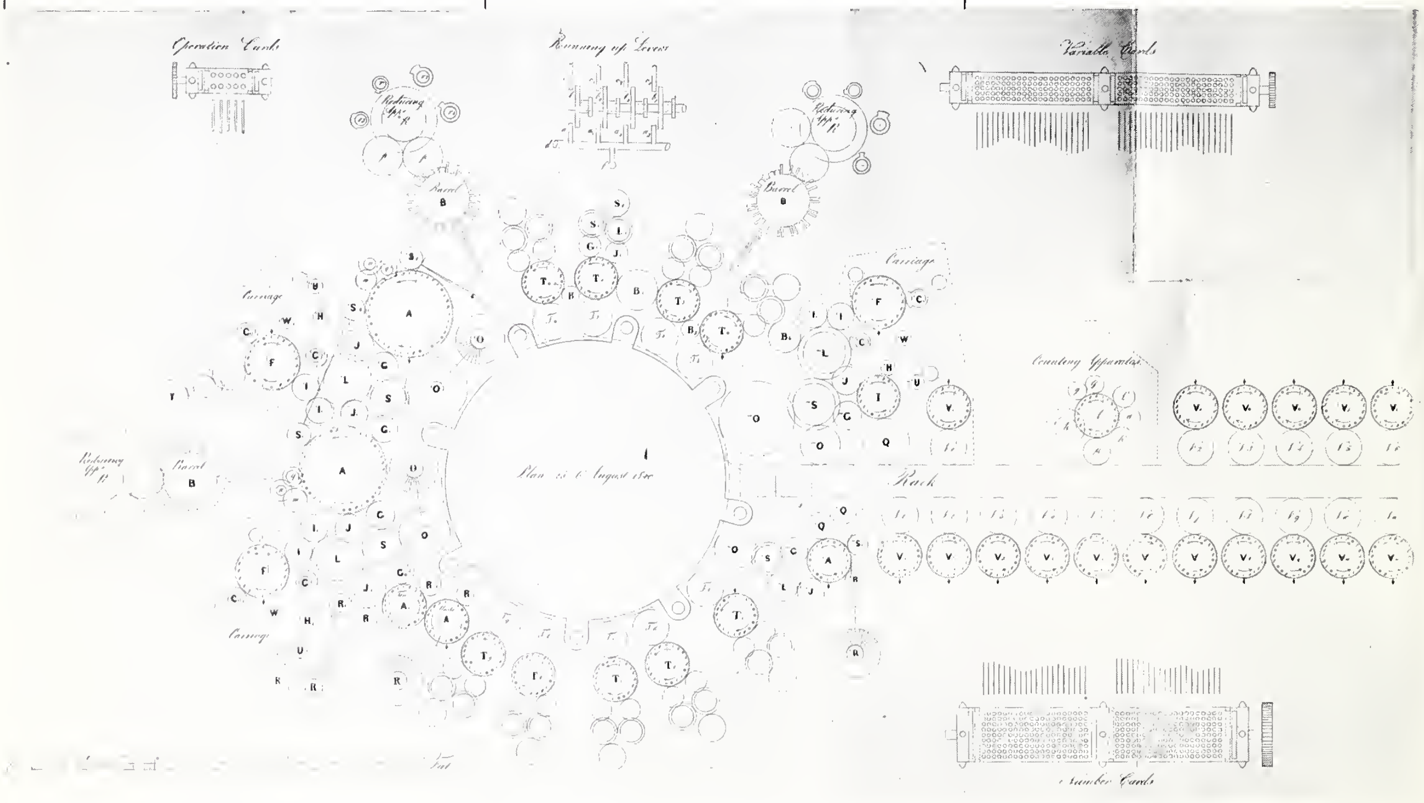
With the advent of arithmetic signs in the 15th century, the popularity of the abacus began to decline in Europe. John Napier further reduced the labor of long multiplication and division with the invention of logarithms. Multiplication and division were then facilitated by adding or subtracting the "logs" of numbers.

Before this technique could be widely used, accurate tables of logs and antilogs had to be compiled and printed. Despite valiant efforts by mathematicians to make these tables accurate, the drudgery of figuring, printing, and copying the numbers led to errors. Often, mistakes were handed down from generation to generation as mathematicians built, all too faithfully, on the wobbly shoulders of those who had gone before.

An alternative to the use of mathematical tables was soon developed—an analog device known as the slide rule, which consists of two numbered scales mounted side-by-side in a manner to permit sliding them easily back and forth. Whereas the modern digital computer counts, the analog device measures quantities. The slide rule scale is arranged so that numbers fall at distances corresponding to their logarithms. Essentially, multiplication is accomplished by adding two lengths together. Division is done by subtracting two lengths.

As with all analog devices, the accuracy of slide rules is limited by the accuracy of measurement. Their use did not solve the problem caused by incorrect tables, but rather introduced a lack of precision. The solution, of course, was to produce reliable tables.

In 1812 this thought occurred to Charles Babbage and John Herschel, two young mathematicians, while they were checking logarithm tables for errors. As recounted by Babbage in later writ-



Schematic of the Analytical Engine designed by Charles Babbage in the 19th century: a grand exercise in futility.

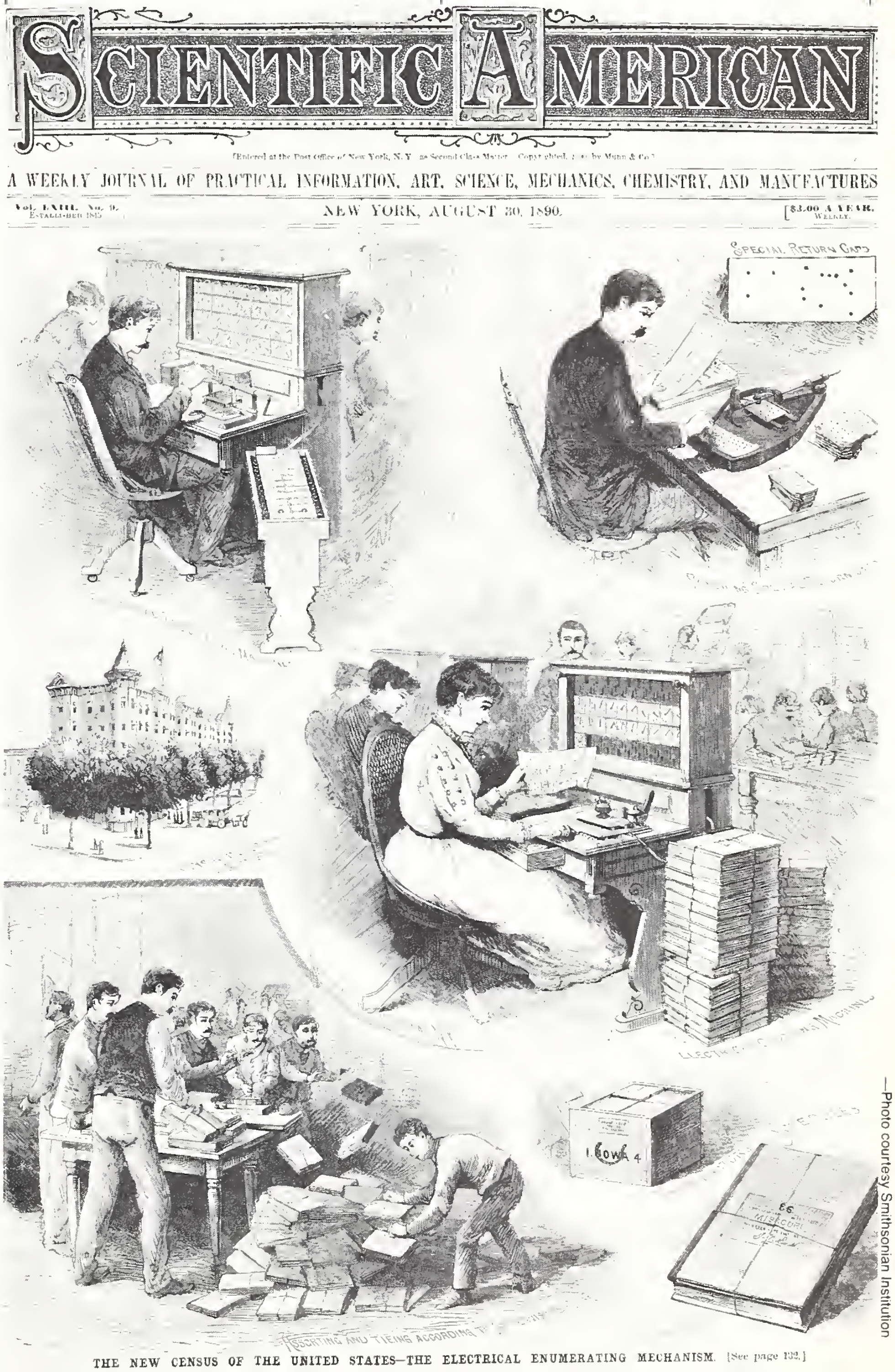
—Diagram courtesy The Science Museum, London ings, he exclaimed: "I wish to God that these calculations had been executed by steam." Herschel replied: "It is quite possible." And so occurred the idea that was to dominate Babbage's life—elimination of error through mechanized calculation.

Because his ideas were so advanced and his standards so high, Babbage experienced one disappointment after another. In many ways the 19th century inventor's work belongs more to our time than to his own.

Babbage's first project, the Difference Engine, was to be a large, complex adding machine designed for compiling mathematical tables. Unfortunately, the machine was doomed to fail. The mechanical tolerances required for the machine to work exceeded capabilities of the time. The accuracy with which gears could be cut was inadequate. Clocks, the nearest mechanical cousins to the Difference Engine, were still laboriously fitted together by hand.

Undaunted by this challenge, Babbage designed new machine tools. He hired and trained a technical assistant. But these preparations cost money and the initial sum provided by the British Treasury soon dwindled away. Five years after beginning the project, Babbage was asking the government for more money. His request was granted. But again, the amount was not—and could not have been—enough.

After almost a decade of work and some £35,000 of government and personal monies, the project was abandoned. If completed, the engine would have been a remarkable piece of work—2 tons of brass, steel, and pewter, cut to tolerances never before attempted.

Embittered by failure, Babbage, a man of considerable wealth, having inherited £100,000 from his father, devoted much time and money to insulting and slandering figureheads of the scientific and political establishment whom he blamed for the engine's failure. But

SCIENTIFIC AMERICAN

[Entered at the Post Office of New York, N. Y., as Second Class Matter. Copyrighted, [illegible] by Munn & Co.]

A WEEKLY JOURNAL OF PRACTICAL INFORMATION, ART, SCIENCE, MECHANICS, CHEMISTRY, AND MANUFACTURES

Vol. LXIII. No. 9.
ESTABLISHED 1845

NEW YORK, AUGUST 30, 1890.

[$3.00 A YEAR.
WEEKLY.

THE NEW CENSUS OF THE UNITED STATES—THE ELECTRICAL ENUMERATING MECHANISM. [See page 132.]

Scientific American *illustrates use of the Hollerith Tabulator in the 1890 census: the era of data-processing begins.*

—Photo courtesy Smithsonian Institution

he did not abandon his goal. After 1833 Babbage elaborated on a "gigantic idea" which he had first conceived while working on the Difference Engine. If built, this massive device, dubbed the Analytical Engine, would have been the first general-purpose computing machine.

Babbage's scheme contained, for the first time, most of the essential features of the modern computer. An arithmetic unit called "the mill" was designed to carry out addition, subtraction, multiplication, and division. A memory unit was to have room for 1,000 numbers, each 50 digits long—a capacity beyond technology until the first electronic computer appeared a hundred years later.

Instructions and data were to be fed into the machine on punch cards, which had been invented in 1800 by Joseph Jacquard for use with his automatic loom. After calculations were completed, resulting numbers would be printed up to 29 digits.

The Analytical Engine was as farsighted and intricate in design as it was impossible to build. Once again, Babbage's ambition had transcended his time. Even with today's technology the engine would be difficult to construct because of the mechanical tolerances required. Still Babbage's efforts were not altogether in vain. His enthusiasm spread to others, notably Herman Hollerith, who designed the first machine devoted to data processing.

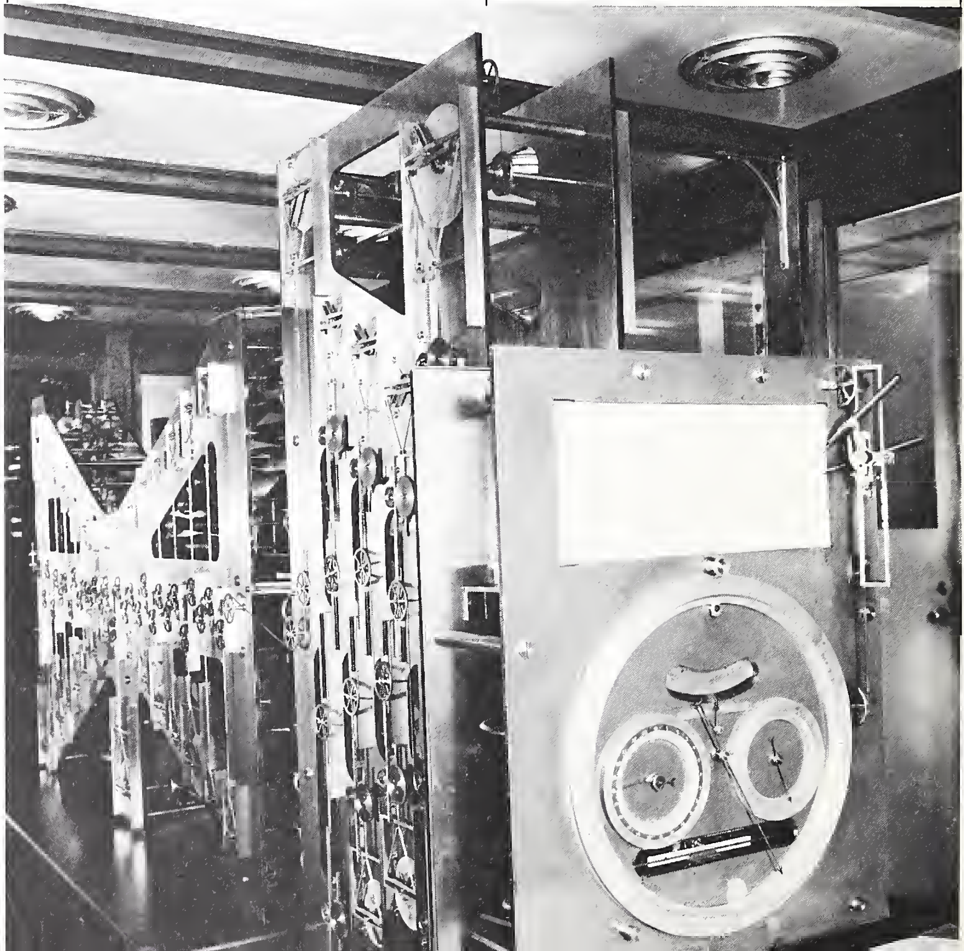
Hollerith's machine, which used punch cards, was the easy winner in a contest staged by the U.S. Census Office to pick an efficient system for tabulating the 1890 census. His device completed the test in half the time needed by his competitors, whose entries used manual methods.

Data in the form of "yes" or "no" answers were translated onto punch cards, which were compiled in a machine that electromechanically sensed positions of holes. Cards passed under a set of brushes that transferred a pulse of electricity through each hole to a metal cylinder.

After forming the Tabulating Machine Company in 1896, which was one of several businesses that later formed IBM, Hollerith built machines for sorting such cards, comparing one to the other, and printing data. To include more information for business use, Hollerith increased his punch cards to the size of the dollar bill of his time, which later became an industry standard.

His inventions opened the door to an era of computing machines, ushered in by the first efficient key-driven calculating machine. Called a comptometer, it was built by Dorr E. Felt from a macaroni box.

The rising popularity of calculating aids and machines in business characterized a shift in attitude toward the kind of work people could or should do. Calculating machines soon entered into head-on competition with people hired as "rapid calculators" by businesses trying to keep pace with expanding markets. Besides the tedium associated with mental calculation, health was also a consideration. Mental calculators, as experts in the trade were called, often complained that their evenings were haunted by unending processions of figures



The Great Brass Brain: predicting tides accurately and efficiently in 1914.

—Photo courtesy National Oceanic and Atmospheric Administration shaped like numbers.

William S. Burroughs, a bank clerk, was forced to change careers because the "monotonous grind of clerical work" had destroyed his health. At the turn of the century, Burroughs entered the comptometer field and from his early venture grew one of today's major manufacturers of digital computers, Burroughs Corporation.

As the calculating machines gained acceptance, more and more applications were found. One called Millionaire, developed in 1893 and widely used in business, found immediate and key uses in science. Percival Lowell began using Millionaire in 1905 to search for a "Planet X" located somewhere beyond Neptune. Calculations were completed in 1914, but the planet, later named Pluto, was not sighted until 1930, 14 years after Lowell's death.

At the same time that manufacturers were converting to mass production techniques, Spanish inventor Leonardo Torres y Quevedo was demonstrating a theory that heralded the oncoming age of the programmed machine in industry. Torres combined electromechanical calculating techniques with principles of automata, demonstrating that such machines can perform any desired sequence of arithmetic operations.

Torres' electromechanical Arithometer, exhibited in 1920, realized theories of automata that he had pioneered 7 years earlier. Arithmetic problems were typed in by the operator, and the Arithometer printed the answers on a typewriter. Torres became the first person to use a system of time-sharing when he linked several typewriters to one Arithometer.

One of his other inventions was a remote-controlled guidance system that successfully steered a boat through Spain's Bilbao harbor, dramatizing the fact that machines could perform tasks formerly reserved for human intelligence. Torres later built the first decision-making device—a chess-playing machine that matched a rook and king against a human opponent's king.

In 1914 *Scientific American* announced the arrival of "a great brass brain" which computed ocean tides on the basis of 37 factors, displaying the results on dials. During the first World War, ships used information from the machine to maneuver into shallow water and elude German U-boats.

After World War I, Vannevar Bush of the Massachusetts Institute of Technology invented the differential analyzer, an analog device assembled from gears, cams, and differentials that mechanically completed the various functions necessary to solve a differential equation. Bush's machine was applied to many different tasks, replacing devices such as "network analyzers" built by utility companies in the 1920's to analyze load requirements. These machines produced scale models of power networks, but they could not predict large power surges that might cause blackouts. The differential analyzer was the first machine with such a capability. Its success seemed to indicate that big, general-purpose analog computers would dominate scientific calculation in the future.

In the 1930's servomechanisms—automatic devices that

Vannevar Bush and the 1930's differential analyzer: "I was trying to solve such problems of electric circuitry as the one connected with failures and blackouts in power networks. I had been thoroughly stuck because I could not solve the tough equations the investigation led to."

—Photo courtesy Smithsonian Institution

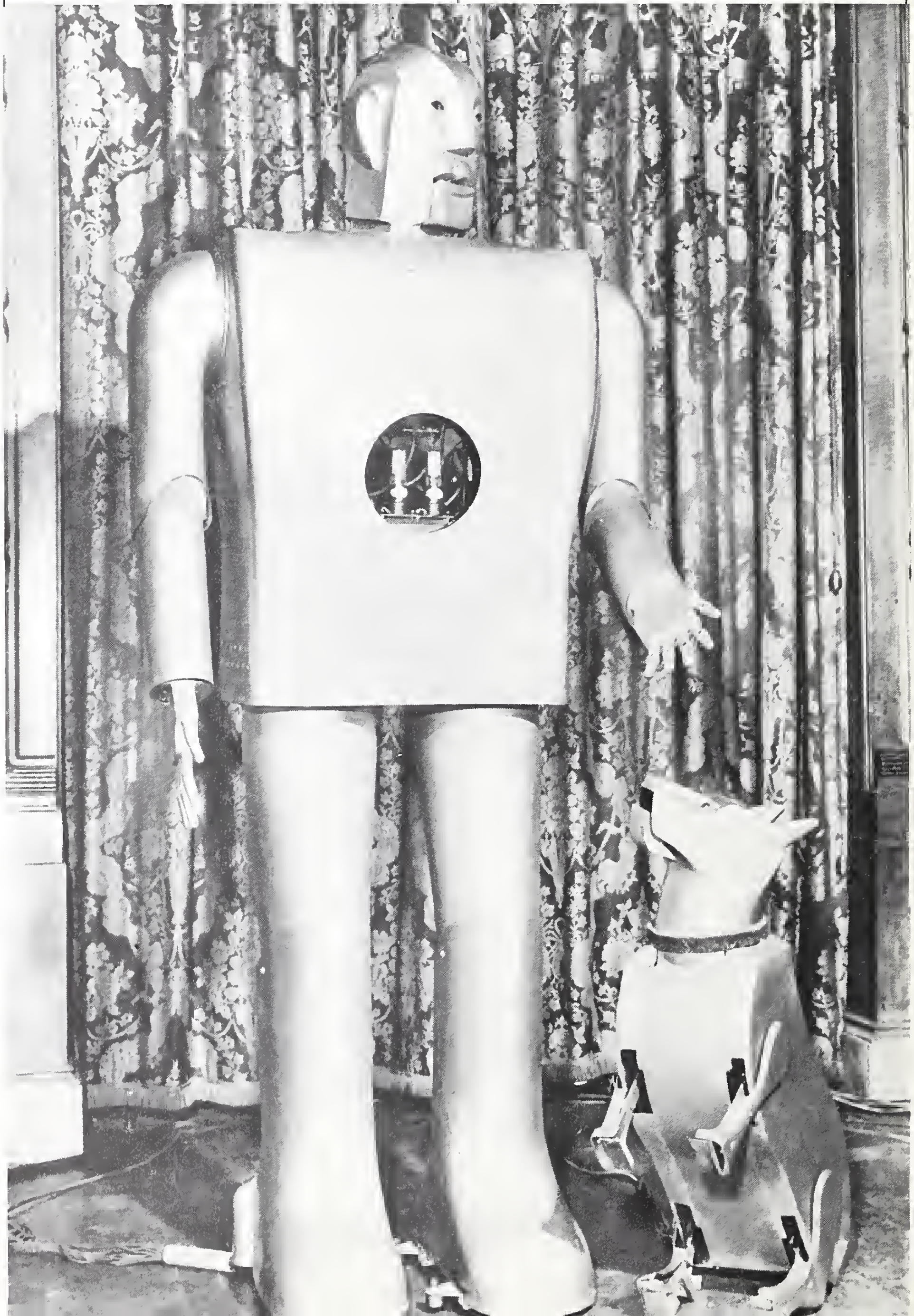
controlled other machines by monitoring their output—came into use. Oil refineries and syrup-production plants were among the first to use these "machines that boss other machines." As control problems were reduced, more and more applications were found for the servomechanisms. Steam turbines, airplanes, and chemical processes were soon included in the domain of the new device.

As machines surprised society with newfound abilities, their creators took to flights of fancy, building robots in exaggerated human forms. Inventors built tin-can contraptions that walked, talked, and responded to mechanical commands. The robots' lifelike actions were an elaborate illusion, as they were controlled by simple automatic devices or, remotely, by human operators. As such, they were no more than novelties, commonly used in product and company promotions or fairs.

Willie Vocalite, built by Westinghouse in 1931, was one of these. Willie had a stovepipe head, expressionless face, and cauliflower ears. At the inauguration of passenger air service between New York and San Francisco, Willie made a speech, wished everyone bon voyage, helped start the engines, and after his official duties were completed, relaxed with a cigarette in the company of a lovely model hired for the occasion.

Eight years later, Willie's metal cousin, Elektro, a stocky, tough-looking robot, appeared at the New York World's Fair with his faithful companion Sparko, the first robot dog. Elektro walked, talked, counted on his fingers, puffed cigarettes, and could distinguish between red and green with the aid of a photoelectric cell. Sparko barked, wagged his tail, sat up, and begged.

In the late thirties, engineers turned their collective genius to problems raised by the second coming of world war. The U.S. Army set out to improve differential analyzers used at Maryland's Aberdeen Proving Grounds to calculate firing tables for artillery batteries. Modifications increased speed and accuracy by a factor of 80, allowing the machine to produce one trajectory every 15 min-



—Photo courtesy United Press International

A.
Elektro and Sparko en route to the 1939 New York World's Fair: tin-can contraptions that walked, talked, and responded to mechanical commands.

B.
ENIAC, the world's first electronic computer, begins operation in 1946: an unwieldy collection of vacuum tubes and relays that could only be programmed by manually changing plug-and-socket connections and by setting switches.

utes as compared to the 20 hours needed by a skilled mathematician. But the machine was limited by its design to processing differential equations: it could only calculate the functions of vectors.

"There exist problems beyond our ability to solve, not because of theoretical difficulties, but because of insufficient means of mechanical computation," Howard H. Aiken said of the analyzer in 1937. He then proposed a new kind of calculating machine.

In 1938 IBM began building a forerunner of the device for Harvard University. It was called the Automatic Sequence Controlled Calculator (ASCC). After its completion in 1944, the ASCC, nicknamed Mark 1, became the first automatic, general-purpose digital calculator.

Mechanical switches called relays routed electrical signals in the ASCC. During its 15 years of use, ASCC proved to be a reliable and effective machine, but its more than three-quarters-of-a-million parts and 500 miles of wiring made maintenance expensive and difficult.

The calculator was mainly used by the U.S. Navy for ballistics and ship design. Science and industry later used the machine to generate astronomical tables and specifications for lens design. It was also used in military studies at Wright Patterson Air Force Base and in research for the Atomic Energy Commission.

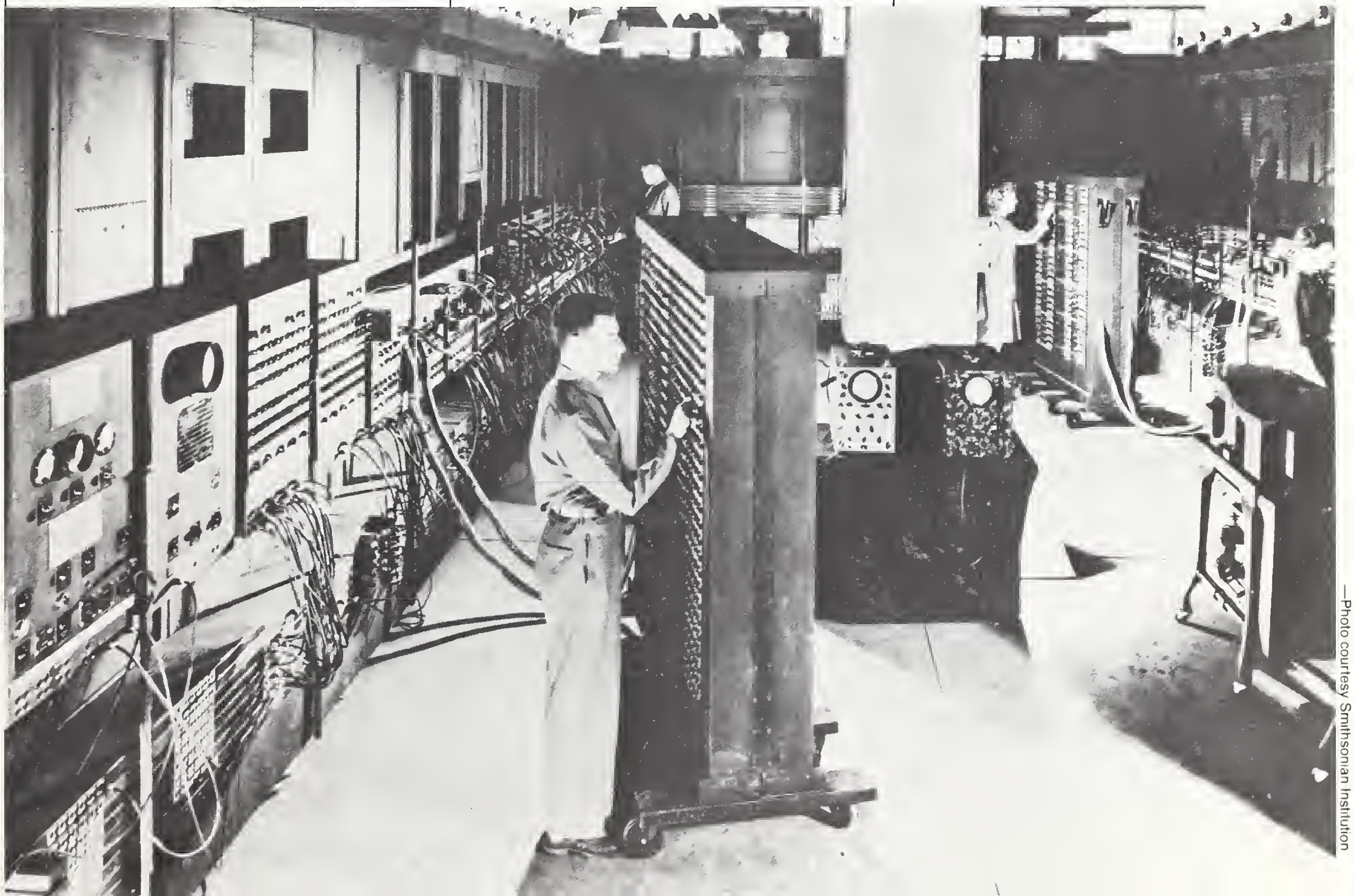
A year before ASCC was finished, John Mauchly and J. Presper Eckert, Jr., of the University of Pennsylvania, proposed the next logical step in mechanized calculation. First described as an electronic difference analyzer, the scientists predicted their new calculator would execute all functions in computing firing tables, producing each complete table in only 2 days. The device promised to get around a major failing of the differential analyzer by allowing input of such data as atmospheric resistance defined by numbers rather than by mathematical formulae.

Built in secrecy at the University of Pennsylvania, the new device, which ultimately became known as ENIAC (Electronic Numerical Integrator and Calculator), was moved to the Ballistics Research Laboratories.

People were necessary to generate firing tables on Bush's differential analyzer, and the human role slowed production. Completion of one table, on the average, took 2 or 3 months.

With the new machine, lengthy and repetitive calculations for each 60-second trajectory could be completed in just 30 seconds. But ENIAC was not completed until 1946, and the huge device, composed of some 18,000 vacuum tubes and 1,500 relays, was never used for ballistic computations. It did find wide-ranging applications in scientific calculation, however. Until the early 1950's ENIAC dabbled in weather prediction, atomic energy research, cosmic ray studies, and thermal ignition.

Germany may have entered the field of electronic computers ahead of America, although little is known about the true dimensions or operation of these machines. The most successful version, Z4, was destroyed in an Allied bombing raid. Designed by Konrad Zuse and built at the German Aircraft Research



—Photo courtesy Smithsonian Institution

Institute, Z4 was used in developing the HS 293, a flying bomb launched from Nazi aircraft.

At the war's end, Zuse could not convince Allied interrogators that he had any scientific expertise to offer, and his research came to a sudden stop. Not until the midfifties did he resume his work, this time as owner of a computer manufacturing company, which was later absorbed by a large German electronics firm.

As technology flourished during the 1940's, a major breakthrough in the burgeoning field of computer science occurred. Although the exact source of the concept is uncertain—John von Neumann, Mauchly, Eckert, or British mathematician Alan M. Turing—it was suggested that instructions could be stored as numbers in the machine itself. The idea raised the mechanized kingdom several rungs on its evolutionary ladder. For the first time, logical choices of program sequences would be made inside a machine.

Earlier, programming ENIAC and Z4 had been extremely tedious: in ENIAC by changing plug-and-socket connections and by setting switches; in Z4 by instructions punched into discarded 35mm movie film. The concept of software programming provided the basis for the next generation of computers.

The first machine with a completely "logical design," which von Neumann described at the time of ENIAC's construction, was to be called EDVAC (Electronic Discrete Variable Automatic Computer). While EDVAC was still under construction in 1948, ENIAC, after special wiring modifications, became the first computer to embody the stored-program concept. Using ENIAC's new capabilities, von Neumann and several meteorologists completed the first computer-based weather forecast. Computations for the hydrogen bomb were begun on ENIAC and completed on its successor MANIAC (Mathematical Analyzer, Numerical Integrator and Computer). MANIAC was one of many stored-program computers that followed in the wake of the new programming concept, although each differed considerably in design. EDVAC, EDSAC, JOHNNIAC (which was named for von Neumann), SEAC, SWAC, and NORC were the first few to appear.

As computers became increasingly powerful, these machines moved into new areas. In his book *Cybernetics,* Norbert Wiener explored the potential uses of automata. In 1948 W. Grey Walter entered the field of cybernetics (the comparative study of automatic control systems) with an electromechanical "tortoise" built to study simple reflex motion. "These machines are perhaps the simplest that can be said to resemble animals," Walter wrote. "Crude though they are, they give an eerie impression of purposefulness, independence, and spontaneity."

Von Neumann, decidedly a "software scientist," hoped to use automatic machines such as the modern computer to draw conclusions about complex natural organisms. He built on the idea of the Universal Turing Machine, advanced by Turing in 1936. Turing described, in theory, a machine that could do any calculation within the realm of human intellect. The

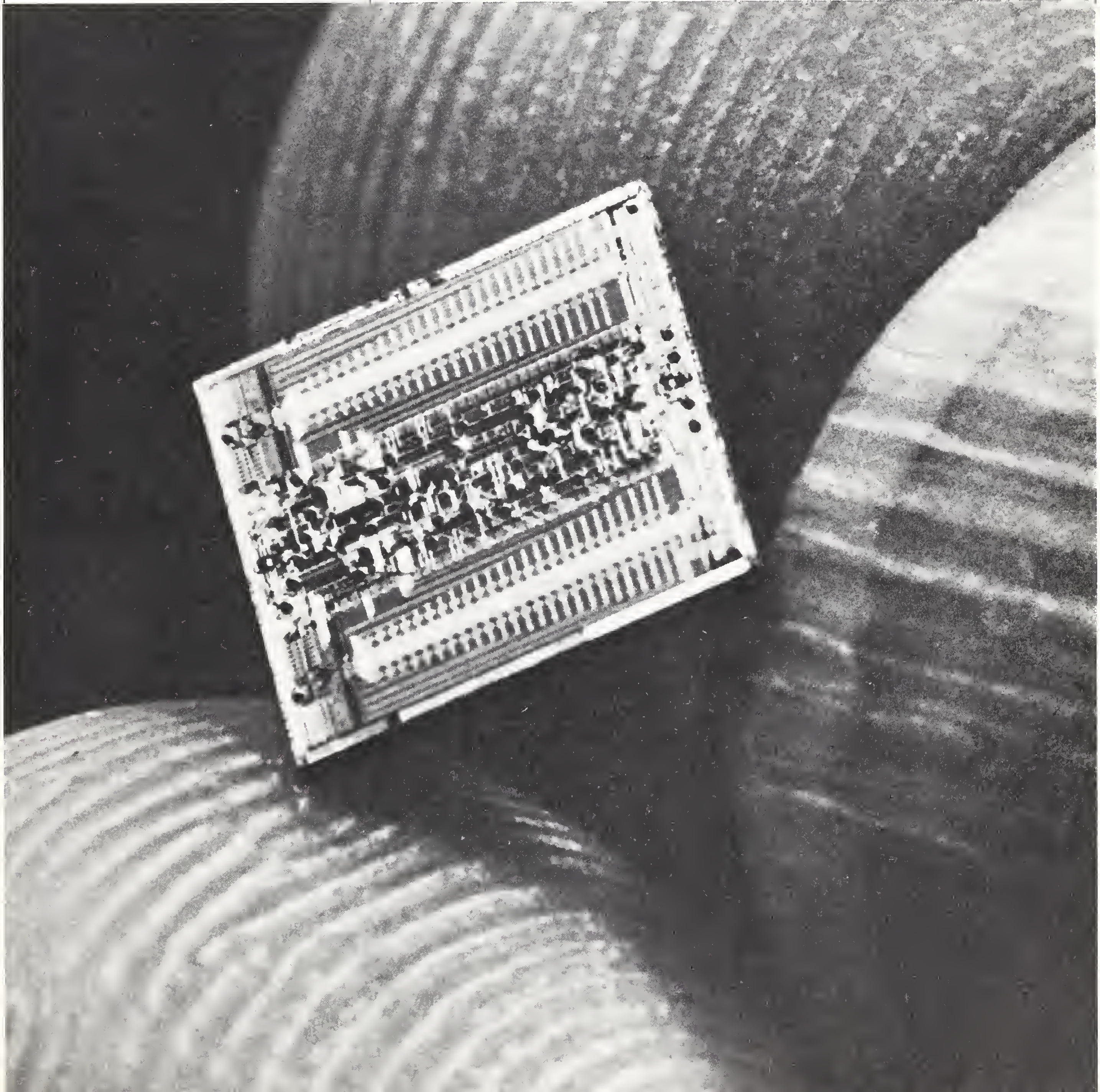
A modern integrated circuit: putting 40-times the memory of ENIAC on a chip the size of an aspirin.

Universal Turing Machine, which contains ideas later built into all general computing machines, provides a standard for measuring the complexity of a computer.

Around 1950 the computer emerged as a general tool. It had become applicable not only to military use, but also to functions in government, industry, commerce, science, education, and social science. The computer's spectacular growth in capability, applications, and numbers surprised most people. In 1954, using cathode ray tubes and magnetic drums for information storage and vacuum tubes for logic and arithmetic functions, experts in the field had estimated that only some 50 companies would eventually find use for computers.

But with the development in the late fifties of computer languages, which simplified programming, and with the introduction of integrated circuits in the midsixties, compressing the equivalent of 1,400 transistors, resistors, and diodes onto silicon chips an eighth of an inch square, the lid restraining widespread application of computers blew off. Private business began using computers to process orders, inventories, and payrolls. Computers set copy for newspapers and processed checks for banks. Airlines used computers to make and record seat reservations.

Even in medicine, an area best characterized as inexact and highly subjective, computers fitted into very specific and important niches. Applications have included such areas as electrocardiogram analy-



—Photo courtesy IBM

sis systems, aids for managing clinical routines, and instrument data collection. Statistical clustering techniques were applied to diagnostic programs and were one of the first diagnostic approaches to prove useful in medicine.

Just as these new methods of numerical calculation were being tested, the first inklings of a new approach to computing were introduced. The first application of this approach—symbolic computing—was the establishment of data banks stocked with patient information. It was an application of symbols to algorithms, matching names to numbers, and it carried along the guarantee that if pattern-matching was properly applied, the answers would be found.

But medical diagnosis, on the same level as practiced by the physician, required much more in the way of programming techniques. To grapple successfully with the problem, prototypes of high level analysis and symbolic representation were developed. Many of these resulted from early work in applying artificial intelligence (AI) to biomedical problems.

At Stanford University, under the direction of Nobel Prize winner Dr. Joshua Lederberg and Dr. Edward Feigenbaum, a team of scientists began the development in 1966 of DENDRAL, a chemistry program whose offspring now rivals experts in figuring out the structures of certain organic molecules. DENDRAL was the unlikely outgrowth of a system called the Advanced Computer for Medical Research (ACME), which had been supported by the Biotechnology Resources Program (BRP) of the NIH Division of Research Resources (DRR) since 1965. Originally ACME was dedicated to real-time analysis of data gained during biological and clinical research. The computer was oriented entirely toward numerical calculation, or batch processing of biomedical research data.

By the midsixties, principal investigator Dr. Lederberg believed that ACME had proved its worth as a data analyzer. At the urging of Dr. Feigenbaum, he proposed that the system host a new type of computer science application, artificial intelligence.

The first attempts to use ACME for AI were undertaken at Stanford, and DENDRAL was the first major effort. In the beginning the machine's size and speed were sufficient. But as DENDRAL grew, and other research was added, more and more computing time and memory space were required. By the early 1970's, Stanford's AI needs completely outstripped the machine's capacity and DENDRAL appeared to be doomed.

During this period Dr. William Baker of the BRP arrived for a site visit. In light of ACME's inadequate capacity and the interest in AI at Stanford and elsewhere around the nation, Dr. Baker suggested that ACME be dropped in favor of a system devoted to the development of this new type of computing. Stanford applied for a grant, another site visit was made, and the National Advisory Research Resources Council of the NIH recommended funding. A machine designed to handle symbolic computation was installed and, through the use of a nationwide communications network called ARPANET,

"All these delays—a thousandth of a second here, a millionth of a second there—we'll have to get the darn thing fixed."

© 1980 by Sidney Harris

a national computer community was established. An advisory group, Drs. Lederberg, Feigenbaum, and Baker, recommended that 40 percent of the computer's capacity be made available to Stanford researchers, another 40 percent spread among the national community, and the remainder assigned to development of the new system.

In 1973 SUMEX-AIM was formed as a community resource for the development of AI techniques. From its inception, the resource has been supported by the NIH Division of Research Resources' Biotechnology Resources Program. Drs. Lederberg and Feigenbaum directed the network that was to become a major medium for the development of projects like DENDRAL among a national group of biomedical researchers. In mid-1978 Dr. Lederberg left to become president of The Rockefeller University. He remains an advisor of SUMEX-AIM. Dr. Feigenbaum is now principal investigator of the resource, which currently includes some 20 autonomous projects, each targeted for application in medicine, biochemistry, or psychology.

Today's AI experts believe that computers will find many nonprofessional and small business applications as the potentials of the field become better understood. Home computers may someday regulate heating and cooling systems, notify fire or police departments in emergencies, and do tax statements, among other chores. At work, letters will be typed, orders placed, bills paid, and files searched by computer. At neighborhood shopping centers, payroll records will be maintained and stock automatically reordered according to need and profitability.

Some scientists characterize this revolution as a race to put more and more function, processing power, and storage capacity onto each semiconductor chip. Since the midsixties, the amount of function on a chip has risen by a factor of 10,000. The cost of a chip, meanwhile, has remained approximately constant at $5 to $50.

By the end of the century, computer buffs predict, this trend of broadening applications will be in full stride, as will the trend toward miniaturization, which made computing available to the general public. A single silicon chip, measuring only a few millimeters square, will be able to follow 20 million instructions per second, using 10 million cells of internal memory storage.

And just as imagination and hardware have gone hand in hand since the early 1900's, scientists predict that programming techniques and the science of software will keep pace with developing technology.

Processes of Computing

The Heuristic Mind

In the course of 30 years the computer has graduated from vacuum tubes and mechanical relays to silicon chips, each one no larger than a pencil eraser. But the transition from ENIAC, one of the first electronic number-crunchers of the late forties, to the "thinking" machines of today required more than advances in hardware. It required advances in programming concepts.

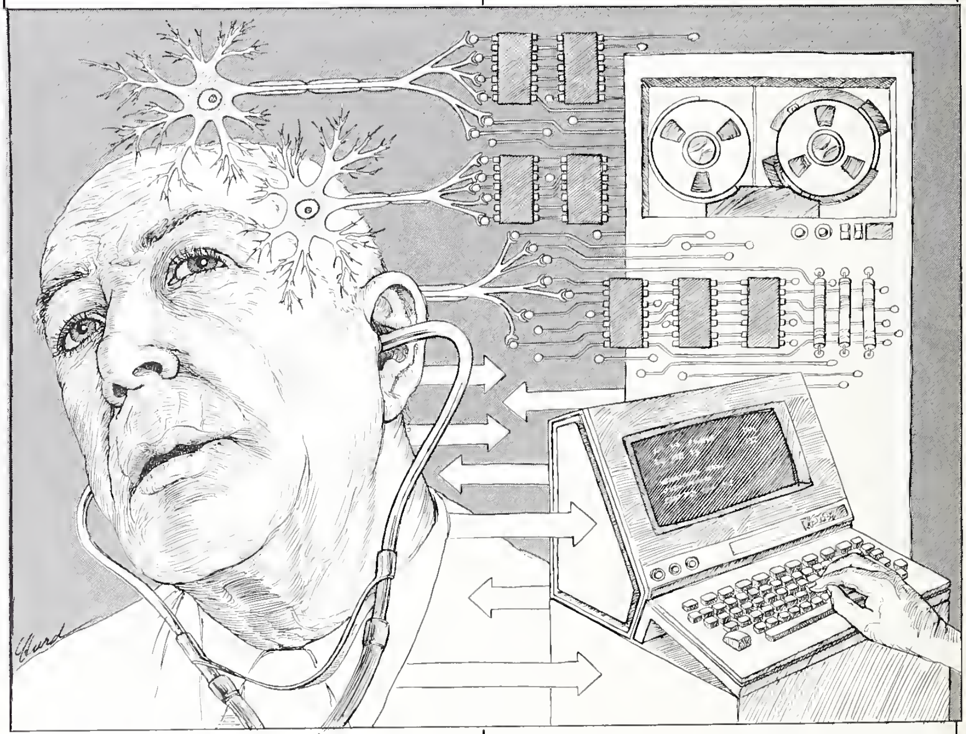
Over the last few decades, there has been an increasing emphasis on the design of knowledge-based systems. At the lowest level, these programs differ from traditional programs in two key ways. They emphasize manipulations of symbolic rather than numeric information, and they use largely informal or heuristic decision-making rules gained from real-world experience rather than mathematically proved algorithms. At a higher level, these tools of symbolic processing are used to construct understandable lines of reasoning in solving problems and to interact with human users.

Symbolic computation is necessary in certain domains, such as medical diagnosis, because comprehensive mathematical formulations do not exist. For example, the relationship of a symptom—such as "burning pain in the upper abdomen"—to disease diagnosis requires the manipulation of symbolic information.

Projects currently in SUMEX-AIM include areas of medicine, biochemistry, and psychology. The key goal of an AI program is to explain conclusions and allow the professional to interact in the decision process.

As a result, AI programs depend largely on decision-making strategies composed of heuristics, or rules based on judgment and experience, which are expressed symbolically. These strategies starkly contrast with numeric computation, which is largely algorithmic, following a mathematically fixed set of procedures when evaluating functions or tabulating results. However, the two classes of computation are not totally dissimilar.

Clinical flowcharts are algorithms used by diagnosticians when deciding how best to man-



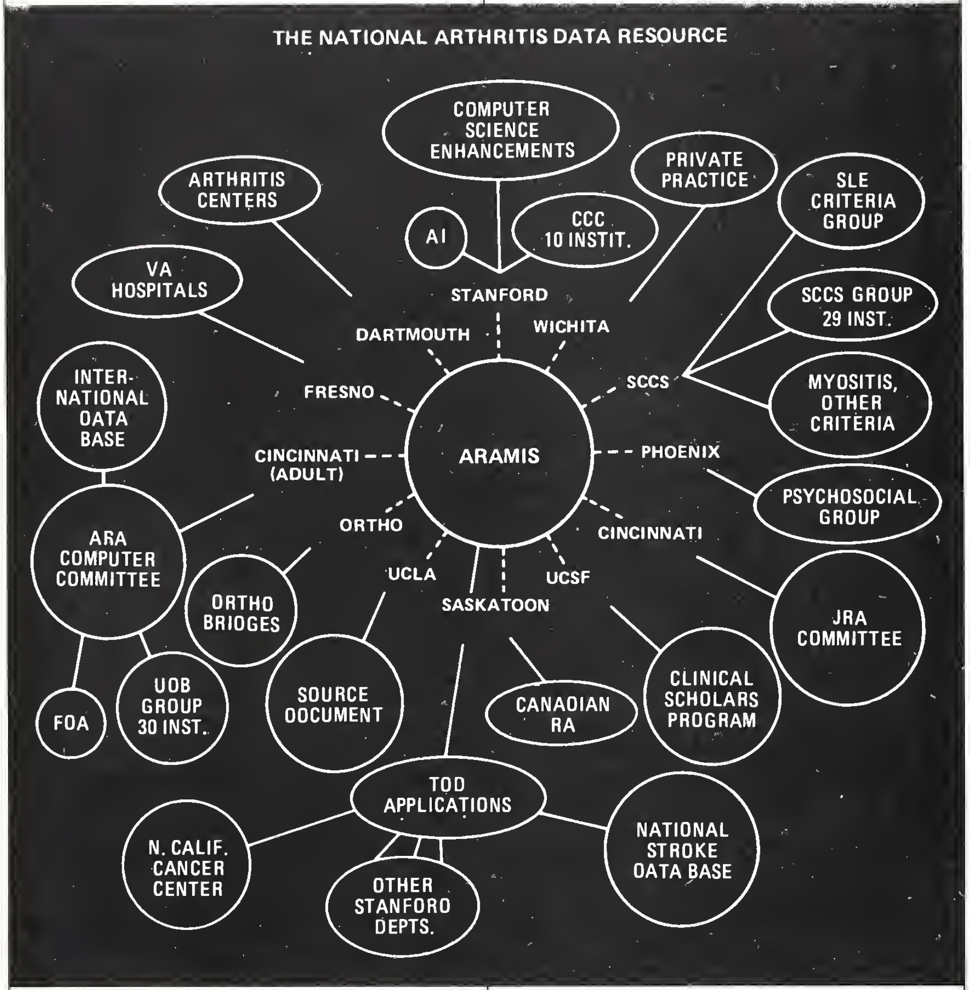
age a patient. Often these fixed procedures are designed by expert physicians for use by paramedics charged with performing certain routine tasks. As such, data are represented symbolically. Because clinical algorithms are relatively simple, computers are seldom necessary.

But automated record-keeping and data banks, more intricate examples of the clinical algorithm, require the computer. In these systems, patient names and histories and other relevant information are manipulated as symbols, and are connected to numeric data that give specific values to the information—for example, patient age: 21. Pattern-matching algorithms can be used to locate records of similar individuals or groups of patients to produce statistical summaries.

Although the earliest systems seldom did more than maintain records, there have been recent attempts to create programs that can complement this function by analyzing the stored information. ARAMIS (American Rheumatology Association Medical Information System) is one of the most successful projects in this category. In addition to search and statistical functions, the data bank offers analysis of prognosis as it relates to a specific type of patient. Programs systematically search the data base to locate case reports and summarize the outcomes of various alternative treatments, matching recorded case histories with descriptions of current patients. In systems such as this, the analysis of alternatives and the decision about the best therapy are solely up to the physician.

More complex decision-making programs attempt to assist the physician in evaluating the best treatment strategy. The decision criteria used in such programs take various forms. Some decision rules may have rigorous statistical justification, while others may be only approximate rules based on human experience and judgment. These latter strategies are called heuristics. Each type can be effective in providing solutions to problems.

In statistical approaches to diagnoses, the decision criteria have



The ARAMIS data bank: meeting needs in the study and practice of rheumatology. (Abbreviations: AI=artificial intelligence, ARA=American Rheumatism Association, CCC=cooperating clinical trials committee of the ARA, SLE=systemic lupus erythematosus, SCCS=scleroderma cooperative criteria study, JRA=juvenile rheumatoid arthritis, Canadian RA=Canadian Rheumatism Association, UDB=uniform data base for rheumatic disease, FDA=Food and Drug Administration, VA=Veterans Administration.)

been codified to a certain degree. Baye's theory of probability is one example. Essentially, Bayesian analysis relates specific patient data to different disease signs exhibited by selected groups of patients. In establishing these relationships, it is sometimes possible to compute the most likely cause for symptoms observed in a patient.

One of the earliest such programs, developed in the 1960's, was used to diagnose congenital heart disease. In some case studies, the program reached diagnoses with accuracy comparable to those rendered by two experienced physicians. As researchers honed and polished the program, applications for other disease areas were discovered. Today many types of diagnostic programs using Bayesian analysis are in operation. But Bayes' theory is just one of several techniques used in medical decision analysis.

Another displays sequences of steps representing various possible actions and events. Sequences of this type resemble tree-shaped networks. Nodes or junctions in the tree are of two kinds. At decision nodes, the clinician chooses from a set of possible actions. One action might be deciding to perform a certain test. At chance nodes, the possible responses of the patient to some action that has been taken are represented. When performing a diagnostic test, the patient's response—whether he develops complications, for example—is a matter of statistical likelihood. By using the decision tree, a clinician can come to a more informed conclusion about the range of alternative strategies.

Modifying the tree by attaching patient-oriented values to decision nodes makes the simulation more realistic. For example, a definitive diagnosis might not be pursued if the required tests were expensive and painful, if the health of the patient were not threatened by this inaction, and if rendering a definitive diagnosis would not significantly improve his health.

The effort to develop these applications into programs using artificial intelligence began in the early 1970's. The intent was to focus primarily on the use of symbolic reasoning techniques. The objectives have been to capture the judgmental or heuristic knowledge of experts for decision-making, and to construct reasoned and explainable solutions for diagnostic problems. Generally the logic built into these programs is composed of six major elements.

- **Plan-Generate-and-Test.** In this framework, the program uses heuristics to select the general area in which the answer is likely to be found. It generates plausible solutions within these boundaries, and tests conclusions against observed data, appropriately revising conclusions until one that best fits the data is uncovered.
- **Domain-Specific Knowledge.** Much of the power that decision-making programs hold is derived from specific rules and knowledge about the target area of application. Such knowledge bases encode factual information about the domain and the heuristic rules used by experts to rapidly find solutions to problems.

"Why don't you check with the local data bank?"

© 1980 by Sidney Harris

- **Flexible Knowledge Base.** If chosen properly, the knowledge base is small enough to be handled adequately by the computer, but large enough to be meaningful to the prospective user. Once the basic program is operating, knowledge can be added, removed, or changed by using an explicit and flexible encoding of the knowledge.
- **Line-of-Reasoning.** Specialists in the target area of an application must be able to follow the logic used by the program when it generates conclusions. Although not strictly necessary, specialists should also agree with the route chosen. To accomplish these goals, computer scientists in SUMEX-AIM team up with experts in target fields to learn the mechanics of reasoning. Human logic is then translated into computer language in the form of symbolic rules.
- **Multiple Sources of Knowledge.** Often several practitioners lend their expertise to the design of AI programs. Textbook knowledge is usually incorporated as well. Having access to knowledge representing varied points of view can speed the process of locating a solution and reduce the chance of overlooking alternative solutions.
- **Explanation.** The program must be able to explain the line of reasoning that led to its conclusions. If not, the user cannot understand the basis for the program's conclusions. Also, through the explanatory function, flaws in the program's logic can be located and fixed without extensive study.

Over the last decade, computer scientists have used these elements to build many types of programs. Some include the ability to learn. Others emulate creativity. Those in the SUMEX-AIM network are devoted to expert problem-solving in medicine, biochemistry, or psychology.

SUMEX and the Science Community

The Seeds of Artificial Intelligence

A typical strategy in some AI research is to choose a problem that is tightly focused and easy to conceptualize, such as a game. This approach offers certain advantages, most notably that ideas can be tested with minimal expense of time and money.

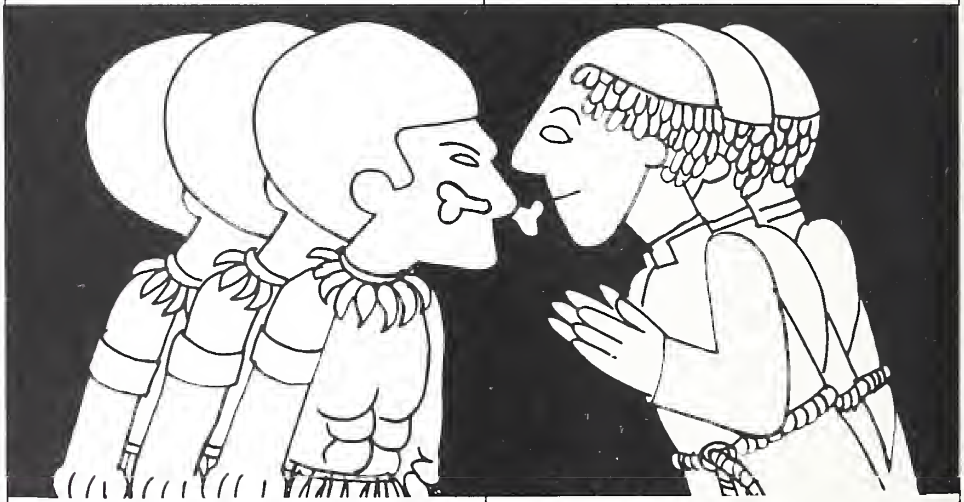
These games are called toy problems because they serve no practical use. An example is the missionaries' dilemma, a puzzle in which three missionaries want to cross a river, but their efforts are stymied by an equal number of cannibals. A boat that holds as many as two people is available, but the missionaries must never be outnumbered, or they will become the main course of that evening's cookout.

Projects in SUMEX-AIM generally shun toy problems. "Their use leads to sterility in that you quickly figure out the solution, but are not faced with the additional challenges that a messy world provides," Dr. Feigenbaum says. "We seek our inspiration from programs directed at diagnosing disease or assisting biologists in planning DNA manipulation experiments, because these problems are open-ended and rich."

The key to designing a successful AI project, he says, is to pick a problem limited enough to be conquered, but not so simple that the program designed to solve it cannot be expanded into a practical tool. Most of the time, projects in SUMEX-AIM are restricted to a subsection of an intended area of application. When that segment is adequately covered, boundaries are carefully extended.

An equally important criterion calls for an association between the project and at least one expert from the target field of application. The collaboration must be a dedicated one, according to Dr. Feigenbaum. "You cannot have the kind of inspirational meeting of minds needed for a project to succeed if the specialist and programmer meet every once in a while," he says. "It takes a quarter-time to half-time effort by the expert that stretches over a number of years."

The seed from which SUMEX-AIM grew embodies this type of collaboration. Known as DEN-



—Illustration courtesy The Design Quarter

A.
The missionaries' dilemma: use of such toy problems in AI research is not productive.

B.
A technician in the Stanford mass spectrometry laboratory: generating data for use in DENDRAL.

DRAL, it began in 1966 when Dr. Feigenbaum told Dr. Joshua Lederberg, then chairman of the Stanford genetics department, about his interest in modeling scientific thought with AI techniques. Chemistry was chosen as the target field for two major reasons. First, much knowledge in the field already existed in machine-readable form. Second, chemistry was the field in which Dr. Lederberg was expert. When the project grew in scope, Dr. Carl Djerassi, Stanford professor of organic chemistry, was recruited.

Applying AI to science was inevitable, according to Dr. Feigenbaum. "As the computer grew in power and the cost of its use decreased, more and more specialties looked to the computer for assistance in information processing," he says. "But very few specialties in medicine and other fields of science could be modeled by formulas and calculations, which are the traditional means of exploiting the computer."

Programs of this type must employ processes similar to those present in human reasoning and, therefore, must be expressed symbolically, Dr. Feigenbaum says. To achieve this, it is necessary to develop techniques by which symbols can be represented and manipulated. But when DENDRAL was conceived some 15 years ago, artificial intelligence was truly a fledgling discipline.

Biochemistry

DENDRAL

The project was initially begun as a prototype to demonstrate that computerized symbolic reasoning could be successfully applied to molecular structure problems in chemistry. The program illustrates well the evolution of AI work.

In solving problems, DENDRAL uses instrument data from a mass spectrometer (MS) and a nuclear magnetic resonance (NMR) spectrometer, together with other constraints on structural features in the molecule. These constraints describe configurations of atoms and provide limits within which the answers, structural candidates for an unknown compound, must fit. Such constraints eliminate the production of undesired substructures which, based on chemical or energetic grounds, are implausible.

Drs. Lederberg and Feigenbaum quickly realized the power provided by supplying several sources of knowledge when analyzing molecular structures. In an early case run on DENDRAL, constraints based on organic chemistry principles alone would have admitted 1.25 million plausible candidate structures for a single compound under study. The scientists responded by adding information from proton NMR analyses, from which the program could infer a few additional constraints. "The set of plausible candidates was then reduced to one—the right structure!" Dr. Feigenbaum recalls. "This was not an isolated result but showed up dozens of times in subsequent analyses."

The original DENDRAL program was restricted to a small number of molecular families for which the program had been given a specialist's knowledge, "namely the families of interest to our chemist-collaborators," Dr. Feigenbaum says. "Within these areas, DEN-

DRAL's performance was usually not only much faster but also more accurate than expert human performance."

Dr. Bruce Buchanan, a member of the DENDRAL team, explains the general approach of DENDRAL. "There are three phases—plan, generate, and test," he says. "In approaching a problem, DENDRAL makes some rough guesses as to what the solution should look like. That is the planning phase. The generation phase works within the established constraints of the plan to develop plausible solutions. Finally, each plausible solution is tested."

Testing is accomplished in two steps, which follow a "model-driven strategy." First, the computer generates sets of instrument data that would be expected to describe each candidate structure. These sets are then compared to actual data about the compound. The closest fits are retained and ranked accordingly. Having enough knowledge about the characteristics of a certain type of compound to do model-driven analysis drastically reduces the amount of data that must be examined, since the data are used mainly to verify possible answers.

DENDRAL's primary limitation was its restriction to only a small subset of organic molecules, the saturated, aliphatic, monofunctional compounds. Work carried out after DENDRAL's early success has focused on the structure-generation aspects of the plan, generate, and test paradigm. From this paradigm, the structure generator, called CONGEN for CONstrained structure GENeration, has been extracted. CONGEN is the segment of the main program that is not closely tied to specific instrumental data and is, therefore, of greatest use.

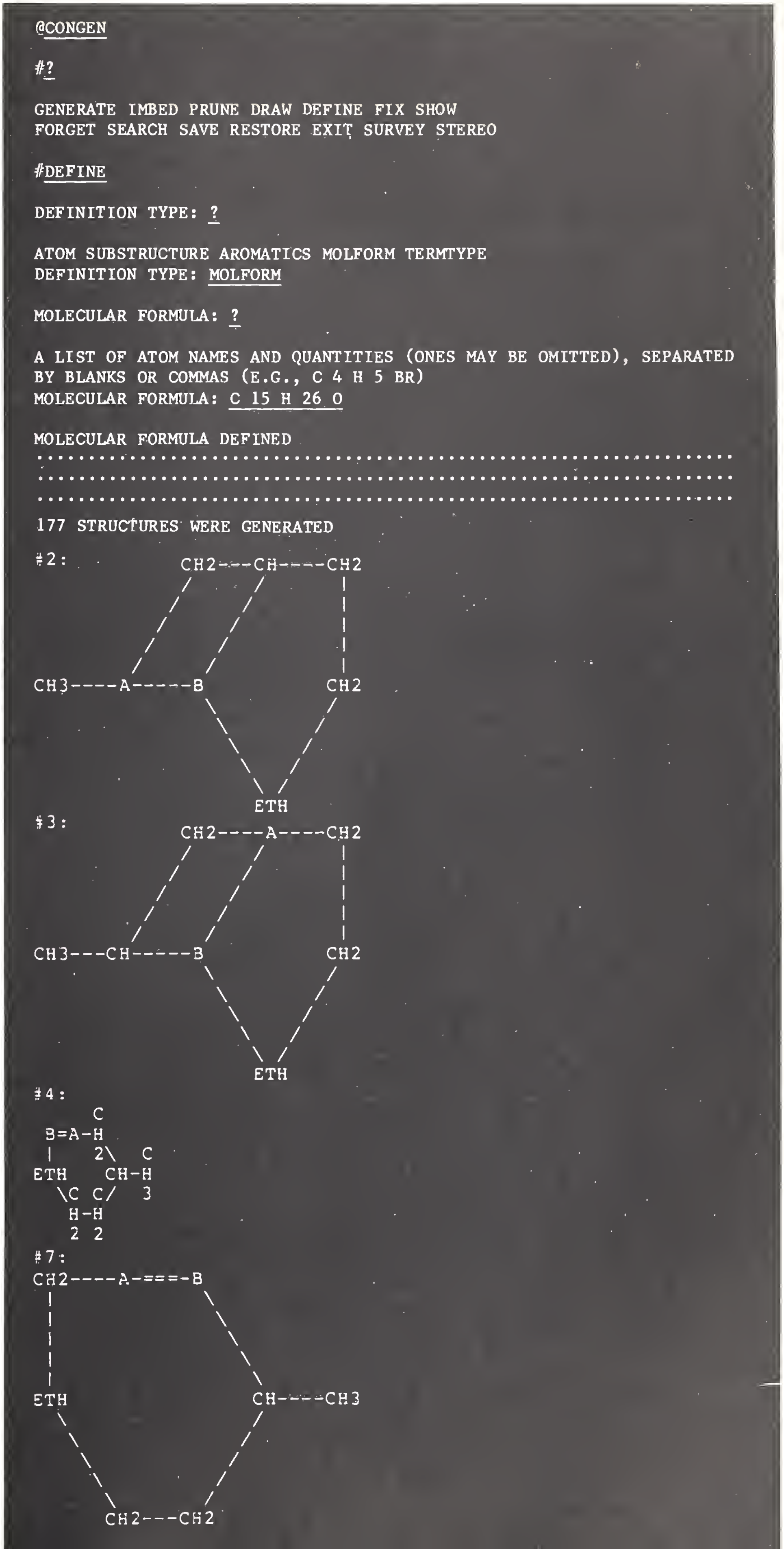
"Chemists have many sources of data for both planning and testing, so the use of DENDRAL as a whole, which would restrict them to NMR and mass spectral data, would be a hindrance," Dr. Buchanan says. "That is why, in the last 3 years, almost all the effort on the project has gone into developing CONGEN, since it has the widest possible applicability."

Now under the direction of Stanford chemists Drs. Carl Djerassi and Dennis Smith, the DENDRAL project has evolved into one of the best known and most successful applications of artificial intelligence. The CONGEN program and related subprograms aid chemists in determining the molecular structure of unknown organic compounds. Because the molecular structure of a compound must be known before its other properties can be studied—properties related to pharmacology or toxicology, for example—DENDRAL promises an important contribution to biomedicine. Some investigators have already capitalized on this offer.

During the past 5 years the CONGEN program has been used successfully by chemists working on biomedical problems at Stanford and other institutions. About two dozen scientists use the program each year when solving questions about the structures of compounds. Investigator affiliations are split about 50-50 between universities and private industry. The program has been exported to several laboratories in the United States. The British government is now supporting work at the University of Edinburgh aimed at linking industrial researchers in the United Kingdom with CONGEN. A copy of the program now runs on the Edinburgh computer. A colleague at the Australian National Research Organization is also spearheading an effort to make CONGEN available in that country.

More recent research efforts have been directed to extending CONGEN's representation of structure even further. The program will soon include principles of molecular stereochemistry, or three-dimensional representation of structures. Stereochemistry is absolutely essential in understanding structures and interactions of molecules in chemical and biochemical systems, Dr. Smith explains. This new work is pointed toward a system of computer-based planning and testing which incorporates chemical and spectroscopic data from several different techniques.

As the forerunner of AI's shift to knowledge-based analysis, DENDRAL holds a special place in computer history. It demonstrated



CONGEN printout: currently one of the most successful applications of artificial intelligence, this program helps chemists determine the molecular structure of organic compounds.

```
@CONGEN

#?

GENERATE IMBED PRUNE DRAW DEFINE FIX SHOW
FORGET SEARCH SAVE RESTORE EXIT SURVEY STEREO

#DEFINE

DEFINITION TYPE: ?

ATOM SUBSTRUCTURE AROMATICS MOLFORM TERMTYPE
DEFINITION TYPE: MOLFORM

MOLECULAR FORMULA: ?

A LIST OF ATOM NAMES AND QUANTITIES (ONES MAY BE OMITTED), SEPARATED
BY BLANKS OR COMMAS (E.G., C 4 H 5 BR)
MOLECULAR FORMULA: C 15 H 26 O

MOLECULAR FORMULA DEFINED
.....................................................................
.....................................................................
.....................................................................

177 STRUCTURES WERE GENERATED
#2:         CH2---CH----CH2
            /     /       |
           /     /        |
          /     /         |
         /     /          |
        /     /           |
CH3----A-----B          CH2
              \         /
               \       /
                \     /
                 \   /
                  \ /
                  ETH
#3:         CH2----A----CH2
            /     /       |
           /     /        |
          /     /         |
         /     /          |
        /     /           |
CH3---CH-----B          CH2
              \         /
               \       /
                \     /
                 \   /
                  \ /
                  ETH
#4:
     C
 B=A-H
 |   2\  C
ETH   CH-H
  \C C/  3
   H-H
   2 2
#7:
CH2----A-===-B
 |            \
 |             \
 |              \
 |               \
 |                \
ETH               CH----CH3
  \               /
   \             /
    \           /
     \         /
      CH2---CH2
```

the superiority of domain-specific knowledge as a means to achieve expert performance and in so doing raised important issues concerning knowledge representation, acquisition, and use.

But, more important than its obvious contributions, the program demonstrated that AI concepts and programming techniques were advanced enough to produce useful tools, although each could deal with only one limited specialty. This example of competence, according to Dr. Feigenbaum, vastly improved the credibility of AI and paved the way for other such systems. "For us, the DENDRAL system has been a fountain of ideas, many of which have found their way into our other projects," Dr. Feigenbaum says.

Meta-DENDRAL

The project in SUMEX-AIM most closely associated with DENDRAL, as might be expected by its name, is meta-DENDRAL. Developed by Dr. Buchanan, professor of computer science at Stanford, the program learns rules about a specific type of compound by examining data from a set of examples. These rules can then be used to interpret data concerning unknown organic compounds. Both DENDRAL and meta-DENDRAL use the same rule-based logic. Criteria set up by expert chemists guide meta-DENDRAL's generation and selection of rules.

Dr. Feigenbaum says the program was evolved from DENDRAL for two reasons. First, it was decided that DENDRAL has laid a foundation firm enough to pursue the deeper study of scientific theory formation. Second, it was recognized that acquiring expert knowledge of a specific domain was the bottleneck in building programs targeted for real-world use.

Meta-DENDRAL was originally intended to complement the parent program. Its job was to formulate rules for interpreting data from mass spectrometer analyses. In such analyses, molecular fragments are separated according to mass and electrical charge. Meta-DENDRAL's output is sets of rules that describe how molecules fragment when studied with mass spectrometry (MS). Meta-DENDRAL also includes evidence supporting each fragmentation rule and a summary of contradictory evidence. Constraints, fed in by chemists, guide generation of rules along desired lines.

The program, like DENDRAL, uses the plan-generate-test framework. The process includes three steps: interpret the data and summarize evidence; generate a set of plausible candidates; test and refine the set of plausible rules.

In the first step, meta-DENDRAL cites each piece of MS data as a highly specific point of fragmentation, then sums up the evidence supporting such fragmentation and the configurations that would cause these atoms to separate. The next step is a heuristic search for general rules that govern the fragmentations. The search begins with the single most general rule and proceeds toward more detailed specifications. This process continues until the program decides that the rules being generated are becoming too specific. Meta-DENDRAL also includes a

Dr. W. Todd Wipke, principal investigator of the SECS project: designing syntheses faster and without the bias of past experience.

criterion for deciding whether an emerging rule is too general.

In the final stage, the program tests candidate rules, comparing positive and contradictory evidence. Those with a negative balance are disregarded. Rules with redundant features or supported by the same evidence are merged.

The end result is a rule-set of comparable quality to those that could be generated by human experts, according to Dr. Buchanan. "In some tests, meta-DENDRAL recreated rule-sets that we had previously acquired from our experts during the DENDRAL project," he says. "In a more stringent test, involving a family of compounds for which the mass-spectral theory had not been completely worked out by chemists, the program discovered rule-sets for each subfamily."

These rules were judged by experts to be "excellent." A paper describing them was published in the American Chemical Society Journal in 1976.

Emphasis during the past year has been to make meta-DENDRAL more efficient. A major overhaul was accomplished, largely reorganizing the methods by which the program works. With these changes, the ability to generate rules concerning a different type of data, carbon 13 nuclear magnetic resonance, was included. Several papers were published in 1979 on the rules generated in this area.

SECS

The SECS (Simulation and Evaluation of Chemical Synthesis) project is aimed at describing the logical principles used when constructing molecules. Developed primarily by Dr. W. Todd Wipke, a chemist and computer scientist at the University of California, Santa Cruz, SECS is intended to promote the development of new and modified drugs, as well as synthetic compounds modeled after those that occur naturally. In particular, the project is concentrated on assisting the chemist to design and select syntheses of biologically important molecules. Dr. Wipke says the computer offers several advantages over conventional methods.

"Using SECS, chemists should be able to design syntheses faster and without the bias of past experience," he explains. "Many more possible syntheses will be considered because of the system's extensive library of chemical reactions, which is larger than any person can remember. And the computer can better process and record the many structures that will result."

Through on-site terminals or telephone links, investigators from university, industrial, and private laboratories are now using SECS. Versions of SECS are available by accessing SUMEX-AIM, or at the University of Pennsylvania Medical School, the International ADP Network Computers, or Merck & Company, Incorporated, among others. Dr. Kenneth Williamson of Mount Holyoke College used SECS to build three-dimensional models of some 50 compounds particularly important in nuclear magnetic resonance spectroscopy. Other scientists have successfully used the program to design chemical syntheses. One chemist used SECS to develop procedures for

—Photo by Carol A. Foote

making synthetic morphine.

"These users have given us a lot of suggestions for improving the program," Dr. Wipke says. "Some have contributed new reactions and quite a few people from industry have actually contributed labor to the project—quite sophisticated labor. In one case, an organic chemist from Hoffman-LaRoche who had worked in the field of heterocyclic chemistry for 10 years spent a year endowing the program with his knowledge of chemistry."

The scientist's lack of experience in computer science was not a problem because SECS uses a special language called ALCHEM, which was developed by Dr. Wipke's group. He says it is composed of declaratives that describe how the environment of molecules influences chemical reactions.

"A chemist can understand the language and read a reaction with only about 5 minutes of explanation," he says. "To actually use the language well takes only a couple of days."

When working on a problem, the program studies data about the natural target molecule and constructs a three-dimensional model for display on a graphics terminal. Based on the analysis, SECS draws from its knowledge base to select reactions that could be used in the last step of the synthesis and then backtracks through the required precursors.

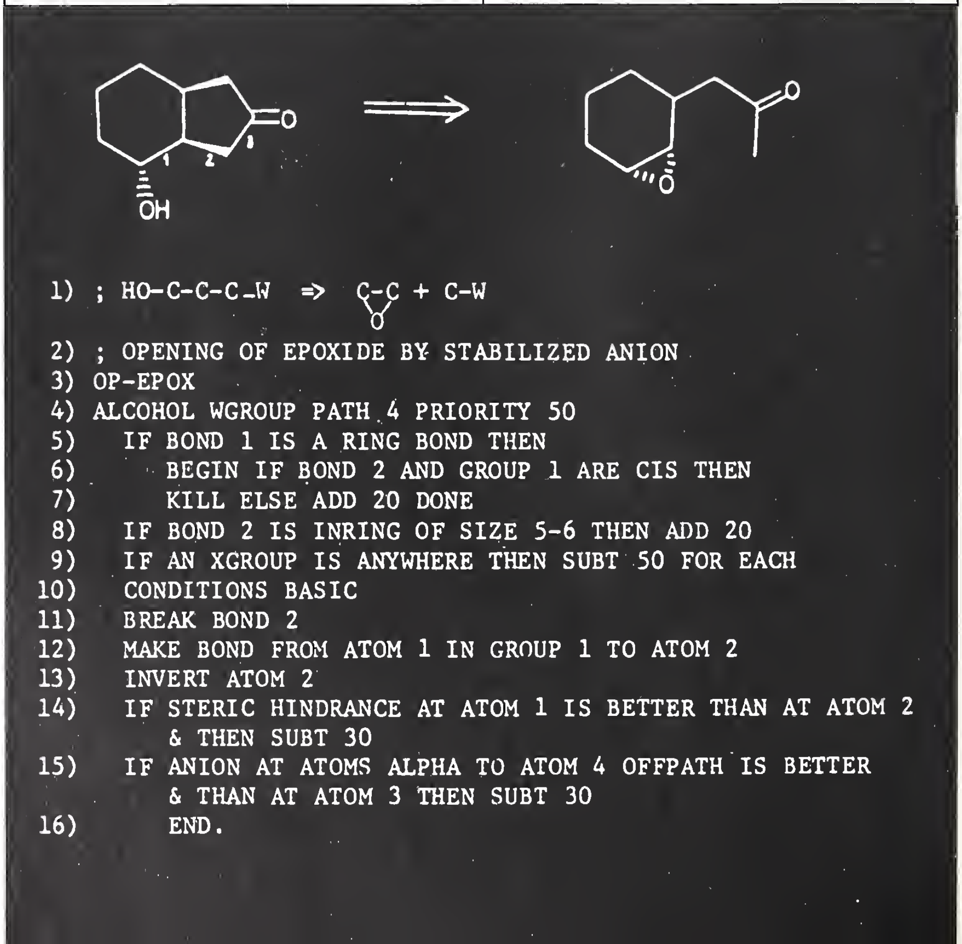
"The system stimulates the chemist's own creativity," Dr. Wipke says. "It presents many different and unbiased approaches to the synthesis." The chemist guides the computer through the process by pointing out the most interesting techniques.

"This is a unique feature of our project in terms of AI research. Usually programs are designed to find one good way to accomplish a task. We are interested in finding all the good syntheses, and that involves dealing with plans, plans that have many branches and many contingencies," he says.

There is another feature that sets the project apart from others in the field, according to Dr. Wipke. "Our program is interactive. We are tackling the problem of synthesis from the viewpoint of how best to use the chemist and the computer as a team and to have each team member doing the tasks for which that member is best suited. A lot of AI has been directed at how to make the computer do the whole thing with very little emphasis on presenting intermediate results to the user in a form that allows the search process to be guided, interrupted, stopped, or redirected."

Dr. Wipke and colleagues, mostly synthetic-organic chemists, are currently expanding the program to include more complex strategies for designing syntheses. "Essentially, the program will have a more precisely directed search, and it will be more selective in what it generates," he says.

But before these strategies can be put into the computer, they must be explicitly defined, which is often difficult to accomplish. For instance, strategies based on principles of symmetry are learned from experience rather than from textbooks, Dr. Wipke explains. For the computer to recognize a symmetrical design, these principles must be dissected and reassembled in the form of software.



```
; HO-C-C-C-W  =>   C-C + C-W
                    O
; OPENING OF EPOXIDE BY STABILIZED ANION
OP-EPOX
ALCOHOL WGROUP PATH 4 PRIORITY 50
   IF BOND 1 IS A RING BOND THEN
      BEGIN IF BOND 2 AND GROUP 1 ARE CIS THEN
      KILL ELSE ADD 20 DONE
   IF BOND 2 IS INRING OF SIZE 5-6 THEN ADD 20
   IF AN XGROUP IS ANYWHERE THEN SUBT 50 FOR EACH
   CONDITIONS BASIC
   BREAK BOND 2
   MAKE BOND FROM ATOM 1 IN GROUP 1 TO ATOM 2
   INVERT ATOM 2
   IF STERIC HINDRANCE AT ATOM 1 IS BETTER THAN AT ATOM 2
      & THEN SUBT 30
   IF ANION AT ATOMS ALPHA TO ATOM 4 OFFPATH IS BETTER
      & THAN AT ATOM 3 THEN SUBT 30
      END.
```

Unlike the current version of SECS, which uses ALCHEM to express rules concerning chemical reactions, strategies will be written using mathematical equations. Expressing knowledge in this form will allow the Wipke team to build an explanatory function into the program. If questioned by the chemist as to why a certain reaction was chosen for the synthesis, the computer will be able to reply, citing strategies of chemistry.

In final form the strategy portion of the program will complement the part that deals only with reactions. "The current program decides what to do by consulting a list of goals," Dr. Wipke says, "and that goal list will be created by this higher level reasoning process which picks out the strategies applicable to the situation and explains why. The program will then select ways to implement the strategies and, finally, decide how to modify the molecule's structure. This multistep procedure allows a view of the problem free from human bias."

Dr. Wipke hopes to demonstrate that computer-based synthesis techniques can also be applied to the study of metabolism. Based on technology from the SECS program, a new computer program called XENO has been developed to predict metabolic pathways for xenobiotic compounds—chemicals not normally found in the body. The objective is to predict plausible metabolites of a given xenobiotic. "What you put in is the chemical structure of the foreign compound," he says. "What you get out is the chemical structure of the metabolites."

Predictions of plausible metabolites result from knowledge of how compounds are activated by enzymes. Many of the mechanisms involved in these processes are known and more are being discovered.

In addition, the metabolite's stereochemistry is predicted. A compound may exist in two forms, each the mirror image of the other. One may be active while the other is not, or they may both be active but produce different effects.

"Stereochemistry in metabolism is a new frontier," Dr. Wipke says. "In the past, instruments were not sensitive enough to explore this angle using the amount of metabolite that was obtained."

In recent months problems have been submitted to the program to test its ability to predict metabolites. Dr. Wipke says XENO has been fairly successful at identifying metabolites found in laboratory studies, and also predicts metabolites that have not been found. When discrepancies occur, Dr. Wipke says, they sometimes can be traced to errors in the knowledge base. "The computer may predict more metabolism than is actually going on in living systems," he says. "That's really not too bad, because the metabolites that can be isolated will always be included in the set of metabolites predicted by the computer. The program defines a set of candidates to look for."

Dr. Wipke and colleagues have now focused on expanding the knowledge base, particularly to include models of more species. Only the rat and the mouse are currently described in detail.

An index of biological activities associated with metabolites—for example, carcinogens—is slated

SECS printout: helping the chemist design and select syntheses of biologically important compounds.

for inclusion. The function will apply pattern recognition to compounds not listed in the index as a means of classifying metabolites.

The XENO project is not the only spin-off from SECS. In 1978 the SECS program led to development of a daughter project that extends computer-assisted synthesis into phosphorus chemistry. Under the direction of Dr. Wipke, Drs. Gerard Kaufman and Francois Choplin at the University of Strasbourg in France created a knowledge base composed of reactions pertaining to phosphorus. In analyzing several compounds and searching the appropriate literature, the new system found most of the existing syntheses and, more importantly, suggested new techniques that appear to be equally good or better, according to Dr. Wipke.

MOLGEN

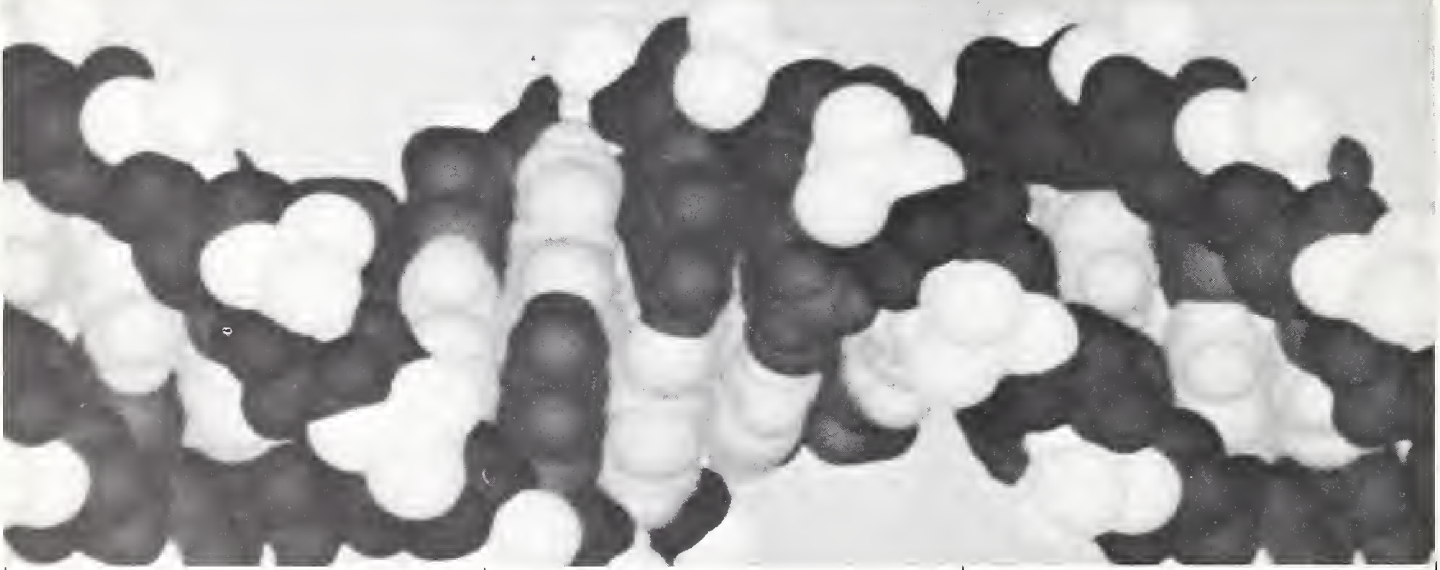
Experiment-planning in the manipulation of DNA is the goal of MOLGEN, a Stanford project being conducted in collaboration with scientists at the University of New Mexico (UNM). Program development is primarily under the direction of Drs. Laurence H. Kedes and Edward Feigenbaum, computer scientists Drs. Mark Stefik and Peter Friedland, and biochemist Dr. Doug Brutlag. MOLGEN's task is to advise geneticists about the design of laboratory experiments. These include methods used to analyze and modify nucleic acids.

MOLGEN is mainly focused on organizing experimental techniques and determining the order in which they should be applied to achieve specified goals, Dr. Brutlag says. "The enormous volume of detailed knowledge makes it likely that good experiments are being missed," Dr. Feigenbaum says. "We believe that an intelligent planning assistant can offer help in anticipating the results of combining experimental methods in many ways."

Dr. Peter Friedland says MOLGEN makes near-expert decisions when selecting physical methods, such as electron microscopy or enzymatic modification, to analyze molecular structure. Eventually the program will be expanded to include the design of synthesis experiments in which methods for building molecules will be described. Functional analyses will be added as well, allowing the program to identify the products of nucleic acids.

The success of MOLGEN as an experiment designer depends on the quality of its knowledge base. Much effort has been expended to supply the base with explicit information about DNA structures, restriction enzymes, a hierarchy of laboratory techniques, and a growing collection of genetics-oriented strategies for discovering information about various aspects of DNA molecules. Most common analytical and manipulative methods have already been put in the base.

Results of the research include some special-purpose programs in the area of molecular genetics. The most useful are highly refined versions of previously existing strategies. Many of these concern determination of the sequences of nucleic acids in DNA. Modifications are focused on technical aspects of the programs; those leading to improved efficiency, for example, and those addressing human en-



—Photo courtesy NIH Division of Computer Research and Technology

DNA's spiral ladder of heredity: helping chemists manipulate the molecule through well-planned experiments is the focus of MOLGEN.

gineering concerns to make it easier for scientists not familiar with computers to use the programs.

In addition to its applied orientation, MOLGEN includes an AI research dimension: use of the knowledge domain of molecular genetics to create a generally applicable problem-solving program. The system is designed to allow generalization into domains beyond genetics in future research and application.

"Integrating the many diverse sources of knowledge is a central problem in constructing MOLGEN because the expert-planning process requires a blend of biological, genetic, chemical, topological, and instrument knowledge," Dr. Feigenbaum says. "The expert's knowledge of experimental strategies must also be represented and put to use."

PROTEIN STRUCTURE PROJECT

Building computer models of protein structures from crystallographic data, particularly electron density maps, is the goal of the PROTEIN STRUCTURE project at Stanford. Electron density maps are data representing the structures in three dimensions. Unfortunately, these maps are usually crude and ambiguous. As a result, the program depends largely on background information, such as the amino acid sequence in a protein, for guidance and support in forming hypotheses about the compound's three-dimensional structure.

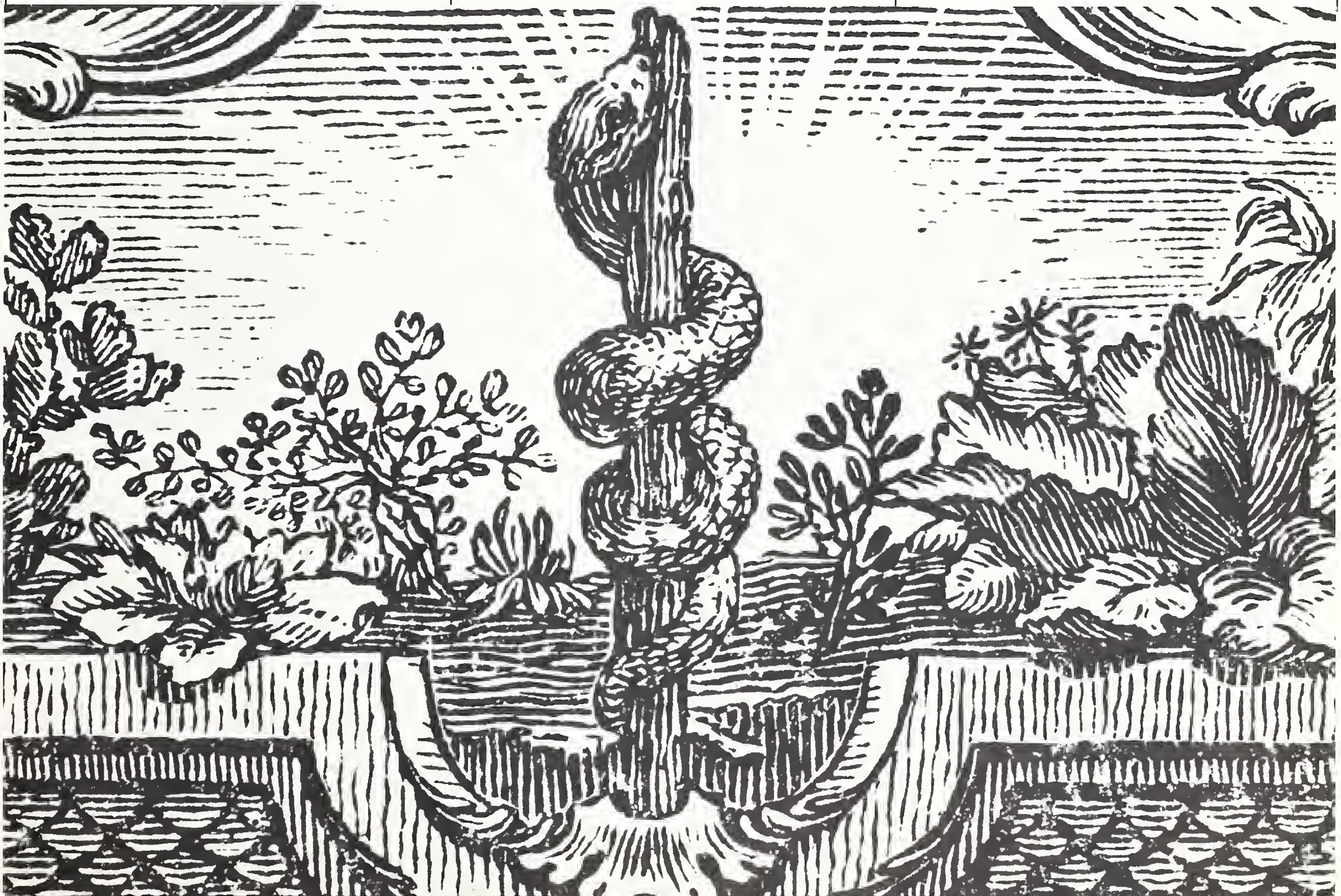
Because the shape of a molecule exerts a major effect on its performance, accurate analyses and representations of molecular structure are seen by medical researchers as essential to understanding the biological function of these complex molecules.

Interpreting electron density maps is the art of a protein chemist, which the system's logic scheme attempts to capture through the use of heuristic rules. Due to the size of protein molecules, which often contain many thousands of atoms, the plan-generate-and-test strategy used by DENDRAL cannot be employed. Rather, the system pieces together hypotheses by concentrating successively on specific areas of the protein. The project is under the direction of Drs. Feigenbaum and Robert Engelmore of Stanford University with assistance from Mr. Allan Terry, at the University of California (UC) Irvine, and the strong collaboration of Dr. Stephen Freer at the UC San Diego.

Clinical Medicine

INTERNIST

Heuristic search techniques are used in all SUMEX-AIM projects, although each differs according to the purpose of the project. Drs. Jack D. Myers and Harry E. Pople, mentors of INTERNIST at the University of Pittsburgh, reasoned that the best way to design a computer program for solving difficult problems is to simulate the mental processes used by people. They are primarily interested in building a program that will aid skilled specialists in solving complicated problems concerning internal medicine. Spin-offs from the program might



be used by physicians' assistants or in rural health care clinics, military outposts, and spacecraft. To be effective, the program must be able to diagnose several diseases if they are present in a single patient, and it must render diagnoses quickly to reflect the current status of the patient.

Drs. Myers and Pople analyzed the diagnostic routine followed by the expert clinician and established a set of criteria:

- Observations fed into the computer must evoke the appropriate hypotheses of disease.
- Hypotheses must generate a list of manifestations that would be present in the patient if the diagnosis is correct.
- The computer must be able to rank models of disease according to their probability of being correct and must be able to decide when the weight of evidence is sufficient to permit reasonably confident judgment.
- The program must be able to group hypotheses into mutually exclusive subsets corresponding to different diagnoses, in order to handle cases in which more than one disease may be present.

Since beginning their work in 1970, Drs. Myers and Pople have developed an operative system, and in so doing have partially achieved these objectives. INTERNIST accepts descriptions of disease manifestations in any order and asks for more information, such as historical items, symptoms, signs, and laboratory data. These facts are not entered in a specific order, but rather as they are gained through tests and observations. As facts accumulate, nodes of recognition are triggered and a pattern begins to develop.

Now, with a specific direction, the computer fits the data together like pieces of a jigsaw puzzle. An interlocking web of programmed data is set up, beginning with categories such as liver disease leading into specifics like hepatitis A. After sufficient data have been fed into the computer, disease models are developed. The models are then compared and ranked.

"The computer holds the profile for each disease in its memory and if the model fits that standard profile very closely, it could make a diagnosis," Dr. Myers says. "If that isn't possible, it will set out asking questions to obtain further information, so that one or more of the models can be confirmed."

A second generation program dubbed INTERNIST-II, which may speed up the diagnostic procedure, is now being designed. Although experimental, the new program has shown promising results, raising hopes that it will lead to a more efficient workup of clinical problems when the program is applied.

Dr. Myers predicts that within 5 years INTERNIST might be diagnosing disease on a practical, rather than experimental, basis. When completed, the system will be able to assist physicians working on difficult cases and paramedics serving in remote or medically underserved areas.

"The computer's assessment of a patient's condition will be regarded as evidence to help the practitioner form a diagnosis," he says. "The program is intended to serve as a consultant, not as a replacement for the physician."

Currently, more than three-fourths of the knowledge appli-

Drs. Jack Myers (pointing) and Harry Pople of INTERNIST: "No one can possibly memorize all the data in medicine."

cable to internal medicine, one of the broadest specialties, has been translated into symbolic data structures and stored in the computer memory. Over the past 2 years INTERNIST's ability to translate this vast store of knowledge into accurate diagnoses has been proved, using a variety of difficult case studies that were published in medical journals or occurred in Pittsburgh teaching hospitals. "In the great majority of cases the program has been effective in sorting out the pieces of the puzzle and coming to a correct diagnosis," Dr. Myers says. "The knowledge base is too incomplete for a comprehensive test in a clinical situation, although it is used on an *ad hoc* basis at Presbyterian-University Hospital, Pittsburgh, for clinical guidance."

Within 1 year, the knowledge base is expected to reach a "critical stage of completeness," according to Dr. Myers. Soon after, field trials of INTERNIST are scheduled to begin at Presbyterian-University Hospital. If successful there, a half-dozen other health care centers will take part in the testing. At each institution, Dr. Myers estimates, 20 case analyses will be run each day. During the trials, physicians' reactions to the system and their pattern of use will be recorded. On the basis of this information, INTERNIST will be revised, if necessary, to improve service to future users.

In the past year, Drs. Myers and Pople have devised a program called ZOG, which makes it possible for a physician only casually acquainted with computer science to master the use of INTERNIST, reportedly within 5 minutes. Tests show that ZOG, developed at Carnegie-Mellon University, is very versatile and easy to use. Dr. Myers says ZOG is important because the computer must be easy to operate if it is to bridge the ever-widening gap between what is known in medicine and what physicians are able to remember.

"No one can possibly memorize all the data in medicine," he says. "There's just too much knowledge and that pool of information is constantly increasing. The computer has a perfect memory and is admirably suited for a large knowledge base."

The data base is being expanded by fourth-year University of Pittsburgh medical students who participate in a medical computing course taught by Dr. Myers. Students are assigned to the project for periods varying from 6 to 18 weeks. Each student focuses on a group of diseases, usually in a specific area. A complex list of disease indicators is gathered from literature and clinical experts on the faculty. Dr. Myers and other clinicians review the data, making any necessary changes. Often students gather additional information. A major continuing effort is required to update the information as new scientific and clinical data become available.

Within the year, the knowledge base may address virtually all the important diseases in internal medicine. The team expects to meet this schedule, in view of the recent full-time addition of Dr. Randolph Miller, previously a junior collaborator on the project, and the participation of Dr. Victor Yu, formerly of MYCIN and now on the University of Pittsburgh faculty.

The first clinical tests of INTERNIST are tentatively scheduled for the early 1980's.

When in use, INTERNIST's wealth of information and diagnostic ability may solve one problem that currently plagues physicians—ineffective use of time. Patients with complex diseases often require the attention of several physicians, or of one physician over a long period of time, before definitive diagnoses can be reached. The computer program was specifically designed to handle such cases, according to Dr. Myers, and its use should speed the diagnostic process.

INTERNIST may also reduce the cost of health care by sparing patients unnecessary tests. When additional information is required to draw a computer diagnosis, the program asks for the procedure that is least expensive and presents the smallest risk to the patient. Invasive methods are always requested last. Dr. Myers emphasizes that the operating cost of INTERNIST has not yet been determined and the expense of using the system may offset the savings in physicians' fees and laboratory tests.

Several groups have indicated interest in the program. The military may want to use the system in remote outposts and submarines. The National Aeronautics and Space Administration may use INTERNIST on board spacecraft or orbiting laboratories.

The program's most valuable application may be in rural clinics where nurse practitioners and paramedics serve health needs with only periodic and short-term contact with physicians. Tucked away in remote areas, terminals linked to INTERNIST by telephone could type out important, sometimes vital, analyses of patients' conditions.

"Suppose the diagnosis comes back 'pneumococcal pneumonia,' which may be something the paramedic thought the patient had all along," Dr. Myers says. "This confirms his idea. Treatment is straightforward and the paramedic would go ahead. On the other hand, if the computer indicates something much more serious or complicated from the standpoint of treatment, I think the paramedic would say, 'Look, this is not for me. We'd better send this patient to a hospital.'"

INTERNIST, the consultant and librarian, is also a teacher. The program is being adapted to list basic disease manifestations and to score users on the number of additional indicators requested, the estimated cost and risk of tests cited to obtain more information, and the number of incorrect diagnoses. The electronic "quizmaster" has been used by members of the university hospital staff who have found it to be an "interesting and educational experience." Other teaching programs are being planned. The most ambitious scheme in the program's future is linking INTERNIST with other computer-based diagnostic systems.

One such interface may include a diagnostic program for hematology at Cornell University. Dr. Myers says hookups to this and other programs may allow INTERNIST to pursue more detailed diagnoses. But INTERNIST's ability to generate diagnoses comparable to those

MYCIN printout: computerized consultant on infectious diseases that diagnoses, recommends therapy, and justifies its decisions.

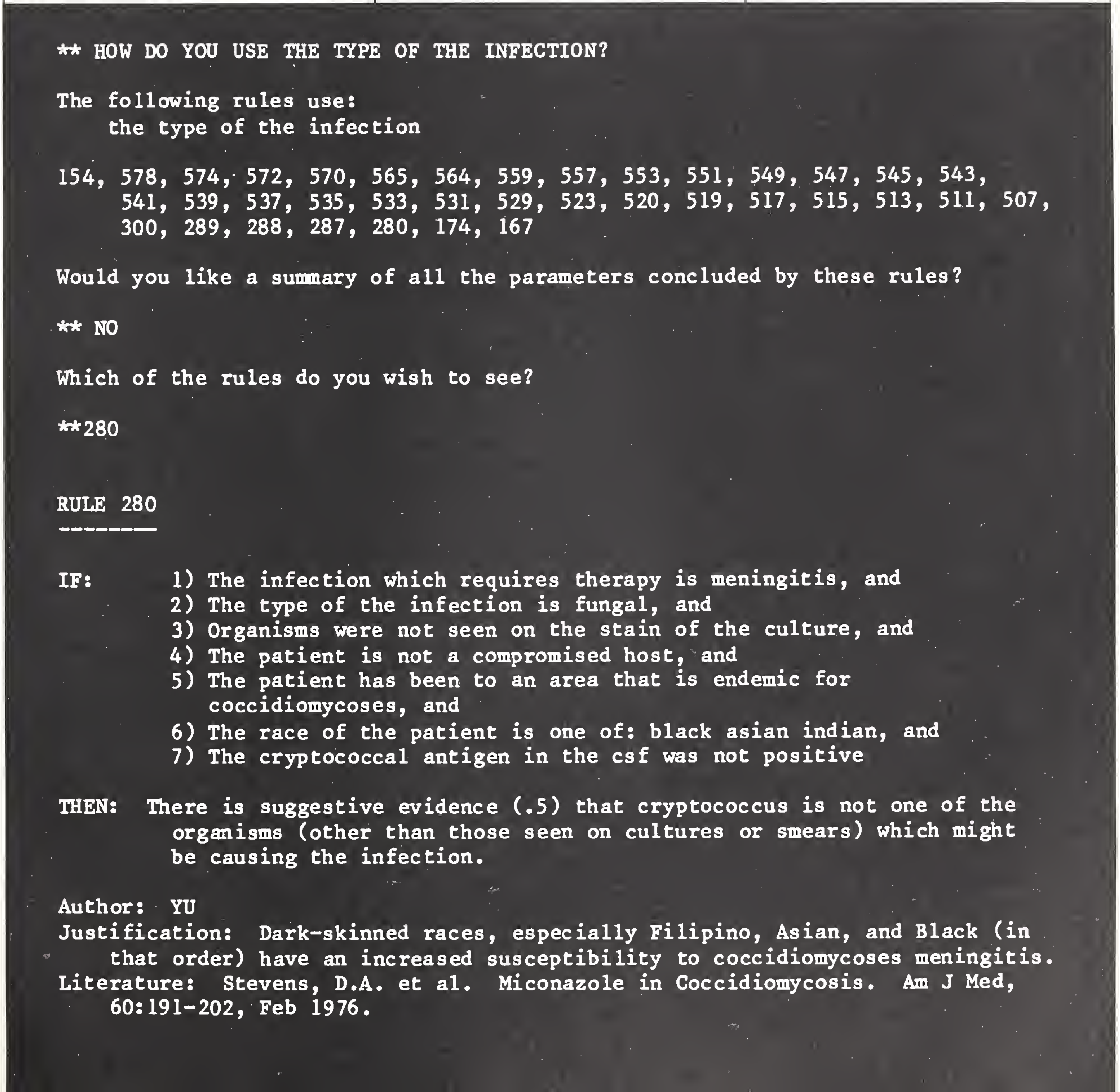
of a human expert is limited to the specialty that it embodies. Although expanding the program to include other areas that have been scrutinized and defined may be the eventual goal, the consensus among AI researchers is that an all-encompassing diagnostic program is still in the distant future.

As Drs. Feigenbaum and Lederberg stated in 1970 when reporting the first major summary of results from the DENDRAL project, ". . . general problem-solvers are too weak to be used as the basis for building high-performance systems. The behavior of the best general problem-solvers we know, human problem-solvers, is observed to be weak and shallow, except in the areas in which the human problem-solver is a specialist. And it is observed that the transfer of expertise between specialty areas is slight. A chess master is unlikely to be an expert algebraist or an expert mass spectrum analyst, etc. In this view, the expert is the specialist, with a specialist's knowledge of his area and a specialist's methods and heuristics."

"Subsequent evidence from our laboratory and all others has only confirmed this belief," Dr. Feigenbaum says.

MYCIN/EMYCIN

MYCIN holds true to the contention that program input must come from expert specialists, not general problem-solvers. The program specializes in the diagnosis and therapy selection for patients with specific infectious diseases. Its goal is to provide sound therapeutic advice, using available informa-



```
** HOW DO YOU USE THE TYPE OF THE INFECTION?

The following rules use:
    the type of the infection

154, 578, 574, 572, 570, 565, 564, 559, 557, 553, 551, 549, 547, 545, 543,
     541, 539, 537, 535, 533, 531, 529, 523, 520, 519, 517, 515, 513, 511, 507,
     300, 289, 288, 287, 280, 174, 167

Would you like a summary of all the parameters concluded by these rules?

** NO

Which of the rules do you wish to see?

**280


RULE 280
--------

IF:       1) The infection which requires therapy is meningitis, and
          2) The type of the infection is fungal, and
          3) Organisms were not seen on the stain of the culture, and
          4) The patient is not a compromised host, and
          5) The patient has been to an area that is endemic for
             coccidiomycoses, and
          6) The race of the patient is one of: black asian indian, and
          7) The cryptococcal antigen in the csf was not positive

THEN:  There is suggestive evidence (.5) that cryptococcus is not one of the
          organisms (other than those seen on cultures or smears) which might
          be causing the infection.

Author:  YU
Justification:  Dark-skinned races, especially Filipino, Asian, and Black (in
     that order) have an increased susceptibility to coccidiomycoses meningitis.
Literature:  Stevens, D.A. et al.  Miconazole in Coccidiomycosis.  Am J Med,
     60:191-202, Feb 1976.
```

tion to identify all the organisms likely to be causing the infection. At present, two major types of infections are thoroughly covered in the knowledge base—blood infections and meningitis.

MYCIN developers Drs. Bruce Buchanan and Edward Shortliffe, in collaboration with Drs. Stanley N. Cohen and Stanton Axline, built on techniques evolved through years of DENDRAL experience and knowledge of how large volumes of domain-specific information can be represented. Their model of logic is the use of "production rules" which represent facts and their interrelationships.

The MYCIN knowledge base currently consists of some 500 such rules. Essentially, each rule is an IF . . . THEN statement, which encompasses a set of preconditions to justify a conclusion. For example, *IF the gram stain of the organism is gram negative, and the morphology of the organism is rod, and the aerobicity of the organism is anaerobic; THEN there is suggestive evidence that the identity of the organism is Bacteroides.*

A.
MYCIN scores better than human prescribers: therapy was classified as acceptable if evaluators rated it as equivalent to their own.

B.
Drs. Bruce Buchanan (left) and Edward Shortliffe of MYCIN: "Building a valuable resource for the practicing physician."

MYCIN is able to identify microorganisms, prescribe antibiotic drugs, and explain, in detail, its advice. When recommending a diagnosis or therapy, MYCIN lists the rules and cites literature references used in the decision-making process. Any or all of the rules are available in detail upon request. The explanatory function allows the physician rationally to reject the program's advice if there is disagreement over its recommendations. Physicians can ask if certain facts, such as the patient's age, were included when forming the diagnosis and therapy recommendations. Or they may request the computer to justify a decision. Rather than simply stating that a clinical indicator suggests a type of microorganism, the computer might say that such an organism is more common than others in a clinical setting and is the likely cause, since the infection occurred after hospitalization. In each case, the computer cites rules and references to medical literature to support its recommendations.

From MYCIN's explanatory function nonspecialists may learn about the complexities of diagnosis and therapy for infectious disease. Furthermore, access to the rules used by the program is a means to expand its knowledge base. Developers can improve the base by entering information that may be missing or inadequately stated. Because there are dozens of exceptions, contradictions, and qualifications for each rule, and because medical research is constantly generating new information for diagnosis and therapy, updating the knowledge base is essential.

Like INTERNIST, MYCIN also has the ability to question physicians. Test results and symptoms may be requested, as well as observations about the patient's appearance.

Although MYCIN has not yet been tried in a routine clinical situation, three separate tests of the program have been very successful. In the most recent, the system's recommendations concerning therapy for patient cases with infectious meningitis were compared to those submitted by specialists, physicians of varying degrees of expertise, and one

Ratings of Antimicrobial Selection Based on Evaluator Rating and Etiologic Diagnosis

Prescribers	Number (%) of Cases in Which Therapy Rated Acceptable by an Evaluator (n=80)	Number (%) of Cases in Which Therapy Rated Acceptable by Majority of Evaluators (n=10)	Number of Cases in Which Therapy Failed to Cover a Treatable Pathogen (n=10)
MYCIN	52 (65)	7 (70)	0
Faculty, 1	50 (62.5)	5 (50)	1
Faculty, 2	48 (60)	5 (50)	1
Infectious diseases fellow	48 (60)	5 (50)	1
Faculty, 3	46 (57.5)	4 (40)	0
Actual therapy	46 (57.5)	7 (70)	0
Faculty, 4	44 (55)	5 (50)	0
Resident	36 (45)	3 (30)	1
Faculty, 5	34 (42.5)	3 (30)	0
Student	24 (30)	1 (10)	3

medical student.

Ten cases involving infectious meningitis were selected by a physician who was not acquainted with MYCIN. All patients had been treated at a county hospital affiliated with Stanford, they were identified by retrospective chart review, and each presented challenging diagnoses.

Two criteria for case selection ensured that the cases would be diverse; there were to be no more than three cases of viral meningitis, and there was to be at least one case from each of three categories—fungal, viral, and bacterial. A detailed summary of each case was compiled. The summary included history, physical examination, and laboratory data. Patients ranged in age from 1 day to 73 years.

Summaries were presented to MYCIN, five faculty members in the Stanford divisions of infectious diseases of the departments of medicine and pediatrics, one senior postdoctoral fellow in infectious disease, one senior resident in medicine, and one senior medical student. None were associated with the MYCIN project. These seven physicians and student were asked to prescribe therapy for each case on the basis of information in the summaries. There were no restrictions concerning the use of reference materials.

Dr. Buchanan recognizes that it is difficult to define precisely the term "appropriate therapy." The recent MYCIN trial defined the term with two control standards. One was simply whether the prescribed therapy would be effective against the pathogen (Table, column 3). This was not the sole criterion, because failure to cover other likely pathogens and the hazards of overprescribing are not considered. The second control was to submit the decision to the judgment of eight prominent specialists of infectious diseases at institutions other than Stanford who evaluated the recommendations (Table, columns 1 and 2). Each had published clinical reports concerning the management of meningitis. In these tests, MYCIN received a higher rating than any of the nine human prescribers. The system's recommendations scored considerably higher than the actual therapy that had been prescribed for the patients.

An important point to be made is that MYCIN and the faculty were relatively selective in the choice and number of antibiotics prescribed (Table, column 3). In contrast, the therapeutic strategy of physicians caring for the patients had been to prescribe several broad-spectrum antimicrobials. In eight cases the physicians prescribed two or three antimicrobials; in six of these eight cases, one or no antimicrobial would have been preferable.

Initial overprescribing of these agents is not necessarily wrong, since redundant or ineffective therapy can be discontinued after a pathogen has been identified, Dr. Buchanan says. But an optimal clinical strategy attempts to limit the number and spectrum of drugs prescribed, in order to minimize their toxic effects and the development of drug-resistant pathogens.

The problem of overuse or misuse of antibiotics is well-documented in medical literature.

—Photo by Bill Smith

For instance, a Stanford University study showed that one of every four persons in the United States received penicillin under doctors' orders in 1977 and nearly 90 percent of these prescriptions were unnecessary. Other studies show that nonspecialists often prescribe antibiotics that differ significantly from those that would have been prescribed by experts in infectious disease therapy.

This misuse of drugs is directly related to the immediate need for treatment required by patients with acute infections. Although culture reports can usually be obtained within 24 hours after the sample is taken, such reports often classify the organism in general terms. Due to the severity of a patient's condition, the physician may not be able to postpone treatment until a precise identification can be made, a process that may require 48 hours or longer.

"In this setting, MYCIN is designed for two roles," Dr. Buchanan says. "It can provide consultative advice to assist the physician in making the best early therapeutic decision possible from available information. And by questioning the physician about the patient, MYCIN pinpoints the items of clinical data that are essential to the validity of the decision."

MYCIN can also plot the steady-state blood levels of various antibiotics, based on such patient variables as body surface area, weight, and level of kidney function. Dr. Buchanan says this capability allows physicians to pick the most effective and least hazardous dose of prescribed drug for individual patients.

The program has been used by physicians for experimental consultation, as well as for classroom and professional demonstration. Computer scientists have studied the program, seeking information about its design and operation. But when judged in terms of being an acceptable clinical tool, MYCIN still must undergo more development. To be practical, it must be able to diagnose all major infections likely to be found in a hospital. This will require further expansion of its knowledge base. Also, refinements for convenience to physicians and ease of operation need to be worked out.

Through collaboration with other scientists working in SUMEX-AIM, Drs. Buchanan and Shortliffe have learned that it is possible to develop clinically useful programs quickly by matching the knowledge of specific application areas to the logic scheme of MYCIN. Extracting and applying the essential parts of the program to other fields has been dubbed the EMYCIN project. All knowledge and references to infectious disease have been removed in EMYCIN, but not the logic behind diagnosis, therapy recommendations, explanation, and knowledge acquisition.

MYCIN has led to the construction of two programs which are already in use. One is called SACON, a computerized consultant that helps engineers pick the proper strategy for analyzing such structures as aircraft wings, rocket engine casings, bridges, and buildings. SACON is used in conjunction with a program called MARC, which offers a large selection of analysis methods, material properties, and geometries suited to modeling the mechanical behavior of structures.

"A year or more of experience with this program is typically needed before the operator can use all the options to his best advantage," Dr. Buchanan says. "The goal of SACON is to bridge this gap by recommending an analysis strategy."

According to Dr. Buchanan, the expert who worked on the SACON project found it easy to translate his knowledge of how to operate MARC into the rule-based logic scheme developed for MYCIN.

Another spin-off from MYCIN is called GUIDON, a tutor for medical students that uses the mother program's knowledge base as a manual for diagnosing infections and prescribing antibiotic therapy. Other manuals derived from MYCIN cover knowledge from SACON and knowledge about pulmonary function, which is included in a separate SUMEX-AIM program called PUFF.

"GUIDON represents a significant twist in AI research," Dr. Buchanan says. "We are transferring knowledge from a program to people, students who want to learn the methods and strategies used by experts in their field."

The program uses AI techniques to represent both subject material and teaching strategies. Course material is represented independently of teaching procedures so that problems and remedial comments can be tailored to each student.

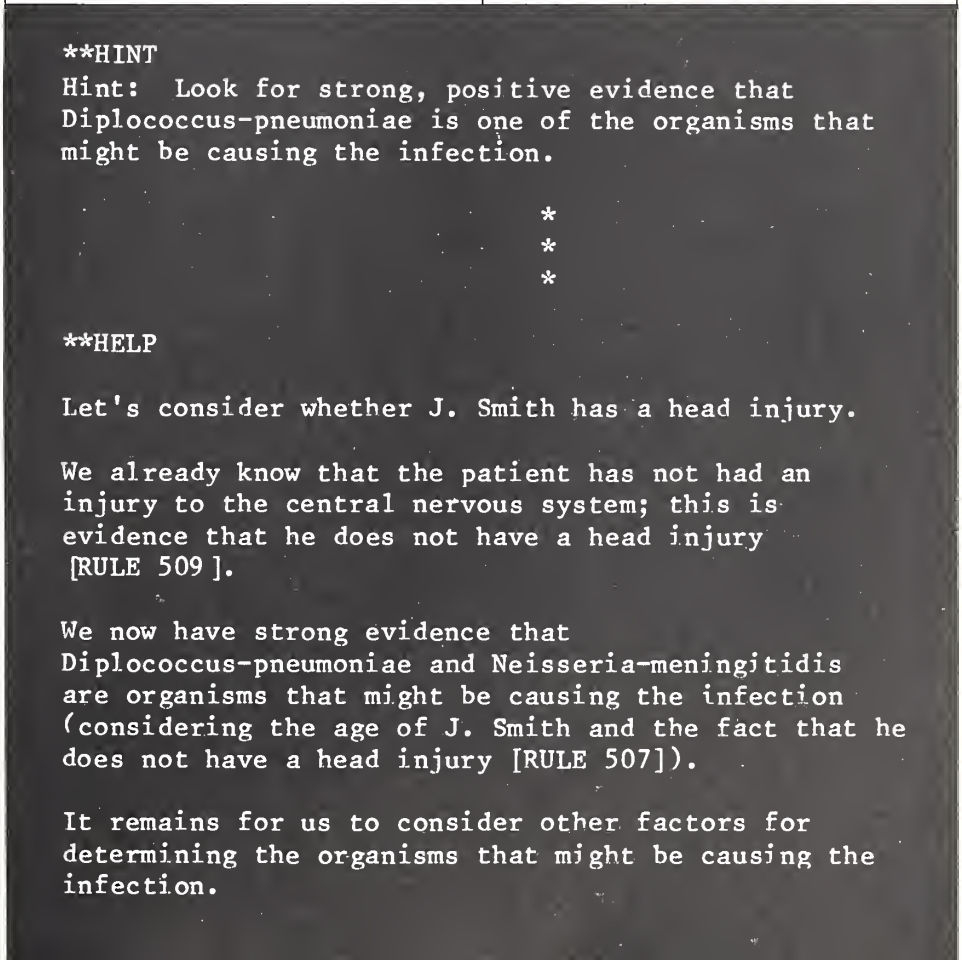
Unlike traditional computer-aided instruction (CAI) programs, which attempt to anticipate every wrong student response and prespecify references to remedial teaching material, GUIDON constructs models of the student's knowledge, reflecting the person's weaknesses, strengths, and preferred style of learning. Through a video display terminal, GUIDON and a student discuss patient cases suspected of involving an infection. The computer helps the student consider the relevant clinical and laboratory data for reaching a hypothesis about the causative organisms.

MYCIN's diagnostic rules provide the underlying expertise used by the tutorial program in selecting topics to be discussed. In addition, GUIDON has 200 tutorial rules, which include methods for guiding the dialogue, presenting diagnostic strategies, constructing a student model, and responding to the student's initiative.

"Traditionally, teaching expertise is compiled into the program, combining all kinds of problem-solving, communication, and pedagogical strategies," Dr. Buchanan says. "In GUIDON we make the important step of explicitly codifying this expertise as a body of rules within the program. The rules state how to teach, how to reply to a student, and how to guide him."

The success of these programs has generated interest in EMYCIN as an academic and business tool. Representatives of several corporations, research centers, and universities are exploring the possibility that the scaled-down version of MYCIN could be adapted to their individual needs, ranging from aircraft design to diagnosis and management of epilepsy.

One of the most promising medical projects to arise is the development of an EMYCIN therapy management system for Stanford's oncology clinic, which provides



GUIDON printout: tutoring medical students with tailor-made clues.

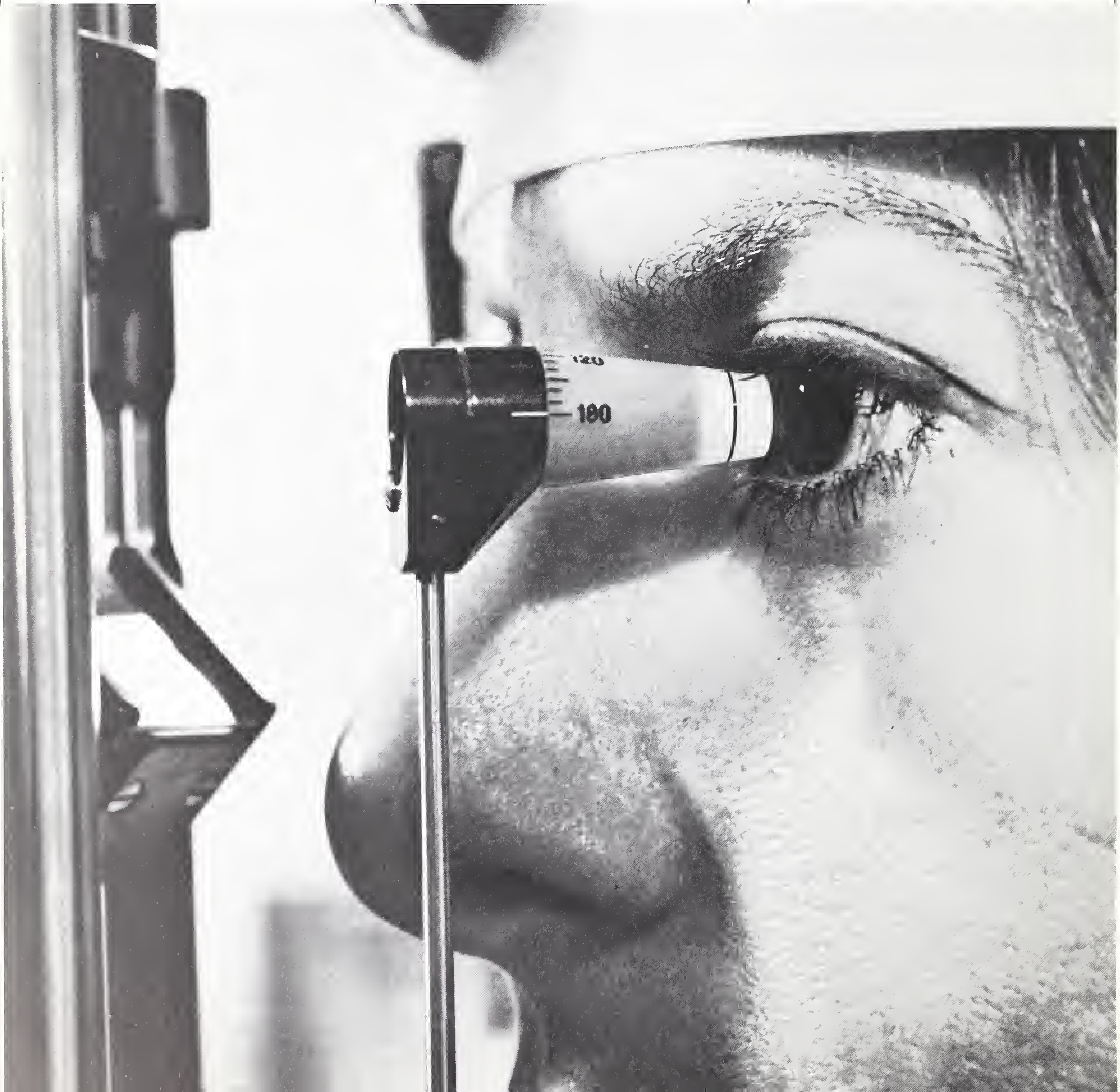
```
**HINT
Hint:  Look for strong, positive evidence that
Diplococcus-pneumoniae is one of the organisms that
might be causing the infection.

                         *
                         *
                         *

**HELP

Let's consider whether J. Smith has a head injury.

We already know that the patient has not had an
injury to the central nervous system; this is
evidence that he does not have a head injury
[RULE 509 ].

We now have strong evidence that
Diplococcus-pneumoniae and Neisseria-meningitidis
are organisms that might be causing the infection
(considering the age of J. Smith and the fact that he
does not have a head injury [RULE 507]).

It remains for us to consider other factors for
determining the organisms that might be causing the
infection.
```

chemotherapy on an outpatient basis. It is separate from the rest of the hospital clinics and its isolation offers certain advantages. "Computer terminals can become familiar to a small subset of doctors, and won't be in the way of people from other clinics," Dr. Shortliffe says. Dr. Charlotte Jacobs, the oncologist who directs the clinic, is collaborating on the system design and implementation.

The clinic has a definite need that can be filled by an AI program of the type being constructed by Drs. Shortliffe and Buchanan. Because it is a research clinic in which clinical studies are routinely performed, some 50 therapy regimens or protocols are going on simultaneously.

"There's usually one person who knows an individual protocol well, but if a resident, student, or fellow delivering care in the clinic has some question, such as whether a patient needs a chest X-ray tomorrow, it's really hard to find the answer. Up-to-date copies of the protocols are maintained but are often unwieldy and the chance of finding the needed information in a reasonable time is often very slight," says Dr. Shortliffe, who spent time in the clinic during his medical residency at Stanford.

The program is intended as a source of just such information. Protocols will be fed into the knowledge base of a program similar to EMYCIN. The reasoning process may eventually allow complicated decision-making, such as matching new patients with a set of appropriate protocols, but during the next year the scientists hope at least to accomplish simpler tasks with their new program.



The protocol for each patient will be available, presenting the day of therapy and procedures required. "A lot that is called for by protocols is currently missed," Dr. Shortliffe says. "Some people don't get appropriate X-ray or laboratory studies called for by the protocol because there is no good system for reminding the clinicians that they should be ordered. Then, when the clinicians try to analyze the data, they sometimes realize that the people enrolled in the protocol have not been categorically following the protocol itself, and the data have suffered for that reason. So, use of the program can improve the quality of studies and the way care is delivered by providing better access to the details of the protocols in an on-line fashion."

The long-term goal of the project is to build on these early algorithms until the program is capable of making inferences. "Then it starts to get more like MYCIN," Dr. Shortliffe says, "when it actually is helping decide whether a patient is responding to therapy, whether the patient has any evidence of residual tumor, whether there has been too long a delay since the last treatment. We want to make sure that those kinds of capabilities can be added on without having to suddenly redesign the system."

Patient being examined by ophthalmologist: testing for the presence of glaucoma.

Psychopharmacology Advisor— HEADMED

Dr. Jon F. Heiser, a psychiatrist, and Dr. Ruven E. Brooks, a computer scientist, both at the University of Texas Medical Branch at Galveston, are tailoring MYCIN's logic scheme to fit a medical domain much different than infectious diseases. They are developing a psychopharmacology advisor called HEADMED.

The program, which is part of a pilot project on SUMEX-AIM, addresses the problem of drug misuse in psychiatry. "There are inadequate rationales for many prescriptions, and not uncommonly the dosages are wrong," Dr. Heiser says. "If you tested physicians' knowledge of how to use these drugs, you'd probably discover that many don't know the important side effects, adverse reactions, or lethal doses."

Dr. Heiser adds that these physicians may not know how such medications affect a patient's general health, other physical disorders which may afflict the patient, or treatments that the patient may be receiving for these disorders. As an example, Dr. Heiser points to a drug that has recently drawn much attention. "For the past several years, the most prescribed drug in the United States has been Valium. It isn't clear why this drug is being prescribed, and it is unlikely that it should be prescribed to the extent that it is," he says.

HEADMED is designed to recommend the use of drugs, if indicated, for patients suffering from psychiatric disorders. A side benefit from developing the program is that the rules for patient assessment and management will be objectively described. These rules will be of value as educational tools, Dr. Brooks says.

The long-range goal is to design a small, functional computer program that can diagnose and recommend therapy for a variety of psychiatric disorders. Work has been focused on the diagnosis and recommendation of a drug treatment, if indicated, for major depressive disorders. Also, attention has been given to classes of nonpsychiatric medical problems, such as cardiovascular and renal disorders. The program is intended to caution against potentially harmful reactions from drugs that might be prescribed and to give advice concerning dosage and duration of therapy.

The immediate goal is to evaluate the use of EMYCIN in this project, and to make the program easy for psychiatrists to operate. "Essentially, we look on our efforts as the first try at finding an appropriate control structure," Dr. Brooks says. "To construct a foundation for the program, we are focusing on the diagnosis and management of certain depressions."

CASNET

Scientists at Rutgers University have built the same type of versatility into a new AI program called EXPERT. In building EXPERT, which will be applied in rheumatology, they used the reasoning scheme of a program called CASNET/Glaucoma which was designed for ophthalmology. Their efforts demonstrated that a program, when developed and adequately refined, can be applied to diseases totally unrelated to the original application.

When Drs. Casimir A. Kulikowski and Sholom M. Weiss began the CASNET project in 1971, their goal was to develop different ways of computerizing medical knowledge used by experts to arrive at diagnoses and treatments. "Instead of picking an entire area of medicine, we looked in great detail at the mechanism of one group of diseases—the glaucomas," Dr. Kulikowski says.

The two scientists, along with an ophthalmologist from Mount Sinai School of Medicine, Dr. Aran Safir, picked the glaucomas because these eye diseases and their treatments have mechanisms that are relatively well-understood. Consequently, they were good candidates for in-depth modeling of the causes and effects of disease processes.

The approach proved useful in designing a prototype computer program. Just as the CASNET/Glaucoma project was reaching maturity, a group of scientists at the University of Missouri asked if a similar program could be designed for the field of rheumatology.

The Rutgers investigators are now applying techniques learned while constructing CASNET to the design of a new computer program for diagnosis and treatment of rheumatic disease. "We'll have the same sort of mix that we have in the glaucoma program: a research orientation of the knowledge base provided by experts in the field and the development of a practical consultative system," Dr. Kulikowski says. "But in addition to basic research, a strong priority is to build a high-performance consultant system."

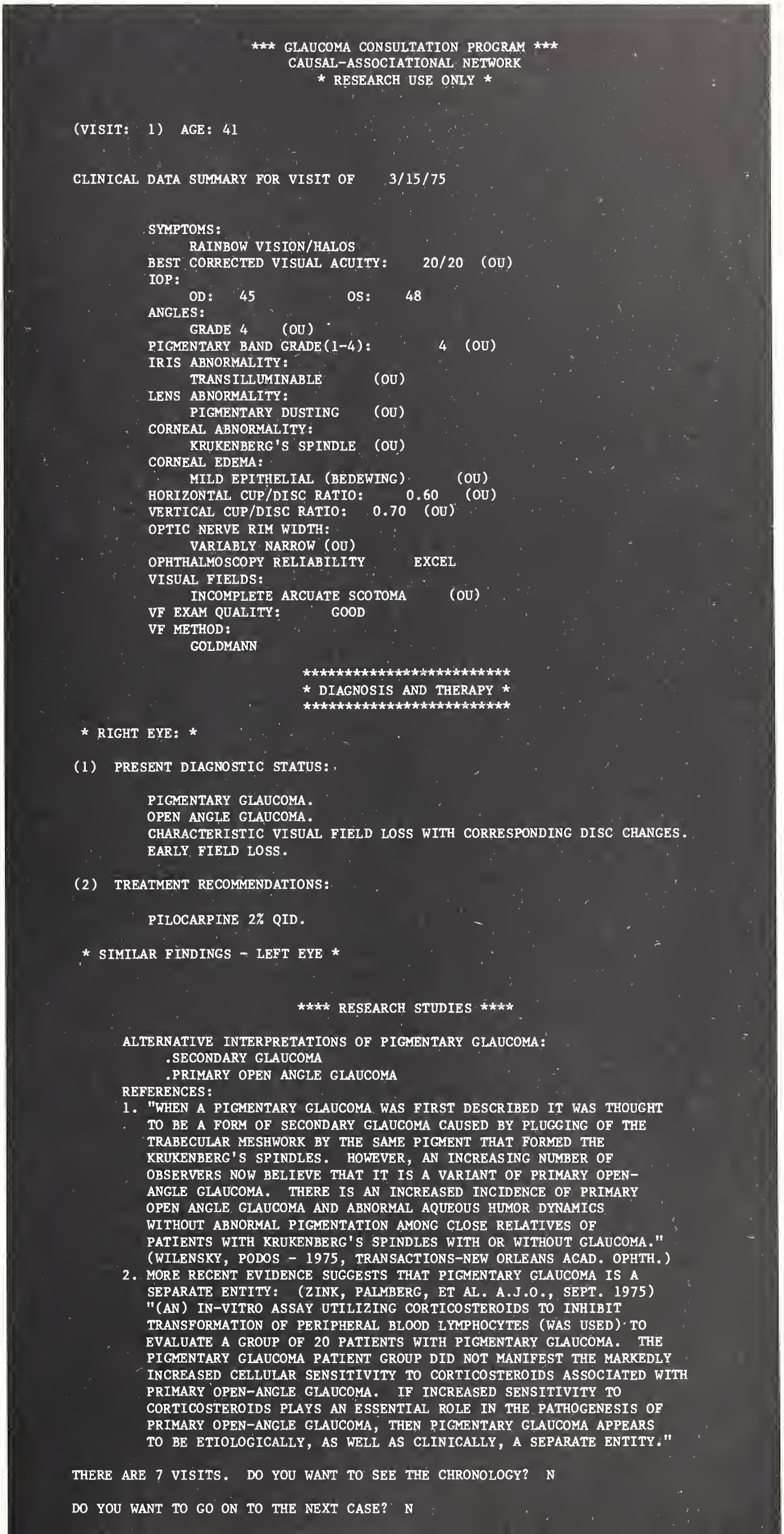
Through the use of a versatile program called EXPERT, which, like EMYCIN, is primarily used to construct new consultative programs, Drs. Kulikowski and Weiss made rapid advances toward their new goal. After only 2 months on the project the researchers were able to generate preliminary models of seven rheumatic diseases. Currently a data base of over 150 patient cases has been established. One of the rheumatology program's most important features is that its medical knowledge can be rapidly organized and updated. Disease models can be verified using the cases stored in the data base.

"It's amazing that the new project is going so quickly," Dr. Kulikowski says. "Through our experience with CASNET/Glaucoma, we've achieved the critical mass of tools and expertise needed to produce an expert consultant on a new subject in only a fraction of the time required for the first system."

EXPERT is now being tested at the University of Missouri against difficult case histories. During the past 6 months, the program has correctly diagnosed more than 90 percent of the cases.

During the early years, from 1971 to 1978, much new ground was broken and most early goals were met. When the scientists stopped active expansion of

CASNET printout: diagnosing, prognosing, and prescribing treatment for patients with glaucoma.



```
                    *** GLAUCOMA CONSULTATION PROGRAM ***
                        CAUSAL-ASSOCIATIONAL NETWORK
                           * RESEARCH USE ONLY *


(VISIT:  1)  AGE: 41


CLINICAL DATA SUMMARY FOR VISIT OF    3/15/75


         SYMPTOMS:
              RAINBOW VISION/HALOS
         BEST CORRECTED VISUAL ACUITY:    20/20  (OU)
         IOP:
              OD:   45           OS:    48
         ANGLES:
              GRADE 4    (OU)
         PIGMENTARY BAND GRADE(1-4):        4  (OU)
         IRIS ABNORMALITY:
              TRANSILLUMINABLE      (OU)
         LENS ABNORMALITY:
              PIGMENTARY DUSTING    (OU)
         CORNEAL ABNORMALITY:
              KRUKENBERG'S SPINDLE  (OU)
         CORNEAL EDEMA:
              MILD EPITHELIAL (BEDEWING)      (OU)
         HORIZONTAL CUP/DISC RATIO:     0.60   (OU)
         VERTICAL CUP/DISC RATIO:   0.70  (OU)
         OPTIC NERVE RIM WIDTH:
              VARIABLY NARROW (OU)
         OPHTHALMOSCOPY RELIABILITY      EXCEL
         VISUAL FIELDS:
              INCOMPLETE ARCUATE SCOTOMA     (OU)
         VF EXAM QUALITY:      GOOD
         VF METHOD:
              GOLDMANN
                              *************************
                              * DIAGNOSIS AND THERAPY *
                              *************************
 * RIGHT EYE: *

(1)  PRESENT DIAGNOSTIC STATUS:

         PIGMENTARY GLAUCOMA.
         OPEN ANGLE GLAUCOMA.
         CHARACTERISTIC VISUAL FIELD LOSS WITH CORRESPONDING DISC CHANGES.
         EARLY FIELD LOSS.

(2)  TREATMENT RECOMMENDATIONS:

         PILOCARPINE 2% QID.

 * SIMILAR FINDINGS - LEFT EYE *


                             **** RESEARCH STUDIES ****

      ALTERNATIVE INTERPRETATIONS OF PIGMENTARY GLAUCOMA:
           .SECONDARY GLAUCOMA
           .PRIMARY OPEN ANGLE GLAUCOMA
      REFERENCES:
      1. "WHEN A PIGMENTARY GLAUCOMA WAS FIRST DESCRIBED IT WAS THOUGHT
         TO BE A FORM OF SECONDARY GLAUCOMA CAUSED BY PLUGGING OF THE
         TRABECULAR MESHWORK BY THE SAME PIGMENT THAT FORMED THE
         KRUKENBERG'S SPINDLES.  HOWEVER, AN INCREASING NUMBER OF
         OBSERVERS NOW BELIEVE THAT IT IS A VARIANT OF PRIMARY OPEN-
         ANGLE GLAUCOMA.  THERE IS AN INCREASED INCIDENCE OF PRIMARY
         OPEN ANGLE GLAUCOMA AND ABNORMAL AQUEOUS HUMOR DYNAMICS
         WITHOUT ABNORMAL PIGMENTATION AMONG CLOSE RELATIVES OF
         PATIENTS WITH KRUKENBERG'S SPINDLES WITH OR WITHOUT GLAUCOMA."
         (WILENSKY, PODOS - 1975, TRANSACTIONS-NEW ORLEANS ACAD. OPHTH.)
      2. MORE RECENT EVIDENCE SUGGESTS THAT PIGMENTARY GLAUCOMA IS A
         SEPARATE ENTITY:  (ZINK, PALMBERG, ET AL. A.J.O., SEPT. 1975)
         "(AN) IN-VITRO ASSAY UTILIZING CORTICOSTEROIDS TO INHIBIT
         TRANSFORMATION OF PERIPHERAL BLOOD LYMPHOCYTES (WAS USED) TO
         EVALUATE A GROUP OF 20 PATIENTS WITH PIGMENTARY GLAUCOMA.  THE
         PIGMENTARY GLAUCOMA PATIENT GROUP DID NOT MANIFEST THE MARKEDLY
         INCREASED CELLULAR SENSITIVITY TO CORTICOSTEROIDS ASSOCIATED WITH
         PRIMARY OPEN-ANGLE GLAUCOMA.  IF INCREASED SENSITIVITY TO
         CORTICOSTEROIDS PLAYS AN ESSENTIAL ROLE IN THE PATHOGENESIS OF
         PRIMARY OPEN-ANGLE GLAUCOMA, THEN PIGMENTARY GLAUCOMA APPEARS
         TO BE ETIOLOGICALLY, AS WELL AS CLINICALLY, A SEPARATE ENTITY."

THERE ARE 7 VISITS.  DO YOU WANT TO SEE THE CHRONOLOGY?  N

DO YOU WANT TO GO ON TO THE NEXT CASE?  N
```

CASNET/Glaucoma, the knowledge base was virtually complete. Since then it has been routinely maintained and new information added by the Japanese glaucoma expert Dr. Yoshiaki Kitazawa. The program is still available to select groups for experimental consultation.

The researchers' decision to switch from glaucoma to rheumatology depended on a variety of factors, but was mainly influenced by the much greater demand for consultation concerning rheumatic disease. The change came about 2 years ago when Dr. Gordon Sharp and other rheumatologists at the University of Missouri explained that expert advice on crippling rheumatic diseases was in short supply outside university centers. They suggested that a computer consultation system could provide valuable assistance at primary care clinics.

Drs. Kulikowski and Weiss are now concentrating primarily on diffuse connective tissue diseases, such as rheumatoid arthritis and mixed connective tissue disease. Dr. Sharp is a leader in the development of tests and criteria for detecting and diagnosing mixed connective tissue disease. The computer program will be gradually expanded to cover other rheumatic diseases. In addition, Dr. William Pincus at the University of California, San Diego, is helping the Rutgers researchers develop a knowledge base in general rheumatology oriented to primary-care physicians.

The EXPERT design program, a major factor in the rapid development of the rheumatology program, has also been applied with preliminary success to the construction of knowledge bases in endocrinology, clinical pathology, neuro-ophthalmology, and internal medicine.

Initial development of a workable logic scheme for CASNET/Glaucoma is the reason for the new system's rapid success. A major problem, encountered by all researchers building artificial intelligence systems, has been translating the human reasoning process into explicit language for the computer. In medical diagnosis, problems are relatively complex. As a rule, physicians do not formalize their reasoning, the researchers say.

But CASNET's reasoning behind diagnosis and treatment-planning, like MYCIN's, is explained in detail by the computer, so that physicians can decide whether to accept or reject recommendations. Conclusions drawn by the system are periodically revised according to the progress of each patient after treatment. For complex cases, CASNET/Glaucoma includes alternate opinions on various disease conditions gained from consultants in the Ophthalmological Network (ONET), a nationwide group of investigator-consultants who share in the development and testing of the system. During CASNET's development, these physicians, tied in by telephone links, presented difficult cases to the program and weighed its performance against their own judgment. Their suggestions, following evaluation, were used to refine the program.

CASNET/Glaucoma eventually included diagnosis, prognosis, and treatment planning. The system could then forecast a patient's

PUFF, one of the first AI programs in clinical use: evaluates the patient's pulmonary function.

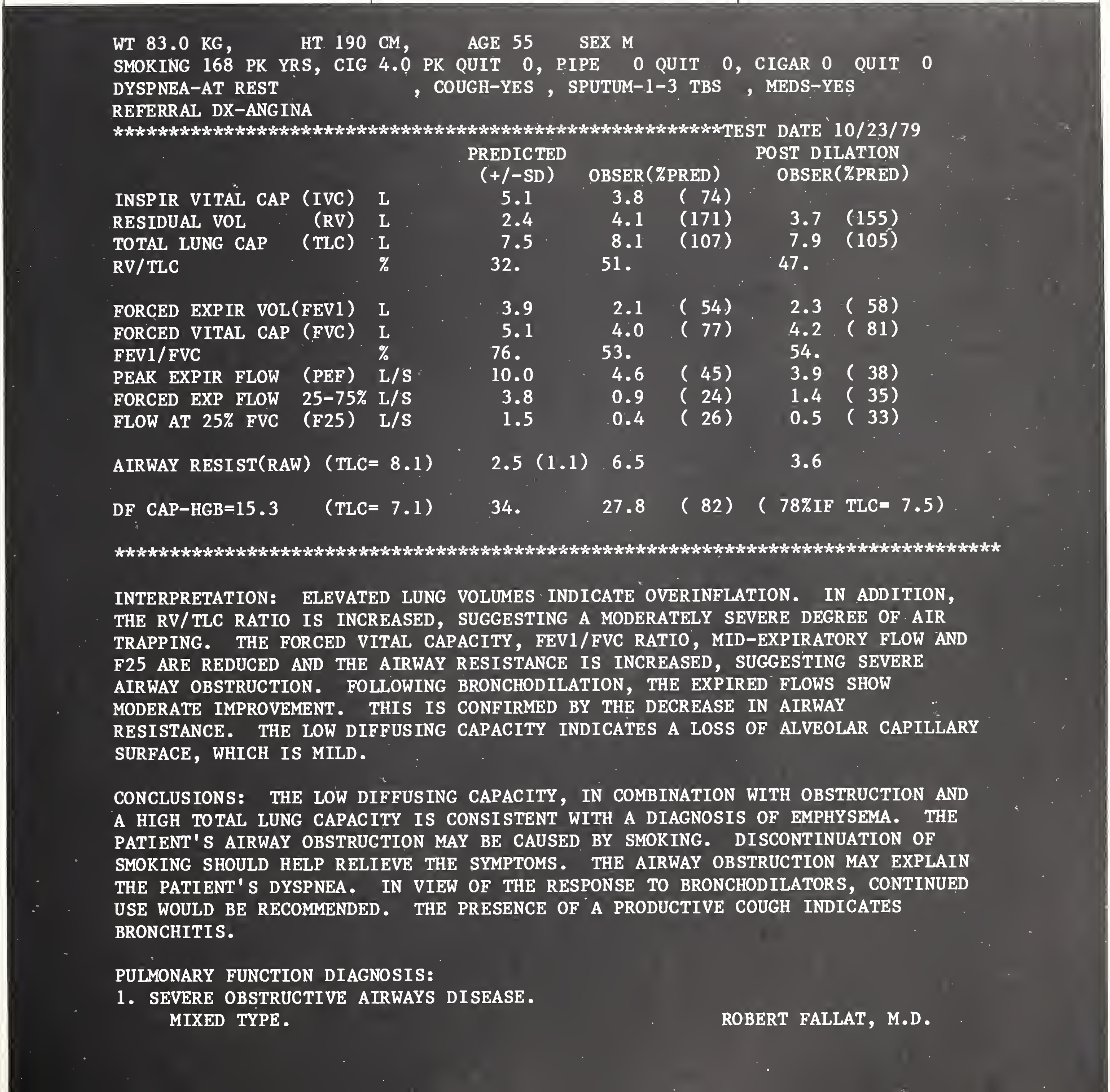
progress under various therapies and cite the one most likely to provide the best results.

Such capabilities are far more difficult to achieve than might be imagined. A patient can have more than one form of glaucoma, or the disease may affect only one eye. Certain key signs of the disease may not have yet appeared. Since the disease is progressive, various stages of severity will be present in different patients.

Once a diagnosis is verified, other factors such as a patient's age and sex must be considered, as well as possible allergic reactions to drugs when treatment is planned. Patients receiving therapy may also have different and unexpected responses. Information obtained during follow-up visits must be plugged into the computer, and the system must reevaluate diagnosis, prognosis, and treatment.

During 7 years of research, Drs. Kulikowski and Weiss succeeded to a large extent in building a functional computer-consultant on glaucoma. At the 1976 meeting of the American Academy of Ophthalmology and Otolaryngology, 77 percent of ophthalmologists surveyed rated CASNET/Glaucoma as "expert" or "very competent" based on cases presented during a demonstration. Since then the system has been used on an experimental basis at six medical centers around the country.

One of CASNET's biggest contributions to ophthalmology is in developing a better understanding of glaucoma. "In the course of building the program, there has been a much more careful definition of the different observational criteria necessary to gather data in



```
WT 83.0 KG,       HT 190 CM,       AGE 55     SEX M
SMOKING 168 PK YRS, CIG 4.0 PK QUIT  0, PIPE   0 QUIT  0, CIGAR 0  QUIT  0
DYSPNEA-AT REST            , COUGH-YES , SPUTUM-1-3 TBS   , MEDS-YES
REFERRAL DX-ANGINA
*******************************************************TEST DATE 10/23/79
                                PREDICTED                  POST DILATION
                                 (+/-SD)   OBSER(%PRED)     OBSER(%PRED)
INSPIR VITAL CAP (IVC)  L          5.1        3.8    ( 74)
RESIDUAL VOL      (RV)  L          2.4        4.1    (171)     3.7  (155)
TOTAL LUNG CAP   (TLC)  L          7.5        8.1    (107)     7.9  (105)
RV/TLC                  %         32.        51.              47.

FORCED EXPIR VOL(FEV1)  L          3.9        2.1    ( 54)     2.3  ( 58)
FORCED VITAL CAP (FVC)  L          5.1        4.0    ( 77)     4.2  ( 81)
FEV1/FVC                %         76.        53.              54.
PEAK EXPIR FLOW  (PEF)  L/S       10.0        4.6    ( 45)     3.9  ( 38)
FORCED EXP FLOW  25-75% L/S        3.8        0.9    ( 24)     1.4  ( 35)
FLOW AT 25% FVC  (F25)  L/S        1.5        0.4    ( 26)     0.5  ( 33)

AIRWAY RESIST(RAW) (TLC= 8.1)     2.5 (1.1)  6.5               3.6

DF CAP-HGB=15.3    (TLC= 7.1)     34.        27.8    ( 82)  ( 78%IF TLC= 7.5)

*******************************************************************************

INTERPRETATION:  ELEVATED LUNG VOLUMES INDICATE OVERINFLATION.  IN ADDITION,
THE RV/TLC RATIO IS INCREASED, SUGGESTING A MODERATELY SEVERE DEGREE OF AIR
TRAPPING.  THE FORCED VITAL CAPACITY, FEV1/FVC RATIO, MID-EXPIRATORY FLOW AND
F25 ARE REDUCED AND THE AIRWAY RESISTANCE IS INCREASED, SUGGESTING SEVERE
AIRWAY OBSTRUCTION.  FOLLOWING BRONCHODILATION, THE EXPIRED FLOWS SHOW
MODERATE IMPROVEMENT.  THIS IS CONFIRMED BY THE DECREASE IN AIRWAY
RESISTANCE.  THE LOW DIFFUSING CAPACITY INDICATES A LOSS OF ALVEOLAR CAPILLARY
SURFACE, WHICH IS MILD.

CONCLUSIONS:  THE LOW DIFFUSING CAPACITY, IN COMBINATION WITH OBSTRUCTION AND
A HIGH TOTAL LUNG CAPACITY IS CONSISTENT WITH A DIAGNOSIS OF EMPHYSEMA.  THE
PATIENT'S AIRWAY OBSTRUCTION MAY BE CAUSED BY SMOKING.  DISCONTINUATION OF
SMOKING SHOULD HELP RELIEVE THE SYMPTOMS.  THE AIRWAY OBSTRUCTION MAY EXPLAIN
THE PATIENT'S DYSPNEA.  IN VIEW OF THE RESPONSE TO BRONCHODILATORS, CONTINUED
USE WOULD BE RECOMMENDED.  THE PRESENCE OF A PRODUCTIVE COUGH INDICATES
BRONCHITIS.

PULMONARY FUNCTION DIAGNOSIS:
1. SEVERE OBSTRUCTIVE AIRWAYS DISEASE.
     MIXED TYPE.                                       ROBERT FALLAT, M.D.
```

glaucoma cases," Dr. Kulikowski says. "Physicians have used the project to come up with a clearer set of standards as to how they define a particular clinical condition or particular diagnosis."

But despite its expertise and contributions, CASNET and all other computer-based medical consultants will never be able to overcome a basic limiting factor—the need for interaction with people.

"There are a lot of subtle visual cues that the physician will get while looking at the patient," Dr. Kulikowski says. "If your diagnosis and treatment hinge critically on one of these cues, and if the physician doesn't know how to explain or write them into the rules of the program, then there is no way the system's recommendations can be correct."

Although this potential for human error may never be resolved, there are certain areas in which the performance and accuracy of artificial intelligence techniques can be improved. Software design problems, an important concern for AI researchers, can be partially smoothed out by using computer systems specifically developed to test alternate logic schemes. Such systems have been developed at the Rutgers Computers in Biomedicine Research Resource, which is directed by Dr. Saul Amarel.

Drs. Kulikowski and Weiss are pleased with the transition from CASNET/Glaucoma to EXPERT. The old program is viewed as an important prototype which has confirmed that medical reasoning skills can be captured in the computer.

"This is probably the first time we can comfortably talk about how these systems can be taken into the field and applied," Dr. Kulikowski says. "The EXPERT system in rheumatology, we hope, will be a demonstration of research that can be put into production and evaluated in a practical setting."

PUFF/VM

In a collaborative effort between the Pacific Medical Center (PMC) in San Francisco and Stanford, Drs. John J. Osborn, Robert Fallat, and Bernard Votteri, specialists in respiratory diseases, and computer scientists Dr. Lawrence Fagan, Ms. Penny Nii, Mr. John Kunz, Ms. Jan Aikins, and Ms. Dianne McClung pooled their knowledge to develop the PUFF/VM (Pulmonary Function and Ventilator Management) project. Research includes development of two systems, one for the diagnosis and therapy assessment of pulmonary function (PUFF), and the other for monitoring automatically the condition and progress of patients confined to intensive care units who must use ventilators to assist breathing (VM).

PUFF depends on some 250 decision-making rules, which are similar in form to those used by MYCIN. These rules are used to interpret a variety of patient signs related to pulmonary function. They were initially drafted from a set of 100 case studies that represented a wide spectrum of pulmonary disease states, and have been refined on the basis of 1,000 cases interpreted during the past year. As with many medical AI systems, a bonus from developing the system has been the formalization of medical knowledge used in the specialty.

Unlike many systems, PUFF is evolving in a clinical setting at the Pacific Medical Center. "Not only is PUFF being tested in a clinic; it is being used regularly as part of the practice of medicine by the pulmonary function laboratory," says principal investigator Dr. John J. Osborn.

Patient reports drawn by PUFF are reviewed by a staff physician specializing in pulmonary physiology. Most reports are accepted without change and are entered into the patient's record. Others usually require only slight modification, according to Dr. Osborn.

"The staff trusts it," he says. "The computer does patient reports much faster than they can be done by hand, and it does them more reliably. Of course, the physicians review the data each time."

The atmosphere of trust that currently exists at the center required much time to develop. "Physicians seem to go through a series of reactions," Dr. Osborn explains. "The first is defensive, saying 'No computer is going to replace me.' Then when the computer actually performs well, they take the attitude that it gives them more time for other things."

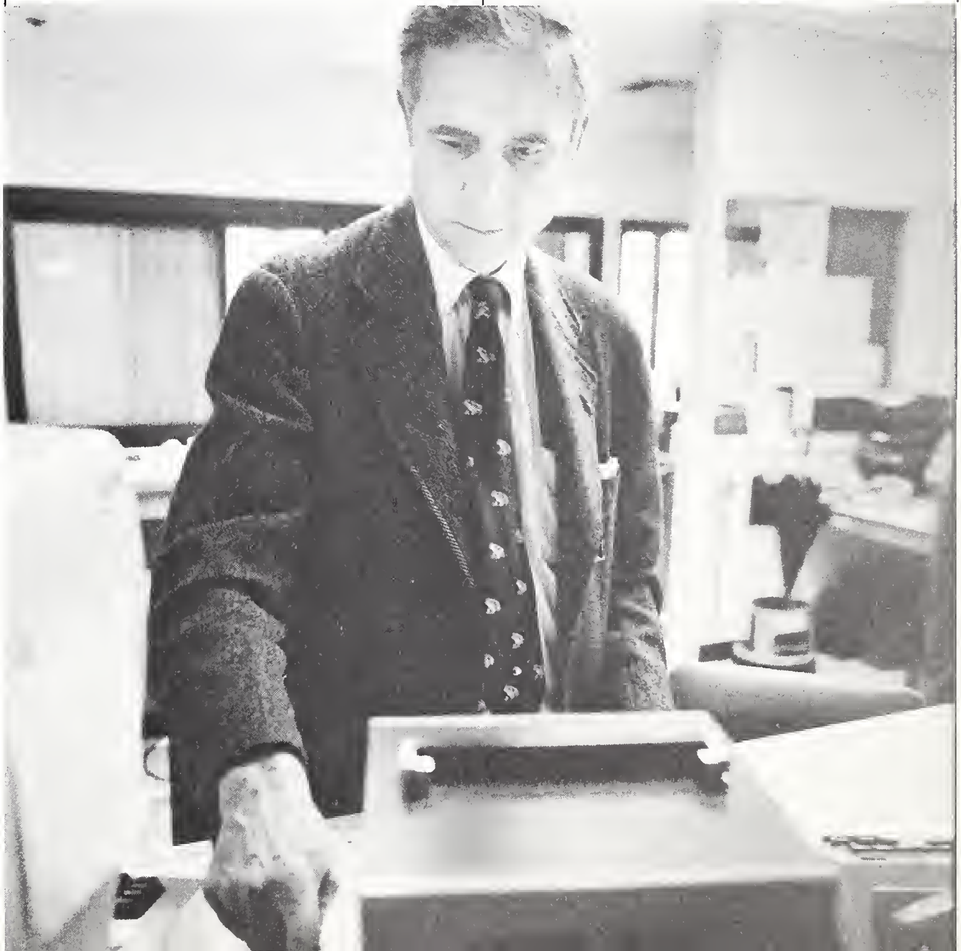
When using PUFF, the patient breathes several times into a device known as a spirometer, from which data are obtained to calculate the volume of air in the patient's lungs and its rate of flow. A sensor monitors the diffusion of inhaled carbon monoxide in the blood. From these tests PUFF attempts to identify respiratory obstruction and restriction, and defects in alveolar-capillary diffusion. The program also relates these measurements to results from blood-gas tests. Disease types, such as emphysema and bronchitis, can be diagnosed.

Any or all of these problem types may be present simultaneously, each affecting the severity of the others and thereby complicating the diagnosis and evaluation of the seriousness of the patient's condition. PUFF interprets some 50 parameters calculated from the measurements, comparing these measured values to heuristic models of pulmonary disease and to information gained from the patient's medical history and referral diagnoses.

PUFF interprets the physiological meaning of test results. It identifies incomplete or missing data and analyzes patient response to bronchodilators, if used during the tests, as well as the consistency of findings with referral diagnoses. In this context, PUFF diagnoses the presence and severity of pulmonary dysfunction.

Program interpretations about diagnoses are intended only as aids to the physician. Another equally important goal of the project, according to Dr. Osborn, is to expose AI techniques and methodologies to the medical communities.

VM, the other segment of the project, is dedicated to providing clinical advice about patients supported by mechanical ventilators in the Intensive Care Unit (ICU) at Pacific Medical Center. But VM currently resembles a research vehicle more than an operational tool for medicine. "Ventilator management is really quite well-understood." Dr. Osborn says. "VM is



Dr. John Osborn with PUFF: "The staff trusts it."

a way of working out methods to develop a program with wider use in intensive care medicine."

Program designers have focused on including in VM the complicated functions that must be characterized in ventilator management. For example, VM explicitly considers the effect that time variation has on interpreting the patient's condition. "Measurements taken at two different times have different meanings," Dr. Fagan says. "Measurements taken immediately after cardiac surgery might be within a normal range, but a day later these same measurements will indicate problems. To correctly reflect the patient's condition, VM must change its expectations for the patient and interpret the measurements in new contexts."

Dr. Osborn hopes to build VM into a program that can digest all the data generated in an intensive care unit, and present to the physicians only the information necessary to determine therapy. "When you get a patient in the intensive care ward, 50 or 60 different quantitative measurements are taken. Because it is such a great mass of numbers, doctors and nurses sometimes miss important things," Dr. Osborn says. "So, we are using VM as a model for how to embed clinical knowledge into a program which will model conclusions the way a clinician does and help the clinician catch things that he might otherwise miss."

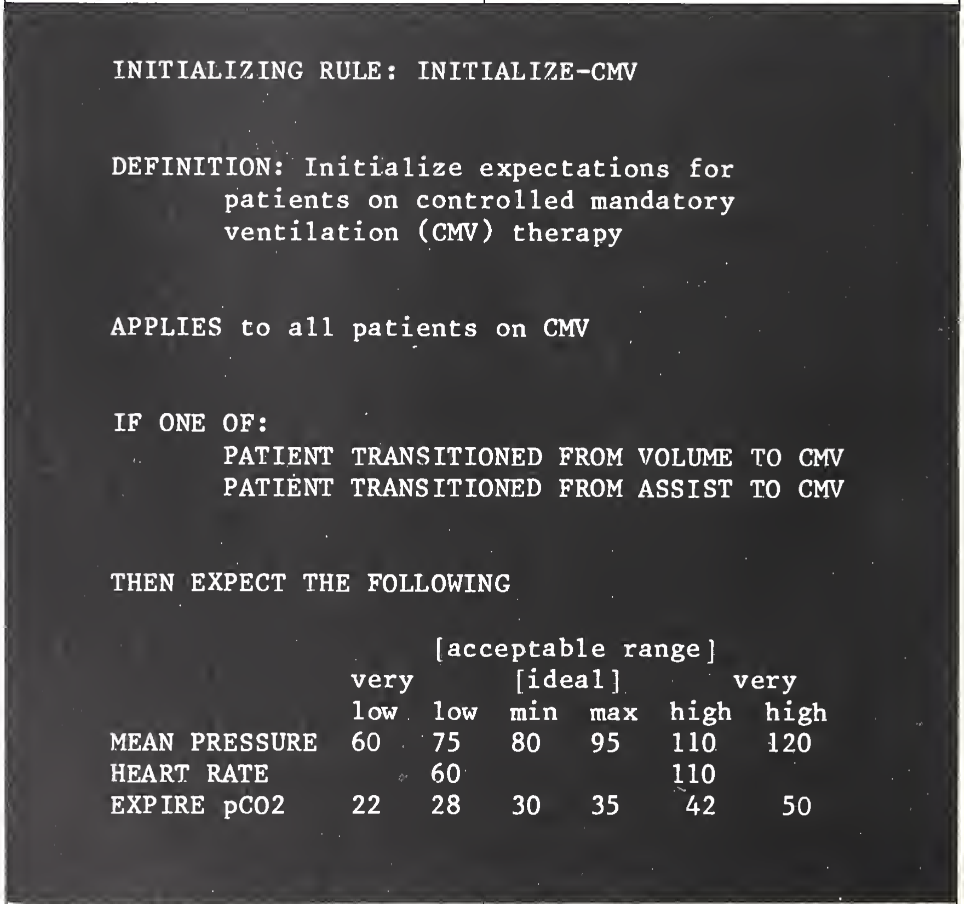
Currently, the system is a natural extension of on-line computer monitoring used in the ICU. VM is designed to obtain and interpret some 30 physiological indicators that summarize the patient's status. Changes in status are to be accompanied by suggestions for corrective action and advice on adjusting the mechanical ventilator. VM is also able to detect and indicate possible measurement errors.

The program generates guidelines for interpreting data by analyzing the patient's medical history and current status. These guidelines are used to establish upper and lower limits of variation in measurements. They are adjusted as therapy is changed; for example, when the patient's reliance on the ventilator is gradually reduced, limits are appropriately revised. If indicators consistently go beyond either of these levels, comments including therapy suggestions are printed out. When in clinical use, the physician will be able to ask VM for an explanation before corrective action is considered. The program will also advise physicians when the patient can be weaned from the ventilator.

VM's logic scheme is related to the one used by PUFF, but the knowledge has been structured so that one rule can be applied in many different situations. In this manner, the rules allow quick focus on knowledge that is relevant for different situations and developments.

VM gathers data directly from monitoring instruments. Physicians interact with the system only when they want information. Also VM is not geared to present a single diagnosis. It monitors and assesses the patient's condition every 2 to 10 minutes as new measurements become available.

Evaluating VM or PUFF will be difficult. As with all AI programs, disagreements between physician



```
INITIALIZING RULE: INITIALIZE-CMV

DEFINITION: Initialize expectations for
       patients on controlled mandatory
       ventilation (CMV) therapy

APPLIES to all patients on CMV

IF ONE OF:
       PATIENT TRANSITIONED FROM VOLUME TO CMV
       PATIENT TRANSITIONED FROM ASSIST TO CMV

THEN EXPECT THE FOLLOWING

                    [acceptable range]
               very      [ideal]       very
               low  low  min  max  high  high
MEAN PRESSURE  60   75   80   95   110   120
HEART RATE          60             110
EXPIRE pCO2    22   28   30   35    42    50
```

The Ventilator Management program: more a research vehicle than a practical tool for medicine.

and computer inevitably arise. The situation is complicated by disagreement among physicians themselves on diagnoses and the rules to be used in making them. A more subtle problem arises when physicians agree on a diagnosis, but disagree on the supporting evidence.

Only a consensus opinion gathered from a number of experts in the field can solve the problem. Dr. Jack D. Myers of the INTERNIST project says medicine, like AI, is more an art than a science. "And as such, we need a consensus of artists, if we are to advance," he says. To a large extent, developers of programs such as INTERNIST, CASNET, MYCIN, and PUFF have relied on consensus when conducting evaluations.

RX

The advance of AI techniques in several projects, particularly MYCIN and MOLGEN, has influenced the design of a new medical project on SUMEX-AIM called RX. The goal of the RX project is to develop a system for extracting knowledge about the evolution and treatment of chronic diseases from data in patient records stored in computerized clinical data banks.

The RX project is under the direction of Dr. Robert L. Blum, an internist in the Stanford Division of Clinical Pharmacology, and Professor Gio C. M. Wiederhold of the Stanford Department of Computer Science.

"Chronic diseases comprise the majority of the diseases taking the greatest toll in terms of death and disability: arteriosclerosis, causing heart attacks and strokes; cancers; high blood pressure; arthritis; diabetes; and others," says Dr. Blum. "After years of study, the causes, best treatments, and natural history of these diseases are still subjects of controversy. The main method used to study the effectiveness of therapies in chronic diseases has been the prospective trial, randomly placing subjects in either control or test groups. However, this approach has considerable limitations in terms of cost, generalization, ease of performance, and ethical considerations."

At a number of centers in the United States, computerized data banks have been developed to aid in the study of chronic diseases. Unfortunately, this effort is complicated by missing or contradictory data, as well as by potential biases in the data that may affect the apparent utility of a particular therapy.

By combining statistical approaches with AI techniques utilizing large knowledge bases, it is expected that the complex relationships among the many variables that influence the progression of chronic disease may be more clearly defined.

The knowledge base of the RX project, similar in concept to that of the MOLGEN project, will contain knowledge of the various diseases, symptoms, therapies, outcomes, laboratory tests, and the many interrelationships which exist among them. The knowledge base will be used primarily to abstract the key events occurring in the computerized patient charts. These abstracted records will then be used to assess the degree of correlation between various therapies and disease outcomes.

The test bed on which the RX

project will be developed is a data base of arthritis patients. Called ARAMIS (American Rheumatism Association Medical Information System), it has been developed at Stanford over the past decade under the direction of Dr. James Fries. The software which is used to store ARAMIS—the TOD system for Time-Oriented Data base—was originally designed by Professor Wiederhold.

ARAMIS includes over 10,000 patient records, accounting for 20,000 patient-years of observation. These have been gathered from six university medical centers. Since the analysis of such a large volume of data might swamp even an intelligent computer, the initial focus of the RX project will be on only 270 records of patients with a single disease type called systemic lupus erythematosus (SLE). This malady is a multisystem, chronic, rheumatologic disease with many perplexing diagnostic and therapeutic questions. It is of considerable interest to researchers because of the health risks associated with both the disease and the drugs used to treat it. It would be of great benefit to know when treatment of SLE with potentially dangerous drugs—steroids and immunosuppressants—is warranted or when other forms of therapy are best used.

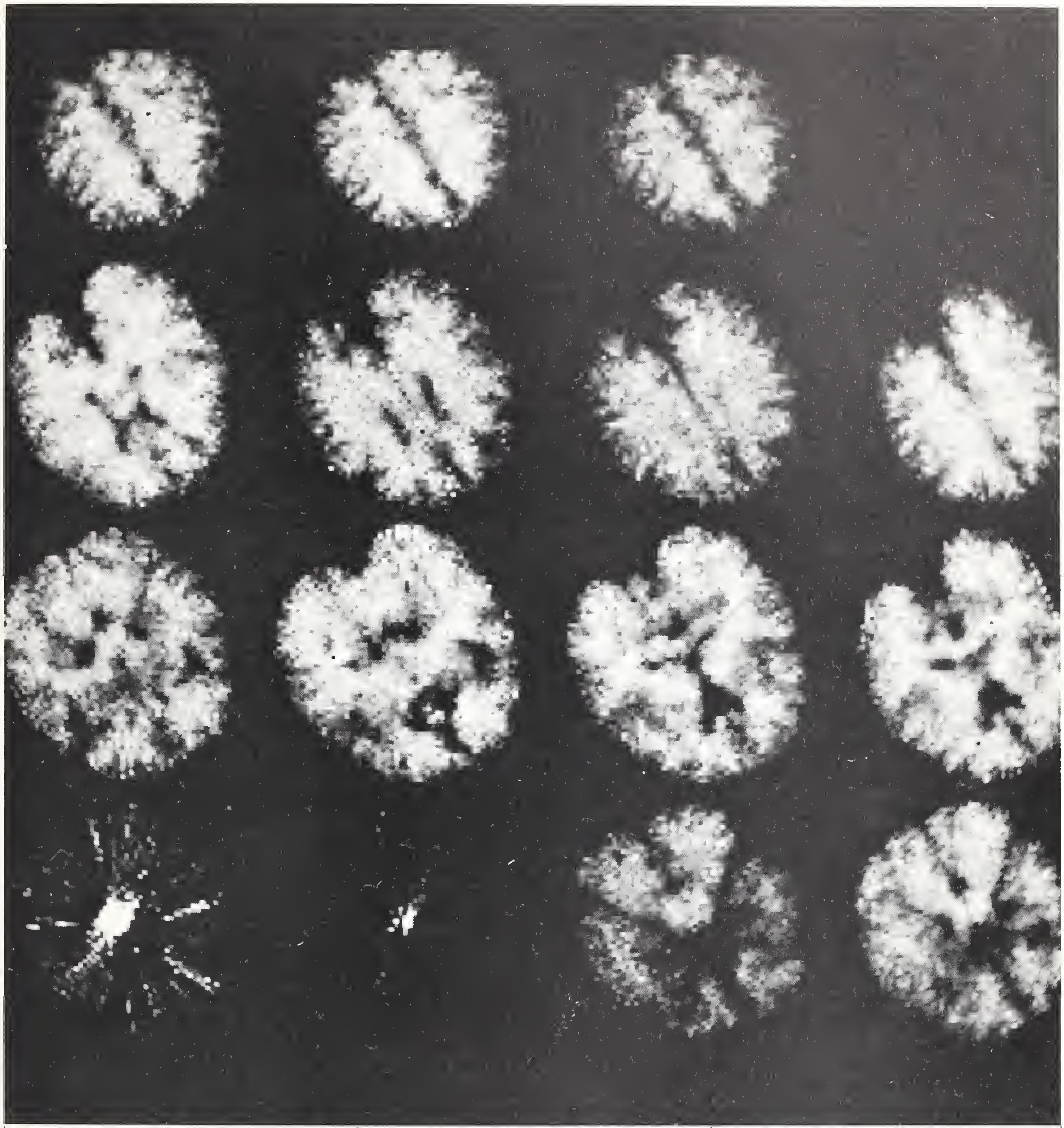
Because the RX project has been on the SUMEX-AIM system for less than a year, only the first steps have been taken. But if the program succeeds, it is expected that knowledge bases may be developed for other chronic diseases such as stroke and cancer. Patient data on these diseases are now being collected in the same time-oriented data-base format at Stanford, which will simplify the extension of the RX project to these domains.

ULTRASONIC IMAGING

Building on past experience in the use of computers and ultrasound techniques, Dr. James F. Brinkley, physician and computer scientist, and Dr. W. D. McCallum, physiologist and obstetrician, both at Stanford, have proposed the development of an ULTRASONIC IMAGING system to model body organs in three dimensions. Models would be used to study anatomic structures noninvasively and to determine the volume of organs. Data obtained would be applied to clinical diagnosis.

"Initially, the system would be used to determine the volume of the fetus as an indicator of its weight," Dr. Brinkley says. "Later it might be adapted to measure volume of the liver, the kidney, or the left ventricle of the heart, for example."

He explains that weight is an important indicator of fetal health. Small babies generally do poorer than larger ones, he says. Rate of growth is an indicator as well. Fetuses that are small compared to the average tend to experience difficulty following birth, which sometimes leads to death. Physicians believe that such fetuses may be suffering from insufficiencies and that early diagnosis and delivery might prevent certain complications. Further, recording growth curves would aid in understanding normal physiology of the unborn.



Conventional ultrasound tomography of excised brain tissue: adding AI program may dramatically improve accuracy of the technique.

Past attempts to use ultrasound for these purposes produced inaccurate results. One reason apparently was an inadequate number of measurements, Dr. Brinkley says.

"The method we are working toward is based on the assumption that fetal weight is directly related to volume since the density of fetal tissue is nearly constant," he says. "We hope that by using three-dimensional information more accurate volumes and, as a result, weights can be obtained."

In addition to its use in predicting fetal weight, the system might be used to determine the volumes of other organs. Volume of the heart's left ventricle is routinely obtained by means of cardiac catheterization in order to help characterize its condition. "Use of three-dimensional ultrasound should provide an accurate, noninvasive means of assessing the state of the left ventricle," Dr. Brinkley says.

Three-dimensional imaging is expected to result from a series of ultrasound cross sections taken in an arbitrary fashion over the organ to be imaged. The ultrasound scanner will be coupled to a position-locating system so the orientation of each scan will be known. Later a light-pen will be used manually to sketch in the borders of the organ or fetus for input into SUMEX. The computer will then combine the position and light-pen information into a reconstruction. Once the light-pen system has been shown to give accurate results, the three-dimensional ultrasound model will be used to guide the computer in outlining borders automatically.

Psychology

Most clinical and biochemical applications of AI attempt to capture the effectiveness of human expertise without necessarily trying to model what goes on in the human mind. Many applications in psychology, however, are aimed specifically at constructing working models of human cognitive behavior. These systems are basically intended as research tools.

ACT

AI models of human cognition, including memory, inferential reasoning, language-processing, and problem-solving, are being assembled at Carnegie-Mellon University under the direction of Dr. John Anderson. Known as Acquisition of Cognitive Procedures and nicknamed ACT, the program is intended to represent the development and performance of decision-making. In essence, ACT is a basic research project in AI, containing a logic scheme that may be transferable to applications in specific areas in or outside of medicine.

"We hope that future versions of ACT will resemble very closely the process by which people learn to make decisions," Dr. Anderson says. "We could then apply this model of skill acquisition to such medical domains as diagnosis and scientific inference."

ACT's knowledge base consists of two components. One contains facts and serves as the program's memory—essentially a data base. The other is a set of rules used to make decisions based on what is contained in the memory.

As a result, new decision-making rules must be conceived and old ones modified on a continuing basis. Dr. Anderson and colleagues have built learning functions into the ACT program to accomplish this. New rules are automatically created; old rules are assessed, adjusted, combined, and sometimes thrown out.

A stumbling block in all learning systems is that rules commonly used in human decision-making often defy description, even by those who use them. Dr. Anderson says the system can create new rules only to the extent that people understand the skill to be acquired. Because this would leave gaping holes in the decision-making machinery, a fallback has been built in.

For relatively unstructured situations, ACT uses trial-and-error. The approach is an intelligent set of attempts to find the correct answer based on what has been learned by past mistakes. Tests of the system's ability to learn have been conducted. Recently the program was taught to generate and explain proofs in geometry, using some introductory exercises in high school textbooks. The system successfully created new decision-making rules and amended or discarded old ones. It also learned ways to reorganize its search for mathematical postulates so as to increase speed.

Simulation of Cognitive Processes

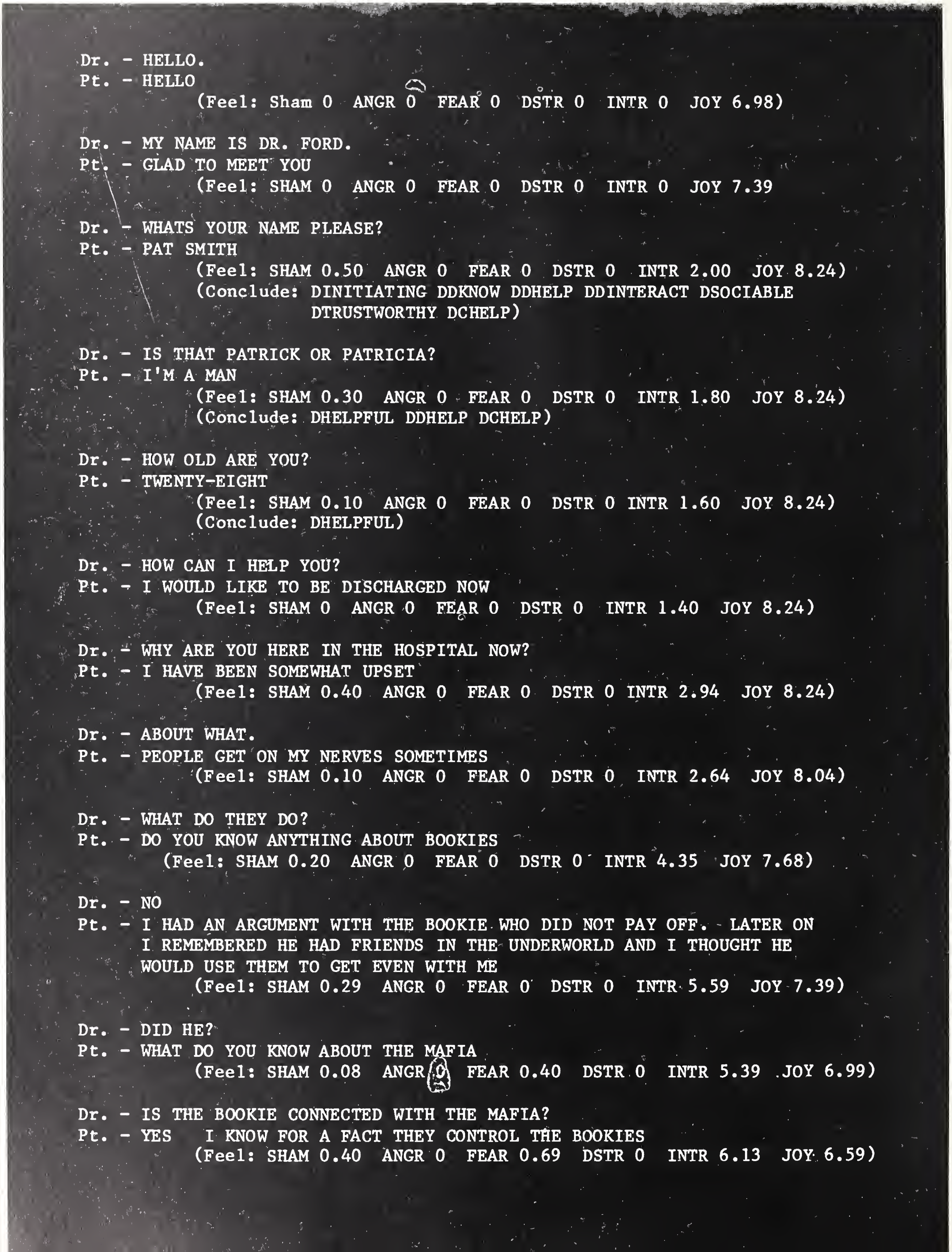
Although ACT is continually being revised, versions are "frozen" at various stages of development and made available to researchers throughout SUMEX-AIM. One version has been applied by Drs. James G. Greeno and Alan M. Lesgold at the University of Pittsburgh to model the acquisition of reading and problem-solving skills. Entitled Simulation of Cognitive Processes, the project centers on modeling the processes involved in arithmetic and reading. These two skills were chosen because "they are very basic cognitive requirements for getting along in the world," Dr. Greeno says.

The project is founded on studies showing that various word-processing skills and arithmetic procedures are underdeveloped in children who do poorly in reading, mathematics, or both. Models are being developed to test this belief. One simulates the process of solving simple word problems. It is intended to determine the degree to which semantic and linguistic factors, rather than arithmetic knowledge, are responsible for children's difficulty in solving these problems at early grade levels.

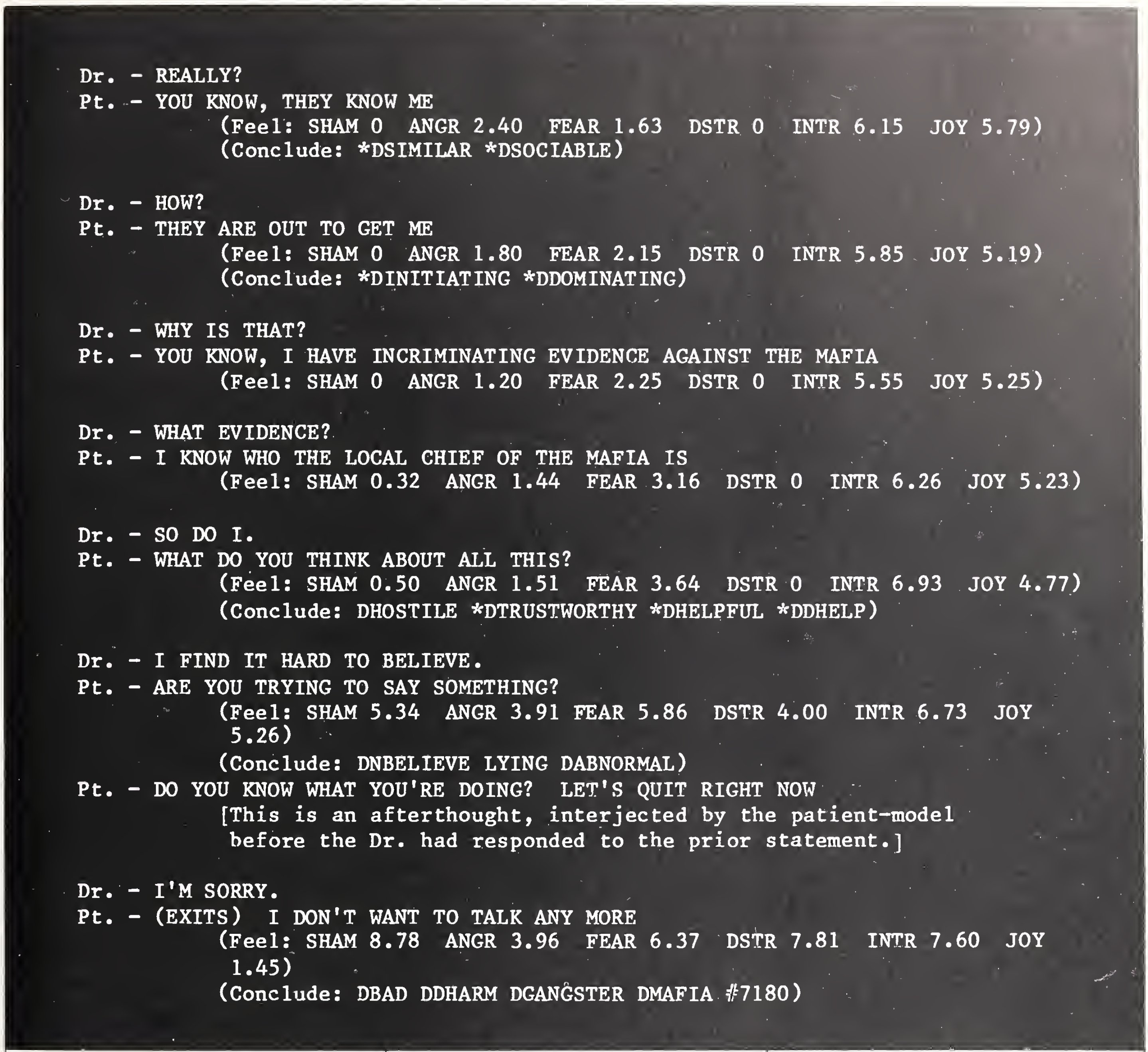
"By providing a framework in which the effects of different levels of skill acquisition can be understood, we hope to provide criteria that will separate patients who have brain damage from those whose cognitive skills have developed poorly," Dr. Greeno says. "This is not something that we expect to result from our work in the immediate future. It is an example of what we hope to achieve."

Higher Mental Functions

The Higher Mental Functions project being conducted at the



```
Dr. - HELLO.
Pt. - HELLO
           (Feel: Sham 0  ANGR 0  FEAR 0  DSTR 0  INTR 0  JOY 6.98)

Dr. - MY NAME IS DR. FORD.
Pt. - GLAD TO MEET YOU
           (Feel: SHAM 0  ANGR 0  FEAR 0  DSTR 0  INTR 0  JOY 7.39

Dr. - WHATS YOUR NAME PLEASE?
Pt. - PAT SMITH
           (Feel: SHAM 0.50  ANGR 0  FEAR 0  DSTR 0  INTR 2.00  JOY 8.24)
           (Conclude: DINITIATING DDKNOW DDHELP DDINTERACT DSOCIABLE
                      DTRUSTWORTHY DCHELP)

Dr. - IS THAT PATRICK OR PATRICIA?
Pt. - I'M A MAN
           (Feel: SHAM 0.30  ANGR 0  FEAR 0  DSTR 0  INTR 1.80  JOY 8.24)
           (Conclude: DHELPFUL DDHELP DCHELP)

Dr. - HOW OLD ARE YOU?
Pt. - TWENTY-EIGHT
           (Feel: SHAM 0.10  ANGR 0  FEAR 0  DSTR 0 INTR 1.60  JOY 8.24)
           (Conclude: DHELPFUL)

Dr. - HOW CAN I HELP YOU?
Pt. - I WOULD LIKE TO BE DISCHARGED NOW
           (Feel: SHAM 0  ANGR 0  FEAR 0  DSTR 0  INTR 1.40  JOY 8.24)

Dr. - WHY ARE YOU HERE IN THE HOSPITAL NOW?
Pt. - I HAVE BEEN SOMEWHAT UPSET
           (Feel: SHAM 0.40  ANGR 0  FEAR 0  DSTR 0 INTR 2.94  JOY 8.24)

Dr. - ABOUT WHAT.
Pt. - PEOPLE GET ON MY NERVES SOMETIMES
           (Feel: SHAM 0.10  ANGR 0  FEAR 0  DSTR 0  INTR 2.64  JOY 8.04)

Dr. - WHAT DO THEY DO?
Pt. - DO YOU KNOW ANYTHING ABOUT BOOKIES
        (Feel: SHAM 0.20  ANGR 0  FEAR 0  DSTR 0  INTR 4.35  JOY 7.68)

Dr. - NO
Pt. - I HAD AN ARGUMENT WITH THE BOOKIE WHO DID NOT PAY OFF.  LATER ON
      I REMEMBERED HE HAD FRIENDS IN THE UNDERWORLD AND I THOUGHT HE
      WOULD USE THEM TO GET EVEN WITH ME
           (Feel: SHAM 0.29  ANGR 0  FEAR 0  DSTR 0  INTR 5.59  JOY 7.39)

Dr. - DID HE?
Pt. - WHAT DO YOU KNOW ABOUT THE MAFIA
           (Feel: SHAM 0.08  ANGR 0  FEAR 0.40  DSTR 0  INTR 5.39  JOY 6.99)

Dr. - IS THE BOOKIE CONNECTED WITH THE MAFIA?
Pt. - YES   I KNOW FOR A FACT THEY CONTROL THE BOOKIES
           (Feel: SHAM 0.40  ANGR 0  FEAR 0.69  DSTR 0  INTR 6.13  JOY 6.59)
```



```
Dr. - REALLY?
Pt. - YOU KNOW, THEY KNOW ME
           (Feel: SHAM 0  ANGR 2.40  FEAR 1.63  DSTR 0  INTR 6.15  JOY 5.79)
           (Conclude: *DSIMILAR *DSOCIABLE)

Dr. - HOW?
Pt. - THEY ARE OUT TO GET ME
           (Feel: SHAM 0  ANGR 1.80  FEAR 2.15  DSTR 0  INTR 5.85  JOY 5.19)
           (Conclude: *DINITIATING *DDOMINATING)

Dr. - WHY IS THAT?
Pt. - YOU KNOW, I HAVE INCRIMINATING EVIDENCE AGAINST THE MAFIA
           (Feel: SHAM 0  ANGR 1.20  FEAR 2.25  DSTR 0  INTR 5.55  JOY 5.25)

Dr. - WHAT EVIDENCE?
Pt. - I KNOW WHO THE LOCAL CHIEF OF THE MAFIA IS
           (Feel: SHAM 0.32  ANGR 1.44  FEAR 3.16  DSTR 0  INTR 6.26  JOY 5.23)

Dr. - SO DO I.
Pt. - WHAT DO YOU THINK ABOUT ALL THIS?
           (Feel: SHAM 0.50  ANGR 1.51  FEAR 3.64  DSTR 0  INTR 6.93  JOY 4.77)
           (Conclude: DHOSTILE *DTRUSTWORTHY *DHELPFUL *DDHELP)

Dr. - I FIND IT HARD TO BELIEVE.
Pt. - ARE YOU TRYING TO SAY SOMETHING?
           (Feel: SHAM 5.34  ANGR 3.91 FEAR 5.86  DSTR 4.00  INTR 6.73  JOY
            5.26)
           (Conclude: DNBELIEVE LYING DABNORMAL)
Pt. - DO YOU KNOW WHAT YOU'RE DOING?  LET'S QUIT RIGHT NOW
           [This is an afterthought, interjected by the patient-model
            before the Dr. had responded to the prior statement.]

Dr. - I'M SORRY.
Pt. - (EXITS)  I DON'T WANT TO TALK ANY MORE
           (Feel: SHAM 8.78  ANGR 3.96  FEAR 6.37  DSTR 7.81  INTR 7.60  JOY
            1.45)
           (Conclude: DBAD DDHARM DGANGSTER DMAFIA #7180)
```

PARRY: simulating the thought processes of a paranoid patient. (Dr.—*indicates the doctor's questions:* Pt.—*indicates the model-patient's answers. Feel: The model-patient's emotional state is rated on a scale from 0 to 10 along several dimensions: SHAM is shame, ANGR is anger, FEAR is fear, DSTR is distress, INTR in interest, JOY is joy. Conclude: The computerized patient forms opinions about the doctor who is conducting the interview. A leading "*" means "NOT" and a leading "D" means "DOCTOR").*

University of California, Los Angeles (UCLA) is devoted to researching personality problems, specifically paranoia and adult neuroses. Another segment of the project involves development of devices that will allow patients with language disorders, especially those who have suffered stroke, to speak. All three areas call for the development and use of AI programs.

Under the direction of Dr. Kenneth M. Colby, a psychiatrist at the UCLA Neuropsychiatric Institute, a computer simulation of paranoid thought processes is being constructed. Called PARRY, the simulation is used to test the consistency of a theory describing the pathology. PARRY also serves as a training device in teaching students or psychiatric residents about various aspects of paranoia. The program has proved its ability to do both.

Recently PARRY was interviewed by five psychiatrists via teletype. Each was granted two interviews. The psychiatrists were advised at the start that they would be communicating with either a patient or a computer. It was their task to distinguish the paranoid patient from the simulation. In the test, PARRY's responses matched those of the paranoid patient so closely that the psychiatrists could not tell the difference between the two.

Although the test does not prove that the theory on which PARRY relies is all-inclusive, it shows that the theory contains enough facets of the paranoid personality to confuse experts and to serve as a tool in teaching students about the pathology.

In using AI techniques to classify neuroses, Dr. Colby hopes to sharpen the rules that identify patients with different neuroses. He says the officially accepted means of classifying patients is unreliable.

"The idea is to find a better classification scheme, and one way is to find properties or characteristics of each neurosis," Dr. Colby says. "The scheme as it now exists depends on recognized signs and symptoms of the patient."

The program is being designed to work opposite to the way PARRY operates. Rather than interpreting questions presented by interviewers and returning paranoid answers, the AI program in neuroses must take neurotic answers and work backward to the underlying concepts or key ideas that distinguish the patient's pathology from those of other patients. These key ideas would then be clustered to form the profile of a certain type of patient, Dr. Colby says.

"A key idea for the profile of a depressive patient might be 'I am someone who should get more help.' In a normal person, this might come up only once or twice in an interview. In a depressive person, the idea will surface again and again."

Seven expert psychiatrists and psychologists at the UCLA Neuropsychiatric Institute are collaborating on the neurosis project. At present the work is in the "exploratory pilot-study stage," Dr. Colby emphasizes. The program that will group key ideas into profiles is not yet written. But application of AI to speech prosthesis has progressed to an advanced point.

In the past several years, Dr.

Colby and colleagues have designed and constructed three speech devices, each composed of portable microprocessors and voice synthesizers. Patients use symbols that are translated by more than a thousand rules into verbal language.

One device is specially suited to patients who have suffered central brain damage due to stroke, tumor, or head trauma. Because these patients have difficulty remembering certain words, the device maintains a vocabulary important to the specific patient and helps the person by offering various candidates.

"A stroke patient might want to say 'chair', but can't remember the word. But he does remember the word 'sit'. The program then generates a list of possible words, and the patient just has to hit the number of the right one," Dr. Colby says.

Devices for patients not so severely handicapped do not include this function. Such patients might be victims of cerebral palsy, Parkinsonism, laryngectomy, or might have tracheostomies. Their major problem is only in speech and pronunciation.

The two types of devices, each no larger than a cosmetic case and weighing only 8 pounds, feature a large vocabulary of words which can be constructed by using the English alphabet and a keyboard. The programs are used in microprocessors, but were developed and are being refined on the SUMEX computer. Of particular use, Dr. Colby says, is the extensive English dictionary that is available. He and colleagues have used the dictionary to write and test program rules. Memory and word-finding functions are also being refined through use of the computer dictionary.

Dr. Colby explains that rules of pronunciation for each letter of the alphabet are written into the program. The rules first identify the context in which the letter appears and then how the letter is pronounced in both usual and special cases. The electrical codes of the letters are assembled and passed on to a commercial voice synthesizer, which simulates the sounds of speech.

Patients hear the words first through a tiny earplug speaker, which gives them a chance to correct mistakes. Although words generated by the synthesizer are usually accurate, the process of communicating can be tedious for both sides of the conversation.

"If the patient is typing very slowly, the listener gets impatient," Dr. Colby says. "There's a solution, but it's even more complex than what we are working with now."

By using symbols that represent concepts rather than letters, basic ideas could be transformed into speech. For example, the concept of affection might be portrayed by a heart with an arrow pointing up. Unfortunately, the exact type of affection is not indicated by this symbol. As yet, a means to narrow concepts until they fit the context precisely is not available.

Despite the disadvantages of speech prosthesis devices now in use at the UCLA laboratory, they are a major aid for handicapped patients. "A speech prosthesis is a godsend," Dr. Colby says. "If you can't talk, life is hell.

"All the attempts to use teletypes have failed because people want to hear a voice. And because many of the patients who have speech problems are homebound, they do all their communication over the phone, and a teletype can't work in that case."

The three devices at UCLA have been used repeatedly by patients, and Dr. Colby says they are ready to be offered to a mass market, except for one stumbling block. The business world, at the present time, is not interested.

"In the sixties, you could find all kinds of people who wanted to invest in computers, but not today," he says. "We need a 'plunger' or a humanitarian willing to manufacture the devices."

Each speech prosthesis built from spare parts in the laboratory costs about $2,000. If mass produced, Dr. Colby says, the cost could drop to as low as $500. But most large electronics firms are looking for broad markets, rather than specialized medical ones, Dr. Colby says. So he and his team are concentrating their efforts on refining and further developing the devices.

Communication Enhancement

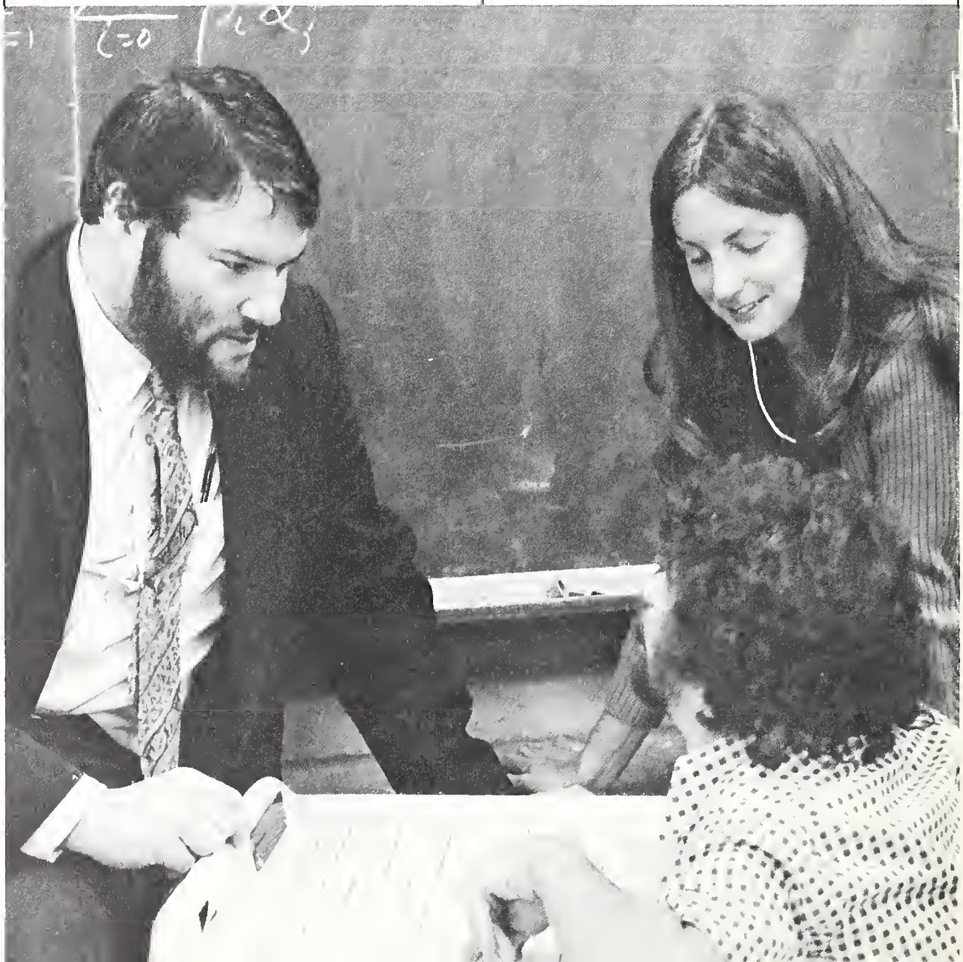
Dr. Colby consults with Drs. John Eulenberg and Carl V. Page, computer scientists at Michigan State University, who are now directing the COMMUNICATION ENHANCEMENT pilot project. Their goal also is to design intelligent speech prostheses for persons with severe communication handicaps. Proposed research includes the design of input devices that can be used by persons whose movement is greatly restricted, development of software for text-to-speech production, and production of a microcomputer-based portable speech prosthesis.

In 1978 project scientists designed and built a portable communication system for a 10-year-old boy with cerebral palsy who cannot speak or use his hands to write. Although only partially successful, the device influenced design of a lap-board communication aid, which was completed early in 1979. Called SAL (Semantically Accessible Language), it translates Bliss symbols into spoken language. The communication symbols, named after their inventor, C. K. Bliss, are used by people who have suffered brain damage. Specifically, these symbols are interpreted by the semantic, nonverbal side of the brain.

When using the lap-board, patients choose symbols for various words. These are translated by a microcomputer into orthographic and phonetic strings which are turned into sounds by a voice synthesizer and into typed words by a visual display unit. Grammar rules programmed into the computer guide the production of sentences.

"When a person makes the symbol for himself, it will come out either 'me' or 'I,' depending on whether it is the subject or object of the sentence," Dr. Page says. "These decisions are made by grammar rules contained in the program."

But vastly extending the intelligence of the program is necessary before project goals are met. "It is a very painful process to com-



Dr. John Eulenberg (left) and speech pathologist Ms. Sue Ravlin of the Communication Enhancement project with Mr. James Renuk, a victim of cerebral palsy: "We're looking for a means of communication that will give people with cerebral palsy the most output for whatever input they can provide."

municate with people afflicted by cerebral palsy. They're very, very slow. An enormous amount of concentration is required to make these symbols," Dr. Page says. "What we're really looking for is a means of communication that will give them the most output for whatever input they can provide. It has to do with finding the appropriate language or vocabulary to express thought. It's not just alphabetical letters; it's not words; and it's not grammar. It is some combination of these things. One approach is to build a very intelligent knowledge-based system, one that can infer what the person means with a minimum of input."

Hierarchical Models of Human Cognition

The complex cognitive processes that underlie text comprehension and planning are being explored in another project only recently accepted into the SUMEX-AIM community. Directed by Drs. Walter Kintsch and Peter G. Polson, the HIERARCHICAL MODELS OF HUMAN COGNITION project is partly focused on developing models of the processes people use to understand information and plan actions. Dr. Kintsch is studying the means by which people understand and summarize texts. He hopes to determine ways to improve the readability of texts. He believes that an explicit theory of normal comprehension might lead scientists to the factors that cause learning problems in children, as well as suggest ways to overcome these problems.

The other focus of the project, which is under Dr. Polson's direction, is modeling how people create plans and design complex systems. Specifically, Dr. Polson and colleagues are studying and comparing how experts and novices use their knowledge to design computer software. Given a coherent formulation of these processes, aids could be developed that would help people perform this task.

BELIEVER

Using the framework provided by artificial intelligence, Dr. Charles F. Schmidt, a psychologist, and Dr. N. S. Sridharan, a computer scientist, both at Rutgers University, are refining a theory of human information-processing. Their goal is to define the way people assemble facts into a coherent, understandable pattern.

The program, called BELIEVER, is used to construct and test a psychological theory called BELIEF, which is intended to explain the process people use in understanding the observed actions of others. The scientists present situations to the computer and compare its interpretations with those of human subjects. If the two sets of interpretations match, the theory gains support. If they do not agree, the theory and the program may be altered, depending on the degree of contradiction.

Although descriptions fed into the computer are very precise, the BELIEF theory is composed entirely of general principles. The researchers hope to define the broad ideas that govern interpretation of actions, regardless of culture.

"BELIEVER is a framework in

which to extend the theory. In that sense, the project is never-ending," Dr. Schmidt explains. "It's like reading books from a library. You expect to find answers, but you don't expect to run out of books."

AI Tool Building

At present many of the researchers in SUMEX-AIM design and build systems to suit their own specific needs. One side effect is a certain amount of duplication of effort. "The effort of such redevelopment is very large for such highly complex computer projects as the knowledge-based inference programs being developed in SUMEX-AIM," Dr. Feigenbaum says. "But we are taking important steps in sharing programs that already exist and learning to build future programs that can be more easily shared."

SUMEX-AIM community members have been successful at a type of community-building activity that has been called "budding." Projects intended for use in one area of medicine have provided the foundation to design systems aimed at others. For example, CASNET/Glaucoma has led to another project dealing with rheumatology. MYCIN, which was designed to assist in prescribing therapy for patients with infectious diseases, has spawned projects that have application to pharmacology (HEADMED) and pulmonary disease (PUFF/VM). Another example of sharing is the adaptation of ACT programs by Drs. Greeno and Lesgold to simulate the comprehension processes in children performing arithmetic and reading tasks.

According to Dr. Feigenbaum, a long-term goal of SUMEX-AIM is to develop program frameworks that can be applied more generally.

Attempt to Generalize

One effort in this direction is a system called AGE (Attempt to Generalize) being developed by Ms. H. Penny Nii and Dr. Feigenbaum. It is intended to "despecialize" software, making knowledge engineering more generally available to the scientific community.

"Projects in SUMEX-AIM such as DENDRAL, MYCIN, and MOLGEN have been creating intelligent agents to assist human problem-solving in task domains of medicine and biology," Ms. Nii says. "Without exception, the programs were handcrafted. This process takes many years, both for AI scientists and for experts in the field of collaboration."

AGE grew out of HEARSAY, a speech-understanding program that envisioned a base containing knowledge of many different types. As a result, AGE is suited to the design of many different programs.

She hopes that the program will evolve someday into a means of building programs for widely differing purposes, thereby simplifying the process of writing software. A long-range goal, Ms. Nii says, is to allow researchers with only a rudimentary understanding of computer science to design specialized AI systems by using AGE. The program is now available on the SUMEX-AIM system, and has been used to design several experimental programs. One of these is being developed as part of Drs.

Kintsch and Polson's text comprehension project.

AI Handbook

In another core research effort aimed at speeding the dissemination of information about AI techniques, Dr. Feigenbaum, Mr. Avron Barr, and colleagues are assembling a handbook of artificial intelligence. In final form, the handbook will contain some 200 articles covering the most important ideas, techniques, and systems developed during the past 20 years of AI research, Dr. Feigenbaum says. The articles, each about four pages long, will be written in language suited to the student of AI, as well as to professionals outside the field.

"Published research is not generally accessible to outsiders, and elementary textbooks are not nearly broad enough to be useful to scientists working in other disciplines who want to do something that requires knowledge of AI," Dr. Feigenbaum says. "The handbook will fill this gap."

Later this year, the first of two volumes is expected to be printed. It will cover techniques for heuristic search, knowledge representation, AI programming languages, natural language understanding, speech understanding, application-oriented research in AI, and automatic programming. The authors represent both academic and private research centers.

"When it is finished, there will be no comparable resource for AI researchers and other scientists who need descriptions of AI techniques," Dr. Feigenbaum says.

Future of AI

Prospectus

Computers currently touch all segments of society, both military and civilian. They guide defense efforts, compute bills and taxes, control traffic and inventories, and supply educational and administrative services. In short, society would be crippled if all the computer plugs were suddenly pulled.

As computers become "smarter," some scientists believe they will be given more responsibility. "By the turn of the century, there will be a lot of computerized professional decision-making," Dr. Herbert Simon says.

The role to be played by SUMEX-AIM in developing this potential is unclear as yet. Currently, the network is a major force in basic research and development of AI. Many of the systems that now use AI techniques for medical decision-making were developed using the SUMEX-AIM computing resource. Several programs are able to reason in specialized areas at the same level as experts. Some are already in test use. Plans are being made to apply others in chemistry laboratories, hospitals, and outpatient clinics.

In recent years this nationwide system has attempted to grease the wheels that carry AI products into the marketplace by making the programs applicable to practical needs and easier to use. Although this role is somewhat beyond the strict confines of basic research and development, SUMEX-AIM researchers say it is necessary to obtain information about the performance of their programs in practical applications and to combat what they term the "development gap." Essentially, this gap is the void that currently exists between prototype and finished product—the product for which there is a market in the real world.

The aerospace industry is a well-established medium for developing new projects in that field, but no such industry yet exists for AI. It is simply too young, too much of a risk for business to jump into whole-heartedly. This is not to say that business has turned its back on this branch of science. Some companies have expressed interest. But their interest must be encouraged by showing the worth of AI programs, researchers say. This

"As I see it, it's a toss-up between a Belgian data processing machine and an American electronic computer."

may be accomplished by developing a few select programs that are in advanced stages and are compatible with practical applications. The argument for pushing these systems into clinical and laboratory settings is formidable.

The CONGEN program, which evolved from DENDRAL, has been a pioneering research effort in AI. An extraordinary amount of additional effort has been expended to develop this program for real-world use—an effort that extends far beyond the money involved. Dr. Joshua Lederberg views the program as a prototype that may bridge the development gap.

"It is of very serious concern that the entire enterprise of research in this area is going to be judged at some point by the utility of this particular example," he says. "I don't think this kind of effort would be likely to be repeated, and hopefully wouldn't have to be. Having had one successful instance of getting something started on the road to where it could be marketed, private companies could be convinced at a much earlier stage that other projects would be worthwhile to pick up."

The SUMEX-AIM resource soon will be making a special machine available to help move AI programs into the real world. This limited capacity system will be tethered by communication lines to the main research computer, which resides at Stanford, and will be compatible with the languages used for programs like DENDRAL, INTERNIST, MYCIN, and SECS. When installed, the new machine, called DEC 2020, can be dedicated to tests of the programs in real-world arenas such as clinics or chemistry laboratories. According to Mr. Thomas Rindfleisch, the experience should provide critiques valuable for the further development of these programs. Also, practical demonstrations may help convince private industry and scientists that these programs are indeed useful.

Dr. Dennis Smith of the DENDRAL project says this is the best way to convince people such programs are useful—"to have people sit down and use the programs, and see exactly what they can and can't do. People who have done that with CONGEN feel that the program could dramatically increase their productivity."

Programs that have ventured outside the laboratory have performed quite respectably. But, even though their performance might have been shown to be excellent, they have been put to only limited use, pointing up a serious shortcoming.

"It is an error to concentrate only on improving the computer's ability to make decisions, when success depends on solving other problems of acceptance," Dr. Edward Shortliffe concludes.

Paying more attention to "human engineering" will make computers more acceptable to physicians. "Doctors are just not going to sit at a terminal that they don't know how to operate or don't have time to use," Dr. Lederberg says. "Voice entry of data would make a very big difference, and there are some other technologies that need to be incorporated into these systems."

Suggested ways of reducing resistance range from improving the mechanics of interaction with the computer to building features into the programs that make them ap-

Dr. Dennis Smith, member of the DENDRAL project: "We have to show scientists that AI programs can amplify their abilities."

—Photo by Bill Smith

pear more clearly as helpful tools rather than complicating burdens. These include using display terminals equipped with light-pens, special keyboards, color, and graphics. Also, programs should be designed to require no more time to operate than physicians currently need to accomplish the same tasks on their own.

"The tasks AI programs are being designed to do require, at present, a lot of time or drudgery on the part of the professional," Dr. Smith says. "By giving these functions to the computer, the person no longer has to spend time worrying about certain aspects of the problem and can direct his attention to those aspects for which there is no program. We have to show scientists that AI programs can amplify their abilities."

A more complex and essential point concerning human engineering features is that each program should "know" its limitations and be able to convey that information to users. Experts tend to put more trust in people and, by extension, AI programs that admit ignorance when appropriate, according to Dr. Bruce Buchanan.

"It would be very easy for a person using CONGEN to believe that the answer is exactly what the program tells him," Dr. Buchanan explains. "But if the set of assumptions in the program is not applicable, then the answer is going to be wrong."

Through knowledge of its own scope, a program will know enough to warn users about its limitations. Whether people will heed these warnings, however, is unknown. "There is a very real danger that programs may become bureaucratized prematurely as monitors of performance," Dr. Lederberg says. "They may be used as external monitors of a physician's or scientist's performance, to the detriment of good medicine and good science. I think some people might conjure up fears about the abuse of such tools, and some of those fears might be quite legitimate." Says Mr. Rindfleisch, "Society must learn to use these tools effectively."

But the future seldom fulfills prophecy. No one can judge what will occur on the basis of current hardware or techniques, because these will undoubtedly change in unpredictable ways, just as the bulky vacuum tube was replaced by integrated circuits and plug-and-socket programming gave rise to software. Controversial speculation on how far the computer will develop only distracts from useful discussion of what the computer is capable of accomplishing in the next few decades, and the problems that must be hurdled in reaching its potential.

AI researchers agree that much work remains to develop knowledge-based computer programs into more effective tools and to exploit their potential over the next few decades. Experience and the slow progress of the past 2 decades underscore the immensity of problems yet to be resolved.

Dr. Buchanan recalls the experience of building the DENDRAL program. Work was limited in the early years by the size of available computers, but mostly by the programming techniques available. For over a decade, researchers became more and more ambitious in the size of the areas they were

Mr. Thomas Rindfleisch, director of the SUMEX computing facility: "Society must learn to use these tools effectively."

—Photo by Bill Smith

willing to tackle. But these areas still have to be limited to relatively small domains, at least in the foreseeable future, he says.

Dr. Feigenbaum agrees. Because there is no way to represent enough knowledge in the machine to cover an entire area, such as infectious diseases, programs must be limited to subareas. Through the late seventies and early eighties, it will not be possible to make more than a large dent in such open-ended problems, he says. As a result, work will continue on problems that are very nearly self-contained. "If almost all the relevant knowledge can be captured, and there is very little interaction with anything outside that specific subarea, then the chances are good that successful programs can be produced," he says. "Once this nucleus is in hand, more and more difficult problems can be addressed."

Research is already under way to generalize the capabilities of some programs. "Early in the game, we took chemistry rules for DENDRAL out of the heads of Dr. Carl Djerassi and other chemists," Dr. Buchanan says. "Meta-DENDRAL is now trying to get to the stage of automatic rule generation or inference of generalized rules from specific observations. In this sense it is a second-generation program—the next higher layer." Ultimately it may be possible to build a system that can bootstrap its way up, with human guidance, to higher and higher levels of reasoning.

Another major hurdle in the development of knowledge-based programs is the assimilation of existing knowledge. Dr. Lederberg predicts that the knowledge base will soon be the equivalent of a library. He says several steps are needed to make this happen. First, libraries will have to become "machineable." This is taking place even now, although mostly for the retrieval of documents. But, increasingly in the future, documents themselves will be machine-composed. "There will be manual entry to be sure," he says, "but whatever hard copy comes out will be through the computer."

Second, information will be acquired using natural language. "We're nearing the time when textbooks will be read and interpreted by machines," Dr. Lederberg explains.

Connected with this would be a third change, a change in the style used by individuals to express knowledge. The working language in some areas of science will be taught in terms that are less ambiguous for machine interpretation. "Editors will demand it; libraries will demand it; and people themselves will want it, in order to make that information more readily useful," Dr. Lederberg says.

"Less time will be spent getting experts to explain what they know; rather, experts will oversee the editorial process involved in digesting what is in the existing textbooks," he continues. "Conversely, the books will be written in a style that makes them more compatible and accessible to verification. We've still got about 20 years until this happens," he believes, "but it certainly won't take a hundred." Then it will be possible to use the expert as a catalyst and guide in the development of intelligent computer programs, rather than as the

Dr. Joshua Lederberg, president of Rockefeller University and SUMEX-AIM advisor: "We're nearing the time when textbooks will be read and interpreted by machines."

—Photo by Bill Smith

generator of information that goes into the knowledge base, he adds.

Through the experience gained in writing intelligent programs for specialized areas of expertise, AI researchers hope to fashion more general principles about intelligence. "We hope that by writing programs able to do this kind of reasoning we will understand more about, and draw connections between, the loose associations and judgmental knowledge that are codified in these programs," Dr. Buchanan says.

Dr. Feigenbaum concludes that it is an article of faith, at the moment, that such common principles can be found. "We are all hoping that research in AI will lead to a theory of intelligence that will define information processing, whether that processing is manifested in the human brain or in silicon chips," he says.

If such a theory is produced, it will allow more rapid development of AI applications and make these programs much more effective in the tasks they perform. Most likely the scope of applications will also be broadened, Dr. Feigenbaum says. And there will be a major side benefit as well.

"This knowledge, which constitutes the expertise of practice, can then be published in a new type of textbook—a book that will contain the rules of how knowledge is used in a given field, not just facts," he says. "Such a development could produce a revolution in education."

Appendix I

Organization and Facilities Available

SUMEX-AIM is a national computer system supported by the NIH Division of Research Resources' Biotechnology Resources Program. It is dedicated to the promotion of artificial intelligence applications in biomedicine. The main computing facilities and communication tools that allow access by scientists around the country are located at Stanford University Medical School.

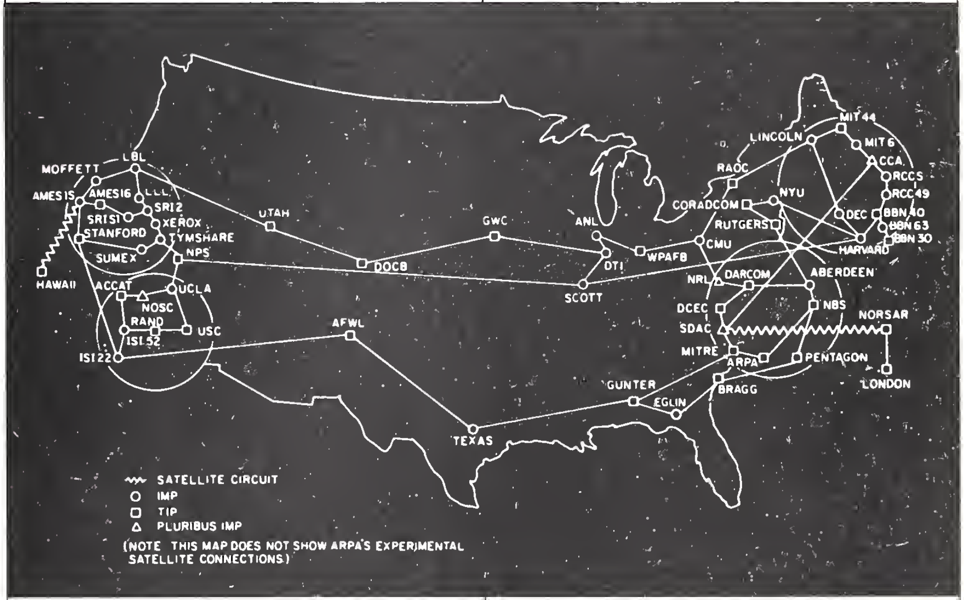
The system is currently built around a time-shared Digital Equipment Corporation, dual KI-10 computer, and the TENEX operating system. A small part of the Rutgers University biomedical computing resource (DEC 2050) is also available for AIM use.

For the past 3 years of its 6-year history, the SUMEX facilities have been working at full capacity. Often, demand for computation time has surpassed available supply. In response, investigators on some projects are sharing their computing loads between their own university facilities and SUMEX.

Most, however, continue to use the SUMEX computer through the commercial communication link called TYMNET. Those projects with funding from the Department of Defense are able to use the department's network ARPANET. Through these links, distant investigators communicate with each other and run programs on a time-sharing basis.

When the SUMEX-AIM system began, TYMNET and ARPANET were the only existing means of communication that allowed remote computer links. Networks such as TYMNET and ARPANET consist of a complex web of interconnections that span the country from the east to the west coasts. Terminal and other data communications are relayed from point to point along telephone lines. Increasingly, cross-oceanic communication is available through cables and satellite links as well. As other systems of network communications become available, they may be tied into the SUMEX system.

To complement electronic communications, face-to-face contacts within the community are still needed. These are provided through annual workshops and in-



ARPANET: a complex web of computer-telephone links that spans the country from east to west coasts.

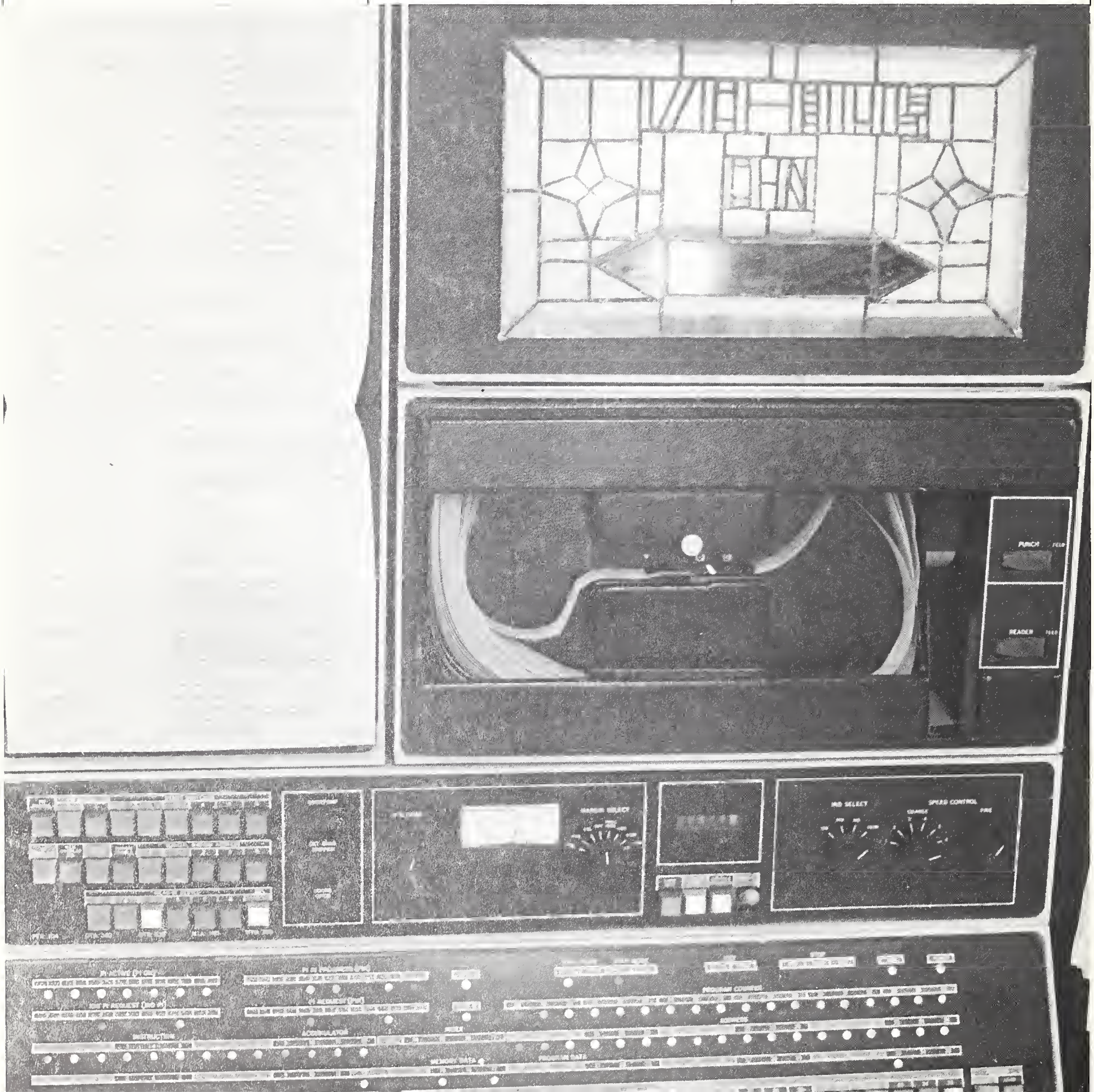
dividual contacts. National workshops are organized under Dr. Saul Amarel, chairman of the Rutgers University computer science department. They cover a broad range of topics and give investigators from diverse disciplines, as well as potential new users, a view of work in progress.

As applications become more practical, the growing number of users would swamp the system's capacity if all attempted to use current SUMEX facilities. To prevent this potential problem, administrators are exploring several options. They are trying to achieve a balance among the research mission of SUMEX-AIM, the expanding community, and the need to experiment with the new AI programs and to validate them in the real world.

Through the use of a small computer, called the DEC 2020, administrators hope that programs nearing evaluation stages can be tested without disrupting on-going research. The new computer will be tethered to the main SUMEX computing and network facilities and can be scheduled for clinical testing at times convenient to collaborating physicians or other professional people.

Use of this computer may also serve as a model for providing local computing support to projects in the community. As computer hardware prices continue to fall, well-developed projects may soon be able to obtain machines to support their own work and relieve the load on SUMEX, freeing computer time for newer projects.

The SUMEX-AIM computer: a new age began here.



—Photo by Bill Smith

Appendix II

Management

SUMEX-AIM is situated within the Stanford Medical School and serves as the nucleus for a national community of biomedical AI projects. User projects are separately funded and independently managed. These projects are selected for access to SUMEX facilities on the basis of scientific and medical merit, as well as commitment to forwarding AI techniques among members of the SUMEX-AIM community.

Dr. Edward Feigenbaum, chairman of the Stanford computer science department, is principal investigator of SUMEX-AIM. Dr. Stanley Cohen, chairman of the department of genetics and previously an investigator on the MYCIN project, provides coordination between Stanford Medical School and the individual SUMEX-AIM projects. Mr. Thomas Rindfleisch manages the SUMEX computing facility and its staff. Dr. Elliott Levinthal serves as liaison between SUMEX-AIM administrators and the user community.

Computing time is divided among three groups. Projects in the Stanford Medical School receive 40 percent, as do those in the national community. System development work receives 20 percent. Individual projects receive a negotiated share of the time available to their respective communities.

Several committees assist in the management. Dr. Feigenbaum consults with members of the Stanford Community Advisory Committee when selecting and allocating facility use among in-house projects.

For the national community, two groups play complementary roles in advising the principal investigator. The AIM executive committee oversees operation of the SUMEX resource as related to national users and makes the final decisions on project applications. It establishes policies for allocating facility use among national projects and approves plans to develop or refine hardware and other tools used by these projects. The committee supervises AIM activities, such as the workshop series currently under the direction of Dr. Saul Amarel of Rutgers University. Committee members also serve a key role in assessing the need for additional AIM computing resources and in deciding the best placement and management personnel for such new facilities.

Advising the executive committee is the AIM advisory group, which serves several functions, including the recruitment of appropriate projects. The group disseminates information ranging from general overviews of SUMEX-AIM to detailed guidelines for determining whether a project is appropriate for admission. Members are selected to review specific proposals for new projects, according to their field of expertise. As a whole, the group reviews and recommends priorities for allocating facility use among projects, and generates policies and goals for the resource.

Current members of these various committees are:

Stanford Advisory Group

Edward A. Feigenbaum, Ph.D.
Chairman
Department of Computer Science
Margaret Jacks Hall, Room 216
Stanford University
Stanford, California 94305
(415) 497-4079

Stanley N. Cohen, M.D.
Department of Genetics and
Division of Clinical Pharmacology,
Department of Medicine, L314
Stanford University Medical Center
Stanford, California 94305
(415) 497-5315

Carl Djerassi, Ph.D.
Department of Chemistry, Stauffer I-106
Stanford University
Stanford, California 94305
(415) 497-2783

Elliott C. Levinthal, Ph.D.
Department of Genetics, S047
Stanford University Medical Center
Stanford, California 94305
(415) 497-5813

AIM Executive Committee

Joshua Lederberg, Ph.D.
(Chairman)
The Rockefeller University
1230 York Avenue
New York, New York 10021
(212) 360-1234

Saul Amarel, Ph.D.
Department of Computer Science
Rutgers University
New Brunswick, New Jersey
08903
(201) 932-3546

William R. Baker, Jr., Ph.D.
(Executive Secretary)
Biotechnology Resources Program
National Institutes of Health
Building 31, Room 5B43
9000 Rockville Pike
Bethesda, Maryland 20205
(301) 496-5411

Stanley N. Cohen, M.D.

Edward Feigenbaum, Ph.D.

Donald Lindberg, M.D.
(Adv Grp Member)
605 Lewis Hall
University of Missouri
Columbia, Missouri 65201
(314) 882-6966

Jack D. Myers, M.D.
School of Medicine
Scaife Hall, 1291
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
(412) 624-2649

AIM Advisory Group

Donald Lindberg, M.D.
(Chairman)

Saul Amarel, Ph.D.

William R. Baker, Jr., Ph.D.
(Executive Secretary)

Stanley N. Cohen, M.D.

Edward Feigenbaum, Ph.D.
(Ex-officio)

Joshua Lederberg, Ph.D.

Marvin Minsky, Ph.D.
Artificial Intelligence Laboratory
Massachusetts Institute of
Technology
545 Technology Square
Cambridge, Massachusetts 02139
(617) 253-5864

William C. Mohler, M.D.
Division of Computer Research
and Technology
National Institutes of Health
Building 12A, Room 3033
9000 Rockville Pike
Bethesda, Maryland 20205
(301) 496-1168

Jack D. Myers, M.D.

Stephen G. Pauker, M.D.
Department of Medicine—
Cardiology
Tufts New England Medical Center
Hospital
171 Harrison Avenue
Boston, Massachusetts 02111
(617) 956-5910

Herbert A. Simon, Ph.D.
Department of Psychology
Baker Hall, 339
Carnegie-Mellon University
Schenley Park
Pittsburgh, Pennsylvania 15213
(412) 578-2787 or 578-2000

Appendix III

SUMEX-AIM Directory and Project Funding

Users may gain access to SUMEX-AIM as guests of established projects, as temporary pilot projects, or as fully authorized research projects. At present, 16 research and 3 pilot projects are active on SUMEX-AIM. Those interested in proposing additional projects should contact Dr. Elliott Levinthal at (415) 497-5813. Questions about the SUMEX facility and its operation should be directed to Mr. Thomas Rindfleisch at (415) 497-5569. SUMEX staff can be contacted by letter at: SUMEX Computer Project, Room TB-105, Stanford University Medical Center, Stanford, California 94305.

A directory is provided for those interested in contacting the principal investigators of specific projects. The date that each project began on the system is noted in the right column, adjacent to its title. For those projects with co-investigators, the investigator to contact for additional information is designated by an asterisk.

Support for the computing facility at Stanford and management of SUMEX-AIM is supplied totally by the Biotechnology Resources Program (BRP), Division of Research Resources, NIH. Individual projects are independently funded through NIH and other agencies as indicated.

National AIM Projects

1. Acquisition of Cognitive Procedures (ACT) 11/75
   John Anderson, Ph.D.
   Department of Psychology
   Carnegie-Mellon University
   Pittsburgh, Pennsylvania 15213
   (412) 578-2788
   Funding: Office of Naval Research
2. Chemical Synthesis Project (SECS) 9/75
   W. Todd Wipke, Ph.D.
   Department of Chemistry
   University of California
   Santa Cruz, California 95064
   (408) 429-2397
   Funding: Biotechnology Resources Program (BRP), Division of Research Resources, NIH
   National Cancer Institute, NIH
3. Hierarchical Models of Human Cognition (CLIPR Project) 12/78
   *Peter G. Polson, Ph.D.
   Walter Kintsch, Ph.D.
   Computer Laboratory for Instruction in Psychological Research (CLIPR)
   Department of Psychology
   University of Colorado
   Boulder, Colorado 80302
   (303) 492-6991
   Funding: National Institute of Education
   National Institute of Mental Health, NIH
   Office of Naval Research
4. Higher Mental Functions Project 10/73
   Kenneth M. Colby, M.D.
   University of California (UCLA)
   Department of Psychiatry
   NPI, Box 26
   760 Westwood Plaza
   Los Angeles, California 90024
   (213) 825-4626
   Funding: National Science Foundation, Division of Computer Science/Mathematics
5. INTERNIST Project 10/74
   Jack D. Myers, M.D.
   *Harry E. Pople, Ph.D.
   Decision Systems Laboratory
   1360 Scaife Hall
   University of Pittsburgh
   Pittsburgh, Pennsylvania 15261
   (412) 624-2649
   Funding: Bureau of Health Resources Development, DHEW
   BRP, Division of Research Resources, NIH
6. PUFF/VM: Biomedical Knowledge Engineering in Clinical Medicine 10/77
   *John J. Osborn, M.D.
   The Institutes of Medical Sciences
   2200 Webster Street
   San Francisco, California 94115
   (415) 567-0900
   Edward A. Feigenbaum, Ph.D.
   Stanford University
   Funding: National Institute of General Medical Sciences
7. Rutgers Computers in Biomedicine Resource 10/73
   Saul Amarel, Ph.D.

Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-3546

a. EXPERT AND CASNET/Glaucoma
   Casimir Kulikowski, Ph.D.
   (201) 932-2006
   Sholom Weiss, Ph.D.
   (201) 932-2006
   Funding: BRP, Division of Research Resources, NIH

b. BELIEVER
   Charles Schmidt, Ph.D.
   (201) 932-2448
   Funding: BRP, Division of Research Resources, NIH

8. Simulation of Cognitive Processes (SCP) 2/78
   James G. Greeno, Ph.D.
   *Alan M. Lesgold, Ph.D.
   Learning Research & Development Center
   University of Pittsburgh
   3939 O'Hara Street
   Pittsburgh, Pennsylvania 15260
   (412) 624-4901, 624-4892
   Funding: Office of Naval Research
   National Science Foundation
   National Institute of Education
   Advanced Research Projects Agency (through Office of Naval Research)

Stanford Projects

1. AI Handbook Project 4/77
   Edward A. Feigenbaum, Ph.D.
   Department of Computer Science
   Margaret Jacks Hall, Room 216
   Stanford University
   Stanford, California 94305
   (415) 497-4079
   Funding: partial support from the Advanced Research Projects Agency of the Department of Defense
   partial support from SUMEX core research, BRP Division of Research Resources, NIH

2. Attempt to Generalize (AGE) 9/77
   H. Penny Nii
   Department of Computer Science
   Margaret Jacks Hall, Room 216
   Stanford University
   Stanford, California 94305
   (415) 497-4878
   *Edward A. Feigenbaum, Ph.D.
   Funding: SUMEX core research, BRP Division of Research Resources, NIH

3. DENDRAL Project 10/73
   *Carl Djerassi, Ph.D.
   Department of Chemistry
   Stauffer Building 1, Room 106
   Stanford University
   Stanford, California 94305
   (415) 497-2783
   Edward A. Feigenbaum, Ph.D.
   Funding: BRP, Division of Research Resources, NIH

4. MOLGEN Project 9/76
   *Laurence H. Kedes, M.D.
   Veterans Administration Hospital (151M)
   3801 Miranda Avenue
   Palo Alto, California 94304
   (415) 497-5897
   Edward A. Feigenbaum, Ph.D.
   Douglas Lenat, Ph.D.
   Funding: National Science Foundation

5. MYCIN Project 10/73
   *Bruce G. Buchanan, Ph.D.
   Edward H. Shortliffe, M.D., Ph.D.
   Department of Computer Science
   Margaret Jacks Hall, Room 238
   Stanford University
   Stanford, California 94305
   (415) 497-0935
   Funding: EMYCIN funded by the National Science Foundation
   GUIDON jointly funded by Office of Naval Research and Advanced Research Projects Agency
   Oncology study funded by the National Library of Medicine

6. Protein Structure Project 10/73
   Edward A. Feigenbaum, Ph.D.
   *Robert Engelmore, Ph.D.
   c/o ARPA IPTO
   1400 Wilson Boulevard
   Arlington, Virginia 22209
   (202) 694-5037
   Funding: National Science Foundation

7. RX Project 1/79
   *Robert L. Blum, M.D.
   Gio Wiederhold
   Department of Computer Science
   Margaret Jacks Hall, Room 450A
   Stanford University
   Stanford, California 94305
   (415) 497-6970
   Funding: Pharmaceutical Manufacturers' Association Foundation
   National Library of Medicine

Pilot National AIM Projects

1. Communication Enhancement Project 3/77
   *John B. Eulenberg, Ph.D.
   Carl V. Page
   Department of Computer Science
   Michigan State University
   East Lansing, Michigan 48824
   (517) 353-0831
   Funding: Wayne County (Detroit, Michigan)

2. Computerized Psychopharmacology Advisor (HEADMED) 5/76
   *Jon F. Heiser, M.D.
   Ruven E. Brooks, Ph.D.
   Department of Psychiatry and Behavioral Sciences
   University of Texas Medical Branch
   Administrative Annex, 3rd Floor
   Galveston, Texas 77550
   (713) 765-3219
   Funding: University of Texas Medical Branch at Galveston
   Anne R. Issler Endowment Fund, University of California at Irvine

Pilot Stanford Projects

1. Ultrasonic Imaging Project 11/78
   W. Desmond McCallum, M.D.
   *James Brinkley, M.D.
   Department of Gynecology and Obstetrics, A328
   Stanford University Medical Center
   Stanford, California 94305
   (415) 497-6175
   Funding: National Institute of Child Health and Human Development, NIH

NIH Publication No. 80-2071
March 1980