Research Resources Reporter

September 1979 Volume III, No. 9

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

Serving Animal Resources, Biotechnology Resources, General Clinical Research Centers, and Minority Biomedical Support

This Issue:

By
Gregory Freiherr

The Problems and Promises of Artificial Intelligence

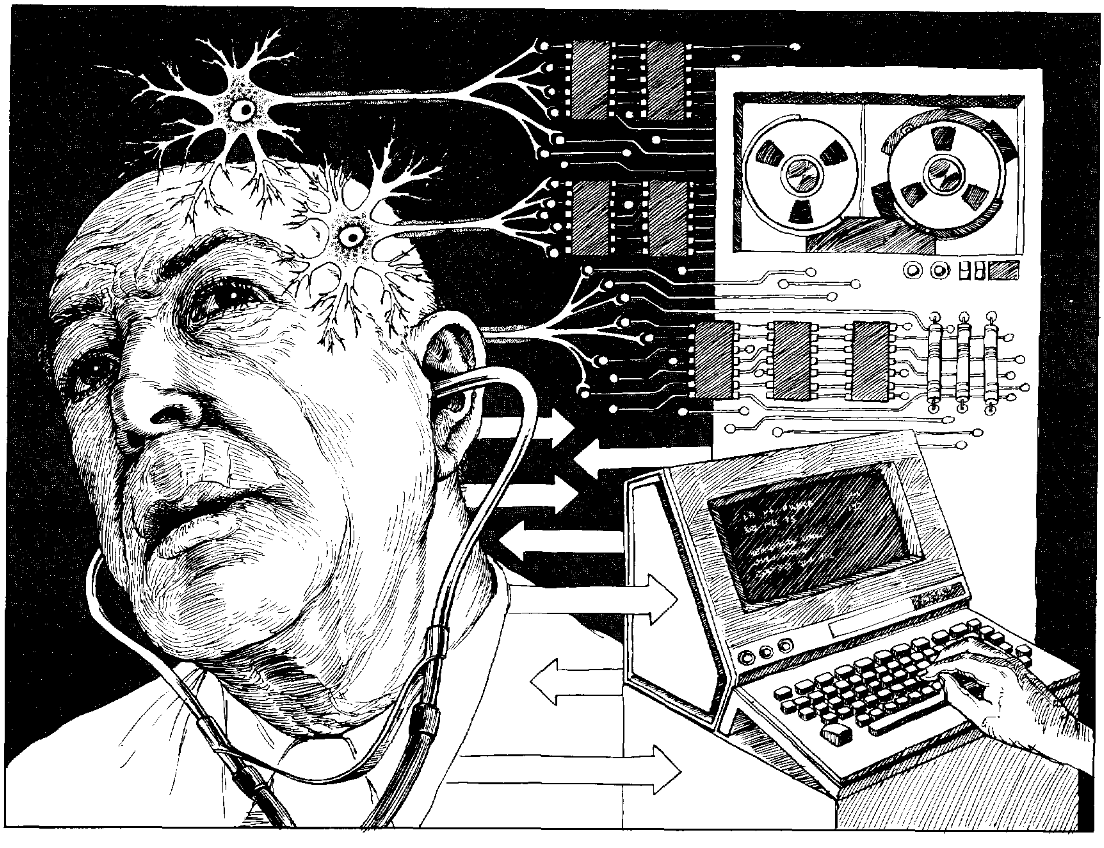
In the past century, science has not only changed our conceptions about the world, it has changed itself. Driven by an explosion of information, specialties in science have sprung up, inevitably giving rise to subspecialties. But staying abreast of developments, even in narrowly specialized areas, is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

The proposed solution is the "intelligent machine," a computer that mimics the expert's reasoning power and can retain in retrievable memory much of the knowledge currently available to experts in a given specialty.

The branch of computer science that embodies such research is called artificial intelligence (AI). It is a multidisciplinary field, comprising teams of computer programmers and experts from intended areas of application.

In the long run, AI promises to amplify the skills of less specialized physicians or scientists and to fill the current gap in professional manpower. These programs may also lead to concise descriptions of the rules and processes used by experts in making decisions. Such descriptions would be of major use in education.

The network of the Stanford University Medical Experimental Computer for Artificial Intelligence in Medicine (SUMEX-AIM), which is supported by the Biotechnology Resources Program of the NIH Division of Research Resources, is a major force in research and development of AI. Many of the computer programs that now use AI techniques for biomedical



INTELLIGENCE *continued*

Dr. Edward Feigenbaum, professor and chairman of the Stanford Computer Science Department, pioneered concepts and techniques for modeling scientific and medical reasoning processes. He and Dr. Joshua Lederberg designed DENDRAL, one of the first AI programs to gain worldwide recognition in chemistry as a practical laboratory aid.

Dr. Joshua Lederberg, president of Rockefeller University, former chairman of the Stanford Genetics Department, 1958 recipient of the Nobel Prize for physiology or medicine, worked with Dr. Edward Feigenbaum to construct DENDRAL and several other such programs, which became part of the Stanford Heuristic Programming Project. Under their direction, the project grew into the SUMEX-AIM network.

decision-making evolved within this computing resource. Most AI systems are still laboratory curiosities. But some are already moving into the real world and some others will make the transition within the next few years.

INTERNIST, the brainchild of Dr. Jack Myers and computer scientist Dr. Harry Pople at the University of Pittsburgh, is targeted at diagnosis and recommendations of therapy in the field of internal medicine. The program is scheduled to begin clinical tests at Pittsburgh teaching hospitals within a year. CONGEN, the offspring of the first AI program developed in SUMEX-AIM, is now being used in laboratories around the globe to analyze the structures of unknown compounds.

The widely used Stimulation and Evaluation of Chemical Synthesis program is focused on helping the chemist design routines for synthesizing compounds, including drugs. The program has led to the development of XENO, an early stage spin-off intended to map the metabolic pathways of various foreign compounds in animal and human systems.

PUFF, too, is in real world usage. At the Pacific Medical Center in San Francisco, clinicians rely on the program to monitor and interpret data concerning the pulmonary function of selected patients. Reports prepared by the computers are so accurate that the attending clinicians seldom need to change them before they are entered into the patient's permanent record.

There are more than a dozen other such programs in SUMEX-AIM. Few have met this kind of success, but investigators of SUMEX-AIM believe that these others also contain the basic ingredients necessary for them to succeed.

According to Dr. Edward Feigenbaum, principal investigator of the SUMEX-AIM network, the key to designing a successful AI project is to pick a problem limited enough to be contained, but not so simple that the program designed to solve it cannot be expanded into a practical tool. Most projects in SUMEX-AIM are re-

Research Resources **Reporter**
Published monthly by the Research Resources Information Center

A publication of the Division of Research Resources, National Institutes of Health

Correspondence should be addressed to:
Research Resources Information Center
1776 East Jefferson Street
Rockville, Maryland 20852
Telephone: (301) 881-4150

Subscription to the *Reporter* is free upon request to the above address.

(Prepared under contract with Tracor Jitco, Inc. NO 1-RR-6-2155)

Director:
Edward Post

Science Correspondents:
Gregory T. Freiherr
Derek V. Goodwin
Roger S. Johnson, Ph.D.

Science Adviser:
Freeman H. Quimby, Ph.D.

NIH Project Officer:
Jim Augustine, Information Officer, DRR

stricted 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 project personnel and at least one expert from the target field of application. The collaboration must be a dedicated one, Dr. Feigenbaum says. "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."

Unfortunately, the novel objectives and approaches that characterize the field of artificial intelligence have led to misunderstandings among the public, the general scientific community, and even, in some cases, those in the field itself. One cause for the confusion is the name *artificial intelligence*.

For centuries, philosophers and linguists have grappled with the question of how to define intelligence. Most have approached the issue by describing the function of intelligence, or the way it appears in behavior. But a concise definition for this term is elusive.

As might be expected, artificial 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 tools of research, machines programmed to do things that would require intelligence if done by people. Dr. Marvin L. Minsky, 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.

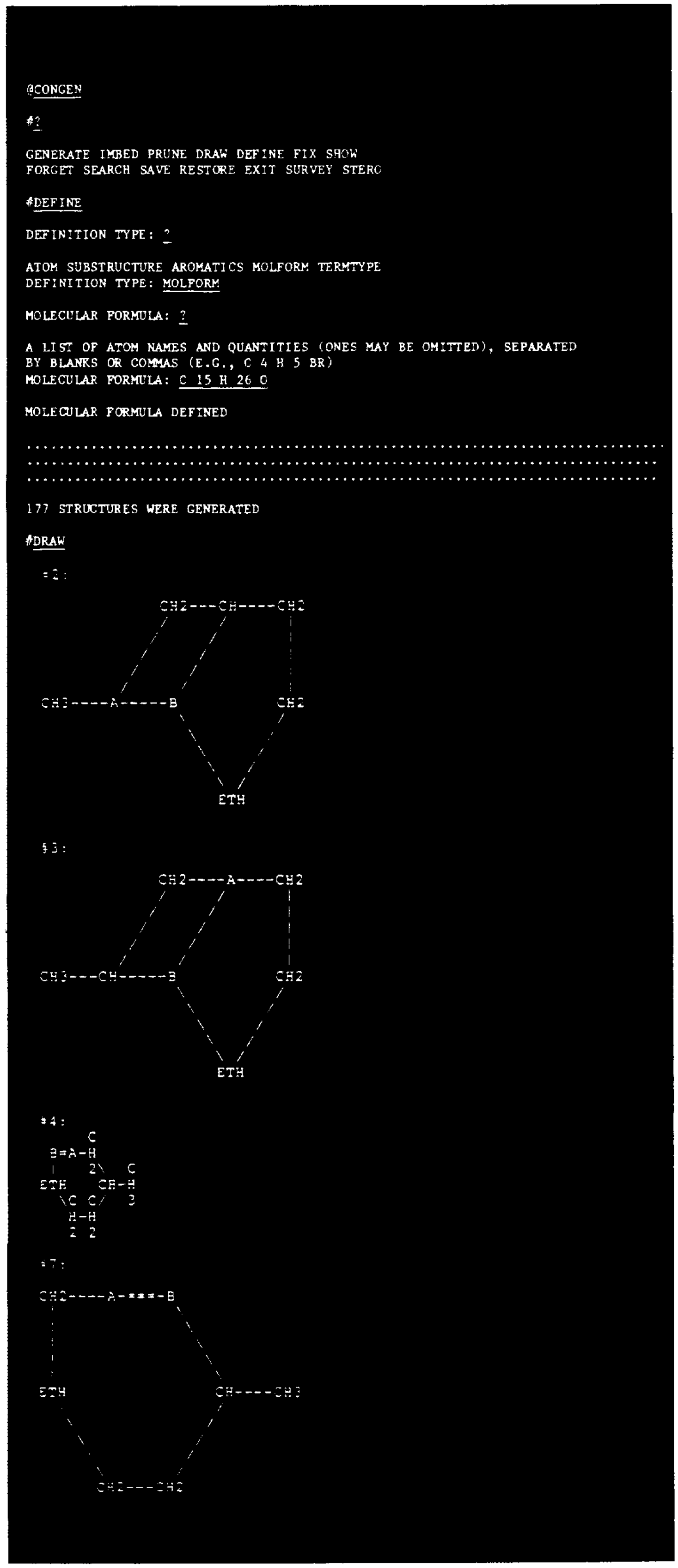
Others take a somewhat different view. Dr. Feigenbaum says the field is not primarily oriented toward technology, but toward investigating the nature of intelligence as information processing, whether the intelligence is expressed by man or machine.

One point of emphasis in current AI research focuses on designing computer programs that capture the knowledge and reasoning processes of highly intelligent specialists. The practical goal of this work is to make such 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 through this work. 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 the resulting programs is limited and much less flexible than the corresponding human intellect. In specialized areas of medical diagnosis and chemical structure analysis, however, some programs can already rival human capabilities. Still, many people are skeptical of the computer's potential.

Dr. Herbert A. Simon, psychologist-computer scientist at Carnegie-Mellon University and a 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 be-

CONGEN, one of the most successful applications of artificial intelligence yet developed, helps chemists determine the molecular structure of unknown organic compounds. By analyzing various data, the program generates candidate structures.



INTELLIGENCE *continued*

come 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 manipulated in the computer to solve 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. Heuristics are decision criteria based on experience and judgment. When used in AI, they 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 as in numerical calculation. At a higher level, these tools of symbolic processing are used to construct understandable lines of reasoning to solve problems and to interact with human users.

Use of complex symbolic structures is necessary when designing computer applications for domains not well-formulated by mathematical terms. 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. A key goal of an AI program is to explain conclusions, and allow the professional to interact in the decision process. Generally, the logic built into these programs comprises 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.

☐ *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 system.

☐ *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.

Mr. Thomas Rindfleisch, senior research associate in computer science and genetics at Stanford, has been the director of the SUMEX computing facility since its beginning in 1973.

These elements are necessary to the construction of AI programs, but they are not enough. Such programs gain their power from qualitative, experience-guided judgments, or heuristic rules, for decision-making. They contrast sharply with programs using numerical calculation, which derive their power from analytical equations.

"Seldom are there questions, in the mathematical sense, that relate measurements of body parameters to the diagnosis of disease," Mr. Thomas Rindfleisch, director of the SUMEX-AIM computing facility, says. "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. They guide decisions about what other information is needed to determine the disease process involved."

Because medical diagnosis is described as an art, represented by imprecise concepts and relationships, AI researchers attempt to identify the principles of decision-making used by experts in drawing conclusions. For example, INTERNIST 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," Dr. Myers says. "As a result, the

Dr. Bruce Buchanan (right), adjunct professor of computer science at Stanford, is the developer of meta-DENDRAL, a learning program that discovers chemistry rules for use by DENDRAL. His program successfully defined rules concerning a group of organic compounds that had stumped experts. With Dr. Edward Shortliffe (left), physician, assistant professor of medicine and computer science at Stanford, he designed MYCIN, which is intended to help physicians diagnose infectious disease and prescribe treatment.

Dr. Dennis Smith, Stanford research associate in chemistry, is a member of the DENDRAL project.

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 the 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 from master to apprentice. That knowledge will now be codified and can become a working body to be used, taught, and critiqued by experts of the field."

In essence, 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, to establish the principles for solutions to new classes of problems.

But before this can happen, certain obstacles must be overcome. One is called, by SUMEX-AIM researchers, the "development gap." Essentially, this gap is the void that 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 development of new projects in that field, but no such industry yet exists for AI. Researchers say this branch of science is simply too young, too much of a risk, for businesses to jump into wholeheartedly.

"Typically, human—engineering issues have been solved by specialists who worry about marketability issues. There is as yet no AI industry with numerous engineers in various companies ready to take ideas out of the lab and convert them into products that people want," Dr. Feigenbaum says.

Some companies have expressed interest, but their interest must be encouraged by showing the worth of AI programs, the scientists say. This may be accomplished by developing a few select programs that are in advanced stages and are compatible with practical applications.

In recent years, SUMEX-AIM has attempted to grease the wheels that carry AI products into the marketplace by making certain programs easier to use. Although this role is somewhat beyond the strict confines of basic research and development, SUMEX-AIM investi-

INTELLIGENCE *continued*

gators say it is necessary in order to obtain information about the performance of their programs in practical applications and to expose these programs to the professionals they are intended to serve.

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," concludes Dr. Edward Shortliffe, physician and computer scientist working in SUMEX-AIM.

Paying more attention to "human engineering" will make computers more acceptable to physicians, he says. "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," adds Dr. Joshua Lederberg, the first principal investigator of SUMEX-AIM and now president of Rockefeller University. "Voice entry of data would make a very big difference, and there are some other technologies that need to be incorporated into the system."

Suggested ways of reducing resistance range from the mechanics of interaction with the computer—which might be achieved by using display terminals equipped with light pens, special keyboards, color, and graphics—to building features into the programs that make them appear more clearly as helpful tools rather than complicating burdens. 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," says Dr. Dennis Smith, a researcher working on CONGEN. "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 now have to show scientists that what we are doing could dramatically increase their productivity."

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. Researchers tend to put more trust in people and, by extension, AI programs that admit ignorance when appropriate, according to Dr. Bruce Buchanan, developer of META-DENDRAL.

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, and some of those might be quite legitimate about the abuse of such tools." According to Mr. Rindfleisch, society will have to learn to judge the effective use of these tools.

But the future seldom fulfills prophecy. No one can judge what will occur, based on current hardware or techniques, because these will undoubtedly change in unpredictable ways, just as bulky vacuum tubes were replaced by integrated circuits and plug-and-socket programming gave rise to software. 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.

The rewards that may result, however, appear to outweigh the effort and expense that lie ahead for those in this field. Through the experience gained in writing intelligent programs for specialized areas of expertise, AI researchers hope to fashion more general principles about intelligence. "There is the hope that by writing programs able to do this kind of reasoning, we will understand more 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." R

This article is based on excerpts from a special publication on artificial intelligence in medicine to be issued later this year by the Research Resources Information Center.