

HERBERT A. SIMON

DEPARTMENT OF PSYCHOLOGY
CARNEGIE-MELLON UNIVERSITY

OUTLOOK STUDY II
COGNITIVE SCIENCE AND LEARNING

AUGMENTATION OF HUMAN PROBLEM-SOLVING ABILITIES

Two themes run through cognitive science research, with its double emphasis on human cognitive processes and artificial intelligence for computers. The two themes are closely bound in several ways. The experience of the past quarter century of research has shown that many powerful ideas for artificial intelligence programs can be borrowed from human cognition, once that is understood; conversely, solutions to problems of designing artificial intelligence programs have offered many hints as to how the human processes might operate. There continues to be a close and fruitful partnership between these two streams of research.

But there is another connection between human cognition and artificial intelligence. Computers and people working in concert may be able to accomplish tasks that neither could accomplish, as well or at all, alone. We have already become used to thinking of the computer as an indispensable ally in large "number crunching" tasks, in science, engineering, and in business. We are just beginning to see the potential of man-computer cooperation over a much wider range of non-numerical symbol manipulating tasks.

One cooperative effort that is becoming visible to all of us is word processing. Here (in using word processing systems for editing manuscripts, for example) the computer serves mainly as a large, flexible, but mainly passive memory. It allows us to examine the text we have written, to revise it, to reproduce it, all without repeated retyping. Yet the computer, in this application, does not really provide any of the intelligence; that comes from the human user. (The computer can, however, construct an index and table of contents for the text, center headings, align margins, and even correct spelling, thus relieving the writer of some of the more tedious chores.)

The next step, which is just at or beyond the threshold of application,

is for the computer to participate in the problem-solving process. Automatic programs for medical diagnosis (for example, INTERNIST and MYCIN) have reached a stage where it is probably worthwhile to call them in as consultants on cases where the symptoms are confusing and conflicting. For a long time to come -- perhaps permanently -- we will want to leave the final diagnosis to the physician, but he can turn to the diagnostic program as an additional source of information and insight.

Similarly, a number of papers have been published in the chemical literature in recent years that report the identification of organic molecules, accomplished automatically by the DENDRAL computer program which draws its inferences from mass spectrograph data. Another program, MOLGEN, not quite as far along, provides advice for the design of research experiments in molecular genetics. Programs for the automatic synthesis of chemical compounds have been in the literature for nearly a decade, and it is believed (the information tends to be held as proprietary) that programs of these kinds are now being used in the chemical industry.

At the present time, there is a considerable revival of interest, in the United States, Europe, and Japan, in "industrial robotry." Part of this activity is concerned with the extension of techniques of factory automation of familiar kinds. Part, however, is concerned with building robots that can protect humans from various kinds of dangers by replacing (or augmenting) them as explorers in space, or as operators of equipment in areas where there is danger from radioactivity or the like.

For the short run at least, the greatest opportunities for using computers to augment human intelligence lie in the direction of interactive schemes that allow human and computer to work in cooperation. This involves such ideas, already partly realized, as organizing computer data banks so as to facilitate

human access to them, programming mathematical assistants (like the existing MACSYMA system) that can carry out routine parts (e.g., integrating an equation) of a mathematical derivation, or programming a system to detect errors in grammar in a manuscript.

As these applications develop, the boundary between what the human does and what the machine does will shift continually, as our knowledge increases about what would be truly helpful to the human, and as we achieve new technical understanding of how to provide that help. Computer intelligence will come to supplement human intelligence in the same gradually and multifarious ways that mechanical power came, in the course of the first Industrial Revolution, to supplement human and animal muscle.

VARIETIES OF LEARNING

The term "learning" refers to any change in a person (or animal) that enhances, more or less permanently, his or her ability to perform some range of tasks. By abandoning the search, popular several decades ago, for a few, uniform "laws of learning," cognitive science has advanced substantially its understanding of the processes that enable different kinds of learning to take place -- for example, acquisition of factual knowledge, of motor skills, of problem-solving strategies, or of perceptual recognition skills.

A principal research strategy for gaining an understanding of learning has been to analyse in careful detail the structure of the tasks that have to be performed by skilled practitioners in various domains, to detect -- using a number of laboratory techniques -- the processes that experts and novices use to perform those tasks, and to model the performers' knowledge, strategies, and other skills. The models now often take the form of computer programs that simulate the actual skilled behavior. Modeling expert and novice behavior then poses the next research question: what kind of process can modify the novice behavior until it becomes expert? What has to be learned, and how can that be accomplished?

What kinds of learning, for example, are involved in the long transition from medical student to physician? The physician must acquire a large body of knowledge -- about human anatomy and physiology, for example, and biochemistry. That knowledge must be supplemented by motor skills (taking a patient's blood pressure), perceptual skills (recognizing a symptom), and strategies for diagnosing illnesses and prescribing for them. Each of these components of the physician's mastery of his profession is organized differently from the others, and probably (although very little is known of the physiology of the matter) stored in different structures in memory.

Task analysis of physicians' skills has now progressed far enough so that the diagnosis and prescription processes can be modeled in considerable detail. In fact, computer simulations of diagnosis have been constructed that can perform at a respectable clinical level in the domains of internal medicine and microbial disease. Models of professional performance have reached comparable levels of competence in a number of other areas, as diverse as chess-playing, synthesizing chemical compounds, solving college-level physics problems, and identifying molecules from mass spectrogram data.

With the progress that has been made in detailed task analysis of complex cognitive tasks, cognitive research is now turning more and more to the learning processes themselves. One kind of learning that has been modeled successfully is the expert's acquisition of the ability to make fine perceptual discriminations among the wide variety of situations that are encountered in practice, and to recognize most of them "instantly" from a small number of cues. A computer program, EPAM (Elementary Perceiver and Memorizer) which employs simple mechanisms to accomplish such perceptual learning, has been used to simulate the chess master's intuitive grasp of the essential features of chess positions. The program has been used also to estimate the variety of familiar "chunks" an expert must be able to discriminate among -- some 50,000 in the case of the chess master. The program has also been used successfully to explain the findings of a wide range of classical experiments on rote verbal learning processes.

The EPAM program and similar efforts are concerned mainly with acquiring the "index" to the body of expert knowledge -- the recognition cues that are used to access that knowledge. Other research in learning is concerned with the content of the knowledge and the way it is organized in memory. It is now widely agreed that human long-term memory can be modeled, in first approximation, as a network of nodes, with little schemas (linked packets of information)

stored at each node, and numerous connections among these schemas. An important research question is how information presented in natural-language text can be transformed so that it can be incorporated in such schemas, augmenting and enriching the information contained in them.

Several research efforts have achieved some success in modeling this acquisition of knowledge from natural-language text. One computer simulation interprets the language of physics problems in college textbooks, creates a representation in memory of each problem situation, with the aid of this representation, sets up the appropriate equations, and solves them. This program does only a moderate amount of new learning, for it is already provided with schemas to represent the common kinds of objects (levers, masses) it will encounter in the problems -- it must simply assemble them appropriately.

Other computer simulations, working with slightly simpler puzzles, actually create new schemas from the input text. Programs of this general kind are now being applied to the study, for example, of how a child can acquire language, building up gradually in memory schemas corresponding to vocabulary elements and grammatical structures.

A somewhat different aspect of learning is the ability of some students to acquire new problem-solving strategies by examining worked-out examples in textbooks -- acquiring, for instance, the skill of solving algebraic equations by studying one or two textbook examples of such solutions. This process, too, has now been modeled successfully for several problem domains.

In the next few years other learning processes, in addition to those discussed above, are likely to be identified and modeled. The modeling efforts themselves are still at an early age of development, but the time should not be far off when educational psychologists can begin to develop and adapt the basic findings for application to the improvement of teaching and learning in the schoolroom. Already ideas are emerging from the research for the improvement of

textbooks (for example, by giving greater emphasis to the perceptual cues needed for discrimination learning) and for helping students in "learning to learn" (for example, by training them in the skills of learning from worked-out examples). These and similar ideas will need careful empirical testing before their usefulness in education can be evaluated, but we can expect a continuing stream of such ideas from the basic research that is now going on.

Only slightly more remote is the use of the learning procedures that have been modeled as a way of "growing" the stores of knowledge needed by artificial intelligence programs. At present, all knowledge acquired by a program must be provided to it explicitly by human programmers. The development of automatized learning procedures would relieve one of the most severe bottlenecks that today impedes the programming of artificial intelligence schemas that require extensive knowledge.

COGNITIVE OPERATIONS

The way in which the mind works can be investigated at several different levels of detail, each level answering a somewhat different set of equations. Three levels are particularly important: the physiological, the level of elementary information processes, and the level of cognitive strategies. As yet, little is known about the connections between the underlying physiological mechanisms and the information processes they support, yet even in the absence of such connecting links, rapid research progress is being made at all three levels.

Elementary information processes are investigated, for example, by studying the times required for reacting to simple stimuli or sequences of stimuli, for learning new materials and for retrieving them from memory. The analysis of errors also provides important cues, helping to discover, from the nature of the similarities between correct and erroneous responses, whether information is encoded in terms of auditory or visual features.

Elementary information processes include such steps as recognizing a visual stimulus, "scanning" lists in long-term memory for the presence or absence of information, retrieving an item from short-term memory, and the like. Characteristically, their duration ranges from a few milliseconds to a few hundred milliseconds. The generally accepted picture, today, of how they are organized hypothesizes (1) a short-term memory of very limited capacity that accounts for the narrow limits of attention; (2) a long-term memory in which are stored the recognition capabilities (the "index"), the organized knowledge schemas, and the motor skills and strategies, and (3) "input" (sensory) and "output" (motor) systems. A great deal of the research of the past two decades has been devoted to elucidating the crucial role that short-term memory limits play in all cognition. More recently, attention has been turning back

to long-term memory, its associative organization and the way in which new knowledge is added to it.

The entire system is viewed as a symbol manipulating system, which "reads" symbols through its sensory organs, "writes" them through its motor organs, and copies, stores, and compares them in its memories. It has the important characteristic that when it compares two symbols, it can behave differently depending on whether the two are judged to be the same or different (conditional branching). In these information processing theories, "symbol" means simply any pattern, capable of assuming a rich variety of forms, that can be stored and operated on in a physical system (a brain, a computer, a blackboard).

The basic hypothesis that guides research in cognition is that it is precisely these kinds of symbol manipulating capabilities, organized appropriately in strategies, that enable a physical system, whether computer or man, to address tasks intelligently. Of course, opinions differ widely as to how wide a span of the activities we call "intelligent" can be accounted for by the hypothesis; but the hypothesis is empirically testable, and at the present time, determining the limits to which it can be pushed provides a powerful and challenging research program for cognitive science.

Research on cognitive strategies, the third level, has made most progress in elucidating problem-solving processes, strategies for concept attainment, and strategies for memorizing. Much progress has been made toward describing the strategies that people use for understanding natural language text, although many complexities of language remain to be unraveled.

PROBLEM SOLVING

One of the first areas to which the new information processing concepts that arose after the Second World War were applied was problem-solving research. The initial tasks studied, mostly simple puzzle-solving tasks in the psychological laboratory, now appear rather simple, but when this line of research was initiated, they seemed to the researchers immensely more complicated than the rote verbal learning tasks and T-mazes with which cognitive psychology had been preoccupied previously. Today, research is moving steadily from puzzle-like problems to problems whose solution requires professional-level skill and extensive subject-matter knowledge.

Even the "simple" puzzles (e.g., getting the three missionaries and three cannibals safely across the river) are hard enough for intelligent adults, often requiring a quarter or half hour for their solution. Research on how subjects solved them in the laboratory (while their verbal protocols were recorded) revealed a few basic strategies that appear to be central to almost all human problem-solving. Problem-solving characteristically involved highly selective ("heuristic") search through large problem spaces, the selectivity being essential if the search was to be successful in reasonable time. It involved frequent use of means-ends analysis, solving problems by comparing present situation with goal, recognizing one or more differences between them, and applying operators stored in memory that would be relevant to reducing the differences.

These processes themselves, heuristic search and means-ends analysis, were not novel, having been familiar for some years both to psychologists and to many laymen. The important ingredients that have been added by cognitive research are, first, an exact specification of what is involved in carrying out these processes, and second, an empirical demonstration that such processes

are sufficient to enable a system to deal intelligently and successfully with problem situations. The precision is achieved by describing the processes as computer programs, the basic operations in the program corresponding to the elementary information processes that have been revealed by research at the second level of detail. The proof of sufficiency is accomplished by using the computer programs to simulate the actual human behavior, and showing that the programs can solve the same problems as the humans, organizing their solution efforts in much the same way that the verbal protocols and other evidence have shown that the humans organize theirs.

Even in dealing with relatively simple puzzle-like problems, people (and computer programs) are often lost in a maze of detail. One technique successful solvers often use is to abstract from the detail, retaining only essential problem features, reach a plan for the solution in the abstract domain, and then using that plan to guide solution when the details have been reintroduced. Such planning mechanisms have also been simulated successfully by computer programs, not only for puzzle-like tasks, but also for elementary high-school geometry.

A key question in understanding problem-solving is how the search efforts are organized and guided by strategies. Today, we have learned a great deal about both stimulus-driven and goal-driven search, and know that most problem-solving is guided by a combination of these. The key idea is that problem-solving behavior consists of a sequence of actions, each initiated or "fired" by recognition of some clues that indicate its relevance in the current situation. The cues may consist of perceived features of the environment (which may sometimes be simply the problem solver's worksheet), or of goals and sub-goals that are held in short-term memory. When the former, environmental, clues predominate, we say the behavior is stimulus driven, when the latter, short-term memory, clues predominate, we say that it is goal driven. The

strategy elements, each consisting of a set of stimulus or goal conditions linked to an action, are known technically as "productions," and a strategy organized in this way as a "production system."

Most contemporary programs for simulating human problem-solving behavior, most simulations of learning processes, and most artificial intelligence programs for performing complex tasks are constructed as production systems. The simplicity of the organizational structure -- all productions have the same basic form, and new productions can simply be added to the system -- has been a major factor in recent successes in devising programs that can learn and simulate human learning.

The same processes that have been identified as fundamental to problem-solving have also figured prominently in research on other cognitive skills -- in particular, concept attainment and the discovery of pattern in data. In concept attainment tasks, the subject must learn to discriminate among a set of figures, picking out "large circles," say, from all the others. In pattern discovery tasks, the subject must discern a regularity and formulate a rule to describe and apply it -- for example, discover the regularity in the sequence "ABMCDM---" and predict what letters will follow. The strategies used to perform such tasks, which form an important component of intelligence tests, are now fairly well understood. The chief abilities required, beyond general problem-solving skills, are skills in recognizing identity of symbols, and the regular succession of symbols on a list like the alphabet or the natural numbers. Again, an important outcome of this research is to demonstrate that a small set of rather simple skills suffices to account for behavior that is commonly used as an indicator of human intelligence.

A number of college-level courses for the development of general problem-solving abilities have been designed on the basis of the research that has been sketched here. Such courses are now being offered at a growing number of

schools, and they seem to be well received, although it has not yet been demonstrated with hard evidence that they represent a cost-effective use of students' time. They raise the age-old pedagogical issue of "transfer": must people be trained in the specific skills they are going to use, or can they be trained in more general skills which can then be applied to specific situations? The correct answer must be "some of each," but the present advance in cognitive science provides us with perhaps the first opportunity we have had to attempt to specify how much is "some." This is a central question in the design of curricula, which will almost certainly receive increased attention in the coming years, and which will likely receive considerable illumination with the help of the new research tools that are available.