among the radiocarbon ages is a question which cannot be answered on the basis of our present information.

Conclusions

The cores from the Norwegian and Greenland seas demonstrate the essential similarity between depositional processes in those seas and in the Atlantic. As in the Atlantic, there is clear evidence of the importance of slumping and of turbidity currents in redistributing sediments and in modifying the depositional record during postglacial time. In contrast, sediment believed to have been deposited during the last ice age is predominantly of glacial marine facies. Spitsbergen and the general region of the Barents Sea are the most probable sources of the ice-rafted material.

The fluctuations in numbers of tests of *Globigerina pachyderma* in cores from the Arctic Ocean indicate a climatically controlled variation in the thickness and continuity of the ice cover of the Arctic Ocean.

From our study of coiling direction in *Globigerina pachyderma* we have found that a northward shift of the 7.2°C isotherm took place at the end of the last ice age, and that the isotherm has never extended into the Norwegian Sea during the last 70,000 years.

The evidence in the cores indicates that the net movement of floating ice must have been from north to south in the eastern half of the Norwegian Sea, in contrast to the south-to-north current now flowing there. The evidence indicates that there was a confluence of ice drifting from the northeast and ice drifting from the southeast off the coast of Norway at about latitude 62° N.

References and Notes

1. B. Kullenberg, *Svenska Hydrograf. Biol. Komm. Skrifter Ny Ser. Biol.* **1**, 2 (1947).
2. E. Philippi, *Die Grundproben der Deutschen Sudpolarexpedition, 1901–1903* (Reimer, Berlin, 1910), vol. 2, pp. 411–616.
3. H. Holtedahl, *J. Sediment Petrol.* **29**, 16 (1959).
4. P. H. Kuenen and C. I. Migliorini, *J. Geol.* **58**, 91 (1950).
5. D. B. Ericson, M. Ewing, B. C. Heezen, *Bull. Geol. Soc. Am.* **62**, 961 (1951); ———, *Bull. Am. Assoc. Petrol. Geol.* **36**, 489 (1952); ——— and G. Wollin, in "Crust of the Earth," *Geol. Soc. Am. Spec. Paper 62* (1955), p. 205.
6. D. B. Ericson, M. Ewing, G. Woolin, B. C. Heezen, *Bull. Geol. Soc. Am.* **72**, 193 (1961).
7. B. C. Heezen and M. Ewing, *Am. J. Science* **50**, 849 (1952).
8. M. N. Bramlette and W. H. Bradley, *U.S. Geol. Surv. Profess. Paper 196–A* (1940), p. 1.
9. M. A. Peacock, *Trans. Roy. Soc. Edinburgh* **54**, 441 (1926).
10. F. A. Henson, *Geol. Mag.* **89**, 293 (1952).
11. E. S. Deevey and R. F. Flint, *Science* **125**, 182 (1957).
12. W. Schwarzacher and K. Hunkins, in *Geology of the Arctic* (Univ. of Toronto Press, Toronto, 1961), p. 666.
13. M. V. Klenova, *Geology of the Barents Sea* (Academy of Sciences of the U.S.S.R., Moscow, 1960), pp 1–367.
14. O. Sverdrup, *New Land* (London, 1904), vol. 2, pp. 1–466.
15. W. H. Dall, *Am. J. Conchol.* **7** (1872).
16. S. Ekman, *Zoogeography of the Sea* (Sidgwick and Jackson, London, 1953), pp. 1–417.
17. A. F. Treshnikov, *Priroda* **2**, 25 (1961).
18. W. Schott, *Wiss. Ergebn. Deut. Atlant. Exped. 'Meteor', 1925–1927* (1935), vol. 3, pt. 3, p. 43.
19. J. A. Cushman and L. G. Henbest, *U.S. Geol. Surv. Profess. Paper 196–A* (1940), p. 35.
20. W. Schott, *Medd. Oceanografiska Inst. Göteborg, No. 18* (1953), p. 1.
21. F. B. Phleger, F. L. Parker, J. F. Peirson, *Swedish Deep-Sea Expedition (1947–1948) Repts.* (1953), vol. 7, pt. 1, pp. 1–22.
22. J. Kane, *Micropaleontol.* **2**, 287 (1956).
23. D. B. Ericson and G. Wollin, *Deep-Sea Res.* **3**, 104 (1956); *Micropaleontol.* **2**, 257 (1956).
24. D. B. Ericson, W. S. Broecker, J. L. Kulp, G. Wollin, *Science* **124**, 385 (1956).
25. O. L. Bandy, *J. Paleontol.* **34**, 671 (1960).
26. D. B. Ericson, *Science* **130**, 219 (1959).

27. We thank A. P. Crary and K. Hunkins for obtaining the Arctic Ocean cores and John Ewing for obtaining cores from subarctic seas. We also thank Janet Wollin for able assistance in the laboratory investigation and Arnold Finck and other members of the staff of Lamont Geological Observatory who directly or indirectly contributed to this study. We are grateful to W. Broecker for the results of radiocarbon dating. Mark E. Burgunker translated the Russian papers to which we refer. The research was supported by the National Science Foundation (grants NSF-G-14197 and NSF-G-23662) and by the Arctic Institute of North America (grant-in-aid project Onr-130). This is Lamont Geological Observatory contribution No. 713.

Information-Processing Models in Psychology

Though hard to present and verify, they open up basic new theoretical and methodological opportunities.

Walter R. Reitman

In 1957, Newell, Shaw, and Simon (*1*) published a description of their Logic Theorist. A program for proving theorems in elementary symbolic logic, the Logic Theorist could be used to generate behavior from a digital computer. Newell, Shaw, and Simon thus were able to demonstrate that it actually was able to solve problems only humans had been able to solve before.

It soon became apparent that information-processing programs like the Logic Theorist were of great interest to psychology. They incorporated strategies and rules of thumb similar to those used by humans, and the behaviors they generated were in some ways strikingly similar to the behaviors of humans working at the same problems. As a result, there has been a vigorous growth in psychological research utilizing information-processing methods and concepts. In addition to problem-solving programs, there now are programs that simulate learning, perception, attitudinal processes, understanding, and even neurotic personality processes, as well as a growing literature dealing with the relation of this work to other forms of psychological theory and research (*2*).

Information processing is still far from integrated into the main body of psychological thought, however, and because it is complex and new, rather than simply a new twist to widely accepted ideas and procedures, the advantages and limitations involved are only imperfectly understood. Therefore, after characterizing the approach and some of the ways it has been regarded and used, we will examine some of the major questions that must be dealt with if these techniques and concepts are to achieve in full the general psychological significance now being claimed for them.

The author is associate professor in the Graduate School of Industrial Administration and the Department of Psychology, Carnegie Institute of Technology, Pittsburgh.

Some Characteristics and Uses

It might be well to say first of all that "information processing" is not synonymous with "data processing" or with "use of computers." As I use the term here, the "information-processing approach" refers to one way of looking at psychological activity. It deals with processes and functions; it emphasizes whatever it is that any particular behaviors *get done*; it is also concerned with the fine structure of behavior. The accomplishments resulting from thinking, problem solving, and psychological activity generally can be accounted for only if we study them in great detail. When we do so, we discover that even simple behaviors appear to be made up of a great many steps integrated into complex sequences. Information-processing computer languages encourage us to specify precisely and explicitly the systems of cognitive structures, elementary psychological processes, and higher-order strategies we induce from behavior in order to account for the achievements we observe. In other words, this approach allows us to view men as dynamic systems analyzing, seeking, and doing things, as purposive organisms manipulating objects and information to achieve ends, rather than as points to be located in some multidimensional space defined in terms of variables derived statistically from tests and measures on large populations of subjects. With the latter mode of thinking, psychologists seek to discover the "basic variables" or "factors" presumed to underlie human behavior. Psychologists with the information-processing viewpoint try to make functioning models of psychological structures and processes by which they can reproduce and thus account for human behaviors.

Proponents of information-processing models and concepts view and use them in a number of more or less distinct ways. I shall outline a few of these very briefly, to provide both a better understanding of the range of uses to which such models may be put and a basis for the discussion of problems that follows.

One of the broadest views on the uses of information-processing techniques is that of Armer (*3*), who suggests that we regard the many kinds of natural and artificial intelligence as elements of a single, complex domain defined by a great many measures of possible functional and structural similarity. Armer thinks of the boundary between research in psychology and in artificial intelligence as flexible and relatively unimportant. We can acknowledge both similarities and differences between humans and information-processing programs, and we should use information about one type of intelligence in studying or constructing another precisely to the extent that we believe them to be similar, no more and no less. The approach depends upon a de-emphasis of the man-machine dichotomy in favor of a multitude of points of similarity and difference; therefore in what follows we shall want to consider whether perhaps there may not be critical aspects of human intelligence that we cannot reasonably expect to represent or reproduce in an information-processing program.

Newell and Simon (*4*) are representative of a sizable group sympathetic to Armer's position but interested specifically in pinpointing the processes and structures underlying *human* intelligence. Typically they collect extensive data from human subjects asked to "think aloud" while solving problems under laboratory conditions, and then they try to write programs that will simulate the behaviors observed in their data. They view a program capable of simulating behavior as a theory of the system of psychological processes and structures underlying the behavior. Such theories are held to have a status comparable to those framed in words or in mathematical symbols and to be subject to the same criteria of adequacy. As we shall see, however, as of now these theories and the behaviors they simulate are in several important ways *not* comparable to the structure and content of most other psychological theories, and we shall want below to consider some of the implications of these differences.

A somewhat different view of the uses of information-processing programs emerges from noting that they enable us to state and explore the consequences of systems of psychological assumptions in a way not possible either with theories framed in words or with direct experimentation. With verbal models, it is practically impossible to be sure that conclusions follow only from explicit assumptions and that they in no way depend upon "unprogrammed" elements entering informally into the argument. With laboratory experiments, we cannot get a test of the theory in and of itself. We must settle for a test of the theory taken together with all the assumptions about manipulations, measures, and conditions that couple the theory by means of operational definitions to the real world. If unexpected results occur, we are unable to say whether the difficulty is in the theory, the ancillary assumptions, or both. In an information-processing model, as in a mathematical model, we can state, manipulate, and deduce implications from our theories in a way that is at once sure, unambiguous, and yet independent of operations relating the theory to data on human behavior. The use of information-processing models in this way may teach us a great deal both about the functions we wish to comprehend and about the theory we have constructed, somewhat as the writer of a programmed text may discover new things about his subject in the course of making the details of its structure explicit. In the long run, of course, there is no point in constructing a model unless we test it against human behavior, and in postponing empirical verification we must take care to avoid letting our work degenerate into abstract exercises in theory construction. In the short run, however, there are considerable advantages in being able to deal with the development of models and with the construction of measurement and empirical-verification procedures as separate and distinct problems.

Note, finally, that since most information-processing models are built around a limited number of key concepts and mechanisms, computer runs of such a system provide concrete experience with the ways its mechanisms work and the things they can or cannot do. Now if we encounter similar functions in studying other psychological problems and activities, we can use our knowledge of these mechanisms in thinking about and accounting for these activities. These ideas about information-processing functions, structures, or mechanisms may be held and used as imprecisely as any other ideas, of course, in which case they retain little more than their conceptual connections with the original programmed model. In other words, just as theory construction without verification may degenerate into an abstract exercise, so a loose use of information-processing ideas apart from the particulars of a program may degenerate into sheer talk. Judiciously employed, however, informal information-processing concepts may be of substantial value as analogical tools in thinking about and developing theories and experimental designs, especially in areas outside the bounds of current programs. Furthermore, to the extent that they are stated clearly and in detail, the possibilities for exploration and evaluation inherent in programs remain open over the longer term.

Programs Serve Multiple Functions

Information-processing systems are distinct from other forms of psychological theory in several ways, and since these peculiarities raise basic questions regarding the psychological significance of such systems, we shall want to analyze them in some detail. Unlike verbal or mathematically stated theories, a program is, simultaneously, a *statement* of a system, a sequence of *computer instructions*, and finally, an *operator* for achieving certain ends, for example, the solution of complex problems. This situation has important advantages, but it also complicates the conceptual status of computer models.

We already have noted that, because the program is both theory and sequence of instructions, one can deduce strictly the implications of the theory simply by running the program and observing the behavior generated under various initial conditions. Similarly, the status of the program as an operator for handling complex tasks makes it of practical interest: whether or not it simulates human behavior, a particular program may well be of value as an artificial means of handling the complex information-processing task in question (*5*). Here, of course, is one of the main links between computer simulation and artificial intelligence.

That a program must run on a computer as well as serve as a theoretical statement obviously implies additional constraints on the program. But what exactly are the effects upon the form and substance of the theory itself? In other words, what follows from the fact that a system intended as a psychological theory is at the same time a running program? Perhaps certain human information-processing functions or complexes of functions cannot be modeled successfully in a computer language. In that case, the information-processing approach will turn out to have been concerned with a relatively special and limited conception of human intellectual activity.

A few psychologists already appear to have concluded that representation in a computer program somehow necessarily implies a "computer-like" psychological theory. Mowrer considers Newell, Shaw, and Simon to "take the digital computer as the model for mental operations," the Kendlers speak of models "dependent upon . . . the operation of computers," and Tracy Kendler flatly concludes that programs of the sort we have been discussing "add nothing to our further understanding of the living mechanisms, but they do provide a better understanding of the computer" (*6*). None of these comments, however, either describes the limitations implied or makes clear just why it is necessary to presume such limitations exist.

Quite another matter is Neisser's (*7*) interesting and sophisticated effort to chart what he takes to be the psychologically significant limits of computer models. Readily granting that computers can be programmed to learn from experience, to behave intelligently and purposively, and on occasion to come up with novel and perhaps even creative results, Neisser nevertheless holds that "the procedures which bring about these results differ substantially from the processes which underlie the same (or other) activities in human beings," and he makes a number of specific points to substantiate his assertion. Unlike computers, he argues, people get bored. "When a program is purposive, it is too purposive." Human memory, furthermore, is less flexible than that of a computer, which can learn and unlearn completely at a command. "A man rarely has single-minded control over what he will learn and forget; often he has no control at all. . . . The result is both stupidity and serendipity." Then, too, computers lack conflicts and "do not get tangled up in conflicting motives. . . . They are good at problem solving but they never solve problem *B* while working on problem *A*." Neisser also suggests that unlike information-processing in a computer, human thinking always occurs as part of "a cumulative process of growth and development . . . in an intimate association with emotions and feelings . . . [and] serves not one but a multiplicity of motives. . . ."

If Neisser's argument has a weakness, it is that he does not always keep separate three interrelated and yet quite distinct issues: the importance of the man-machine dichotomy; the limitations of current information-processing models; and the prospects for computer simulation as a psychological tool generally. One may share Neisser's doubts "that machines will think as man does," and still note that Newell, Simon, and others also have pointed out the fundamental dissimilarities between men and computers. All they assert is that *functions* in the one may be reproduced with the help of the other. The aspects of human intelligence they feel they have incorporated in their models are in their programs, not in the computer, which is a tool rather than a simulacrum.

By the same token, it is hard to argue with Neisser's assessment if it is read as a critique of current programs. Somewhat related analyses have been made by Reitman, Grove, and Shoup (*8*) as well as in what no doubt is the most detailed and thorough treatment of the limitations of current programs available, Newell's own (*9*). These independent discussions reinforce Neisser's argument as it applies to the programs available at the time he wrote. But though Neisser prods model builders to a broader view of human intelligence in their programs, he also several times expresses doubts that ideas like his own can in fact be included in information-processing models, at least in the near future.

The hardware peculiarities of a computer clearly make some types of models easier and cheaper to work with than others. It is not clear, however, that the choice of a program as the vehicle for a theory effectively rules out any particular theory. The force of this classic Whorfian view as it applies to what we may say and think in a computer language remains to be seen. As evidence for the defense, we may note that Reitman, Grove, and Shoup, working with psychological assumptions in many ways similar to Neisser's, have programmed a model of human thinking (Argus) that displays many of the characteristics of human information-processing he cites. A system based upon a model of active cognitive structure derived from Hebb (*10*), Argus is not at all singleminded, is much less in control of what it remembers and forgets, and also is much more prone to conflict and serendipity. In fact, the assumption that work on a problem may and will be interrupted by the occurrence of interesting and unexpected ideas about other matters is built into the basic structure of the program.

In sum, Neisser makes an invaluable contribution by analyzing and pointing out important aspects of human intelligence not yet touched upon in information-processing models; but the overall import of his argument is perhaps a shade too dark.

Detail, Change, and Description

If we compare information-processing programs with psychological theories set forth in mathematical notation, the difference in the sheer volume of detail is striking. For example, the description of a recent version of the Newell, Shaw, and Simon General Problem Solving program (GPS) runs to more than 100 pages (*11*) and even so covers only the main details of the system. Furthermore, the discussion assumes a knowledge of an earlier basic paper on GPS (*12*) and a knowledge of Information Processing Language-V (IPL-V), the computer language in which it is written (*13*). Finally, the appendix, which simply *names* the routines and structures employed, takes another 25 pages. Unless one is familiar with similar systems, a thorough grasp of the dynamic properties of so complex a model almost certainly presupposes experience with the running program and its output.

What accounts for this mass of detail? We can suggest three sources. The first is the complexity and comprehensiveness underlying the aims of the research. GPS is to reproduce and account for highly complex human behavior, blow by blow, and at a level of detail never before attempted for so demanding a task. Secondly, languages like IPL-V are extremely flexible, since they are designed to express a great variety of systems. To make this possible, they have implicitly built into their structure many *fewer* assumptions than mathematical languages have about the context and organization of the objects and processes we may define in them. Consequently, however, when we define a particular model in a program, a good deal more detail about its objects and processes, how they are handled, and what they do must be stated in the language explicitly. Like a custom-built house or car, an IPL-V program provides much more latitude for individual choice, but at the cost of requiring a good deal more information about the particular arrangements decided upon. Thirdly, a good bit of detail is required to define matters so that the computer actually is able to carry out the necessary operations. By contrast, in the case of a mathematical model set down on paper for human readers, prior knowledge of such detail is assumed. It should be observed, incidentally, that only the last of these three sources of detail is a direct consequence of the statement of the system as a *program*. The first two are integral to the structure of the theory itself. They have nothing to do with the computer, which is best thought of here as a highly accurate and rapid labor-saving machine utilized solely to crank through the operations that follow from the theory.

Given this volume of detail, however, how is the theorist to communicate his theory? And when he describes what the system does, how is he to make clear the set of constraints that limit the conditions under which it does it? Some details obviously are of more importance than others. There are, for example, several ways to segregate the various program components by function and level. But these distinctions are informal: one does not generally know to what extent the results at higher levels are strictly independent of the particular low-level coding procedures the theorist has specified. As a consequence of the detail problem, though particular sets of psychological assumptions may be *embedded* in a system, they do not, of themselves, *define* the resulting program that incorporates them. In any given case, the operation of the system depends both on the psychological assumptions and on the specific encoding by means of which the assumptions are made concrete and supplemented by such additional detail as is required to specify the system fully so that it generates observable implications about behavior.

To make the communication problem completely clear, suppose that GPS or some other system were able to simulate in detail a wide range of problem-solving performances. Faced with such results, a number of psychologists interested in human cognitive processes might undertake to acquire a thorough understanding of the system. All others, however, would be without direct access to the theory itself (if we take literally the assertion that the program or the system itself *is* the theory). Their understanding of the system would be entirely dependent on an intermediary.

Contrast this with the situation which obtains in the case of typical mathematical models, for example, Estes's work on rational learning curves of the form

$$E_t = (1-1/N)\ (1-c)^{t-1},$$

with E_t the number of errors made on the tth trial, N the number of alternate responses available, and c a parameter such that $(1-c)^{t-1}$ is the number of unlearned items at the beginning of the tth trial. Here the theory is entirely explicit. Anyone with a knowledge of the mathematics involved may grasp the system and reason about it directly, entirely without an intermediary. Another student of the problem who thinks that a particular derivation or result depends critically on some special feature, assumption, or boundary condition of the model can investigate the model himself to see whether his expectations are confirmed (*14*).

One might perhaps object that anyone with a knowledge of IPL-V, a copy of the program deck, and access to a computer also could reason directly about systems such as Argus or GPS. Certainly this is true, but it does not negate the fact that, given equally good knowledge of the mathematical language underlying Estes's model and of the information-processing language underlying GPS or Argus, communication and independent analysis are just a very great deal easier in the one case than in the other. For most practical purposes, in fact, the difference in amount is so large as to be a substantial difference in kind.

Many of the same remarks can be made with respect to the problem of *system change*. Here again the basic difficulties have to do with the mass and complexity of the detail. When a mathematically stated theory is modified, one may compare the two forms directly and draw one's own conclusions about the extent and implications of the changes. But when a system like GPS is changed (*15*), one is in practice entirely dependent upon an intermediary. The intermediary, usually the system designer, will try to point out the major changes and their implications. But he obviously cannot go into detail, and his remarks are not likely to provide much of a foundation for rigorous reasoning about the system by the reader. Once again, as things now stand, there is simply no convenient, explicit form in which to communicate, and as a consequence the distance between the information-processing model and the reader is very much greater than with mathematical models.

Experimental Verification

Modern experimental methods and their associated statistical techniques provide an efficient, precise, and widely accepted basis for determining and communicating to others the extent to which a psychological model satisfactorily accounts for a class of behaviors. But the framework typically presumes conditions, for example, the existence of a metric and independent trials, that generally are not satisfied by the data information-processing models seek to simulate. How then can we determine for ourselves or for others that a particular information-processing model in fact adequately accounts for any given sequence of behaviors?

There are, of course, a number of different types of models. In one kind, for example, Feigenbaum's (*16*) model of rote verbal learning, the model corresponds to a generalized or abstract individual. In these cases, as Feldman (*17*) suggests, conventional statistical approaches may be quite helpful. Feigenbaum and Simon (*18*) report just such a test of predictions from that verbal learning model, and with excellent results.

Models of information processing in single individuals such as GPS, however, generate long strings of decisions, evaluations, and actions. Satisfactory tests of adequacy with this kind of highly conditional sequential data are more difficult to achieve.

Feldman considers three possible approaches for such cases. The first is based upon simple counts of differences between the streams of behavior generated by subject and program. Such counts are of limited usefulness, however, because of the sequential dependencies which characterize complex behavior. That is, if the model errs at a step, it is difficult to decide whether subsequent errors should be attributed to further difficulties in the model or instead treated as indirect consequences of the earlier error (and thus not to be counted against the model).

Feldman's second technique is conditional prediction. Whenever program and subject differ, an error is scored just as in the simple difference-count procedure. Now, however, the program is set back and on the track. That is, the program's decision is withdrawn and the subject's substituted instead. In principle, this procedure should correct for the inadequacy of simple difference counts in dealing with conditional dependencies. That is, subsequent differences between program and subject should reflect new errors by the model, rather than the propagating effects of previous errors.

Feldman's interest in conditional prediction stems from its usefulness in connection with his own work (*19*) on a model of the cognitive processes underlying behavior in the binary-choice experiment (in which a subject is asked to predict which of two events, E_1 or E_2, will occur on each of a long series of trials). In this situation, the technique does indeed provide very useful information both about the overall performance of the model and about its local strengths and weaknesses.

Use of a conditional-prediction measure assumes that the initial states of the subject and the program are in correspondence as each unit of behavior begins, either because the last unit has been predicted correctly or else because the program has been reset. In Feldman's own model, the results of a unit action are localized in effect. The behavior of its "pattern evocation" mechanism, for example, is to generate a list of patterns which in turn serves as the input to the "pattern selection" mechanism that follows. If the subject's protocol suggests that the pattern-evocation procedure has erred at some point, the error may be scored and the system then set back on the track simply by modifying the generated pattern list appropriately.

Other systems, as Feldman himself notes, may be much harder to set back on the track in a practically meaningful sense. Argus, because of the constant changes taking place in the states of its active cognitive elements, provides a good, albeit extreme, illustration of the problems involved. Suppose we wanted Argus to simulate subject S as he solved problems while thinking aloud. Suppose, further, that our protocol were good enough so that we could match subject and program at the end of every Argus step. What, now, if the system errs? To realign Argus and the subject, we would have to know all of the modifications made by each in their respective cognitive-element systems (so that we could replace the incorrect modifications with the correct ones), or else we would have to have a complete record of the current state of the subject's cognitive elements (which would then be reproduced in Argus).

Argus, then, is a system that in practice cannot be set back on the track in Feldman's sense. The changes made in the course of its operation are too *numerous* and too *diffuse*. Furthermore, even if we were to trace every modification in the state of the cognitive structure made by the system during a step (which would be possible with the IPL-V system's monitoring facilities, though extremely time-consuming), there is no way in which we could obtain the corresponding and equally necessary information from the subject. Even programs like GPS would appear to involve changes sufficiently complex and diffuse so as to render the conditional-prediction method of uncertain value in such cases. Thus the method, though valuable in connection with some types of programs, does not in fact offer a readily generalized approach to the verification problem.

The third approach Feldman considers is one first proposed by Turing. The basic notion here is very simple. As Feldman observes, "We are all familiar with the advertisements that challenge the reader to distinguish between oleomargarine and the 'high-priced spread.' This same type of test might be used to see whether an expert could distinguish between the subject's protocol and the machine's. . . ."

Feldman considers this method interesting but limited. It will tell the researcher "whether or not he has achieved a program that will produce behavior that is indistinguishable from human behavior." But since it depends on the unknown detection and decision-making characteristics of judges, it lacks operational rigor. Furthermore, it probably is a poor test of a system's ability to simulate the behavior of a *particular* subject. The use of judges is standard procedure in many experimental and psychometric settings, however; and though one cannot define their characteristics as measuring instruments, one ought not underestimate the value of knowing whether or not such a sample of humans is able to distinguish the simulated behavior from the original. Those less concerned to reproduce in detail behavior of particular individuals than to generate information-processing behavior globally related to human activity may well find Turing's test of interest. It provides a crude measure of family resemblance, and of course it in no way rules out development and application of other, more precise measures of similarities and differences.

One other factor also complicates verification. The emphasis upon simulation sometimes is attended by a deemphasis of the usual distinction between reproduction and prediction of data, at least by the usual standards of experimental psychology. To see just what is involved, consider Feldman's discussion of his own working procedures.

> The completion of the model was a lengthy task involving the iterative procedure of proposing a detailed model, testing the model against the data, modifying the model, testing again, and so on. During this procedure, almost every part of the model originally proposed was modified or replaced. In addition, modifications of the model led to new interpretations of the data, and these interpretations led to further modifications of the model.

At least in some current practice, then, simulation does not necessarily involve prediction of new or previously unseen data. In the instance Feldman describes, and in similar cases, there is a continuing interaction between the data and the model being constructed to fit them. Often, furthermore, the model builder has a range of protocols to choose from, so that there is the possibility of additional interaction between his selection of a particular protocol for intensive study and the type of model he tries to build. This amounts to still another unknown factor one must consider in attempting to evaluate reported results, and all of the usual caveats regarding capitalization on error apply.

Computer Simulation of Thinking

In view of the difficulties presently associated with the description of information-processing models and of the extent to which they account for the complex human behaviors they are intended to simulate, how do those who view them as psychological theories in fact go about assessing and communicating their results? What kinds of comparisons do they make, and what kinds of conclusions do they draw?

A recent article by Newell and Simon (*4*) summarizing their work on GPS and their evidence for its utility as a theory will serve as a reasonably typical instance of the procedures followed. After describing the main concepts incorporated in GPS, they suggest that one may evaluate the adequacy of such a theory at several levels.

> At the grossest level, we may ask whether the program does, in fact, solve problems of some of the sorts that humans solve. This it demonstrably does . . . [At an intermediate level] the general kinds of means-end analysis that the General Problem Solver uses are also the methods that turn up in the subjects' protocols. We have examined in fair detail some 20 protocols of subjects solving logic problems. . . . Virtually all the behavior in these protocols falls within the general framework of means-end analysis. The three goal types we have described account for about three-fourths of the subjects' goals, and the additional goal types that appear in the protocols are closely related to those we have described. The three methods we have outlined represent the vast majority of the methods applied to these problems by the subjects.

The third and most specific level involves a highly detailed comparison between program output and the line-by-line behavior of individual human problem solvers. One advantage of the status of information-processing models as programs is that when a program such as GPS is run, the computer may be instructed to trace and print out in as much detail as we wish to have the sequence of comparisons, judgments, manipulations, and decisions the program generates. This enables us to compare almost word for word what the program considers, tries, or rejects with the subject's ongoing reports of his own activity. Newell and Simon present some 30 lines each of human behavior and corresponding program trace, noting both the many rather striking similarities and the several sorts of differences to be observed between the two records. They then conclude that, all in all,

> The fragmentary evidence we have obtained to date encourages us to think that GPS provides a rather good approximation to an information-processing theory of certain kinds of thinking and problem-solving behavior.

Work on verification of information-processing models is proceeding currently in several directions (*20*). Newell and Simon plan to include in a forthcoming book several *hours* of protocols of human problem solvers—materials running to several thousand words. There will be highly detailed comparisons between these data and the output of versions of GPS designed to simulate the behavior of the individual subjects originating the protocols. With this volume of material available, those who wish to do so will be better able to study the primary evidence and come to their own conclusions about GPS and the information-processing approach generally.

Efforts also are being made, along somewhat different lines, to provide more precise and adequate measures of the fit between original and simulated behavior. Recall that relatively minor variations in a program may generate major behavioral changes by throwing the program off one track and onto another. Major discrepancies between human and simulated behavior also may be expected to stem at times from correspondingly small differences between the sets of information processes involved. But suppose one sample of behavior is used to specify a program, and then a second sample from the same subject, previously unseen by the programmer, is used in evaluating the program. If the program fails to predict the new sample perfectly, the programmer modifies the program until it does so, and then counts the number of elementary changes he has had to make in the original program to achieve this perfect simulation. If substantially all of the program actually is called upon during the test run, and if the number of differences is small, one might be willing to consider the original program a very good approximation of the system of processes underlying *both* samples of the subject's behavior. As the differences increase, of course, there is increasingly less reason to be satisfied with the model. Note that the count of changes will vary to some extent with the skill of the programmer making the changes, and it provides no real *measure* of the amount of change as it stands. But talking about the similarity of one *program* to another, as opposed to the similarity of one behavior to another, constitutes an interesting and promising innovation in the search for improved verification and communication procedures.

Conclusion

A psychologist considering learning enough about information processing approaches to use them in his own research faces questions of several sorts. Knowledgeable persons disagree somewhat about the relative importance of the human functions modeled so far, and understandably more about the range of functions likely to prove amenable to this kind of approach. There may be differences between men and machines so fundamental that such models will be unable to go beyond what already has been achieved. Alternatively, at the other extreme, information-processing formulations may apply so completely that much of psychological research will become the study of complex information manipulations in natural systems.

There also is disagreement about the relation between programmed models and the more usual forms of psychological theory. At what point does the complexity of a theory become so great as to render it useless as a scientific tool? How seriously ought one be concerned about the fact that these models cannot yet be subjected to testing and verification in any generally accepted manner? We should note the complexities of the behaviors they attempt to explain, and the impressive and detailed correspondences between real and simulated behavior that have been adduced. But it remains true that experimental psychology evaluates theories by ground rules these models do not now meet.

Having reviewed their drawbacks, we ought not forget the things information-processing models *do* do. They enable us to think about and represent functions involved in extremely complex human activity, in a form that is precise, objective, and as detailed as we wish to have it. They also allow us to generate behavioral consequences from a computer, and thus to study the strict implications of our theories. For the kinds of complex human behavior discussed here, there is no other approach remotely comparable in any of these respects. Finally, the drawbacks, though serious, may be temporary; the advantages and contributions are permanent. In short, unless one believes in an absolute functional dissimilarity between men and machines or in the absolute primacy of experimental-test conventions, there is good reason to view these models as one of the two or three most important bodies of new methods and ideas to become available to psychology in this generation.

References and Notes

1. A. Newell, J. C. Shaw, H. A. Simon, *Proc. Western Joint Computer Conf.* (1957), pp. 218–230.
2. E. A. Feigenbaum and J. Feldman, Eds., *Computers and Thought* (McGraw-Hill, New York, 1963); E. B. Hunt, *Concept Learning* (Wiley, New York, 1962); J. McV. Hunt, *Intelligence and Experience* (Ronald, New York, 1961); G. A. Miller, E. Galanter, K. H. Pribram, *Plans and the Structure of Behavior* (Holt, New York, 1960); S. Tomkins and S. Messick, *Computer Simulation of Personality* (Wiley, New York, 1963).
3. P. Armer, reprinted in *Computers and Thought*, E. A. Feigenbaum and J. Feldman, Eds. (McGraw-Hill, New York, 1963), p. 389.
4. A. Newell and H. A. Simon, *Science* **134**, 2011 (1961).
5. H. A. Simon, in *Management and Corporations, 1985*, M. Anshen and G. L. Bach, Eds. (McGraw-Hill, New York, 1960).
6. O. H. Mowrer, *Learning Theory and the Symbolic Processes* (Wiley, New York, 1960 H. H. Kendler and T. Kendler, *Psychol. Rev.* **69**, 1 (1962); T. Kendler, *Ann. Rev. Psychol.* **12**, 447 (1961).
7. U. Neisser, *Science* **139**, 193 (1963).
8. W. R. Reitman, R. B. Grove, R. G. Shoup, *Behav. Sci.*, in press.
9. A. Newell, in *Self-Organizing Systems, 1962*, M. C. Yovits, G. T. Jacobi, G. D. Goldstein, Eds. (Spartan, Washington, D.C., 1962).
10. D. O. Hebb, *Organization of Behavior* (Wiley, New York, 1949).
11. A. Newell, "A Guide to GPS–2–2 Program" (Carnegie Institute of Technology, Pittsburgh, 1962).
12. A. Newell, J. C. Shaw, H. A. Simon, *Proc. Intern. Conf. on Inform. Process.* (UNESCO, Paris, 1960), p. 256.
13. A. Newell, Ed., *Information Processing Language-V Manual* (Prentice-Hall, Englewood Cliffs, N.J., 1961).
14. Simon, for example, has shown that this particular mathematical model for paired associate learning in fact is a special case of a more general function. Noting that this more general function "is the product of two functions and assumes that the discovery and fixation aspects of learning are independent," he then is able to demonstrate that Estes's data do not provide a sharp test of his assumptions about the course of the fixation process in learning [H. A. Simon, *Psychometrika* **27**, 417 (1962)]. Good descriptions of a computer model such as GPS may make possible informal consensus regarding its general properties and limitations, but they do not and cannot provide anything of this type of direct access to the system under consideration itself.
15. GPS is a useful illustration of the magnitude of the change problem, since Newell (see *11*) provides a summary of its history:
"The first version, called GPS–1, was coded in IPL–IV . . . The most complete description of GPS existing in the published literature . . . gives only the highest level organization . . .
"GPS rapidly outgrew the small storage capacity of JOHNNIAC [the first computer on which it was run] and it was recoded in IPL–V to run on the 704–709–7090 series machines, which have 32,576 words of fast storage. The new program was called GPS–2–1. Functionally it was almost identical to GPS–1, but substantial organizational changes were made. The change to GPS–2–2 involved somewhat smaller organizational changes, but required a separate designation, since both versions were running at the same time . . . Additional versions, GPS–2–3 and GPS–2–4, now exist. They involve more substantial organizational changes from GPS–2–2, and will be documented separately."
16. E. A. Feigenbaum, *Proc. Western Joint Computer Conf.* (1961), pp. 121–132.
17. J. Feldman, in *Computer Applications in the Behavioral Sciences*, H. Borko, Ed. (Prentice-Hall, Englewood Cliffs, N.J., 1962), p. 337.
18. E. A. Feigenbaum and H. A. Simon, *Psychol. Rev.* **68**, 285 (1961).
19. J. Feldman, *Proc. Western Joint Computer Conf.* (1961), pp. 133–144.
20. A. Newell, personal communication.

21. I thank the Carnegie Corporation of New York for encouragement and financial support, and B. Green, U. Neisser, A. Newell, and T. Standish for suggesting improvements in the manuscript. The article is based on a forthcoming book tentatively titled "Cognition and Computer Simulation" (Wiley, New York, in press).