LEONARD UHR

Tryouts toward the Production of Thought

This paper explores what we might mean by terms like "sense," "perceive," "know," "consciousness," "self," "short-term memory," "analog," "schema" and "cognitive model," by trying to find what they refer to in actual computer programs. The different types of systems for perception are briefly surveyed, to indicate structures that they must have in common. Then more wholistic cognitive systems are described.

One system (Uhr, 1975a, b, 1976a) called a SEER[1] attempts to handle simultaneously all the cognitive processes—including perception, remembering, deductive problem-solving, language processing, acting, and learning—that are usually studied separately by psychologists and computer scientists. SEERs are first attempts to develop well-integrated wholistic cognitive systems that are designed to do a variety of things, albeit drably (much as most human beings go about their everyday tasks of thinking), rather than one particular difficult thing well (like chess or proving theorems). They are therefore required to decide what type of thing to do, and when, and to choose, organize, integrate, and orchestrate the sets of cognitive transforms needed to carry out the chosen actions, cutting across the separate cognitive systems.

RATIONALE

Our problem is the development of a science of information-processing entities, a science coordinate with physics—the science of matter/energy. Psychology and the building of "artificial intel-

Note: The author's research described in this paper was partially supported by the National Science Foundation and the University of Wisconsin Graduate School.

ligences" will someday be viewed as the natural and artifactual applications of our science, just as geology and engineering are today viewed as the natural and artifactual applications of physics. Therefore I shall treat the "modeling of intelligent thinking entities" and "artificial intelligence" (or "simulation") as one and the same science. That we act as though we have two separate sciences is just another indication of how primitive our present science is. Think what geology or engineering would be like without any science of physics.

Our problem is not to simulate all the details of human thought, or to direct all the power of the computer's speed and size to specific peculiarly well-suited intellectual tasks, but to increase our understanding of the processes whereby a system comes to perceive, know, understand, manipulate, and interact fruitfully with its larger world. We then can particularize our general model into specific models, for example, of ants or humans. (But remember, this process is not as simple as it sounds. Even physics progresses fruitfully; applications often must be handled by cookbooks, not by theoretical analysis. So it is with our science.)

I think most workers in the field would reject this view, if only because it suggests that our problems are huge and that we have hardly begun to solve them. But we shall never solve our problems by evading them. We must take the proper path, and we may turn out to be luckier than we think. Perhaps we shall find the path to be short or discover that we have already come a longer way.

Some Prototypical Approaches to Perception

> Some forty years ago, in a society of physicists and physiologists, I proposed for discussion the question, why geometrically similar figures were optically similar. I remember quite well the attitude taken with regard to this question, which was accounted not only superfluous, but even ludicrous. Nevertheless, I am now as strongly convinced as I was then that this question involves the whole problem of form-vision. That a problem cannot be solved which is not recognized as such is clear. In this non-recognition, however, is manifested, in my opinion, that one-sided mathematico-physical direction of thought, which alone accounts for the opposition (Mach, 1906, p. 109).[2]

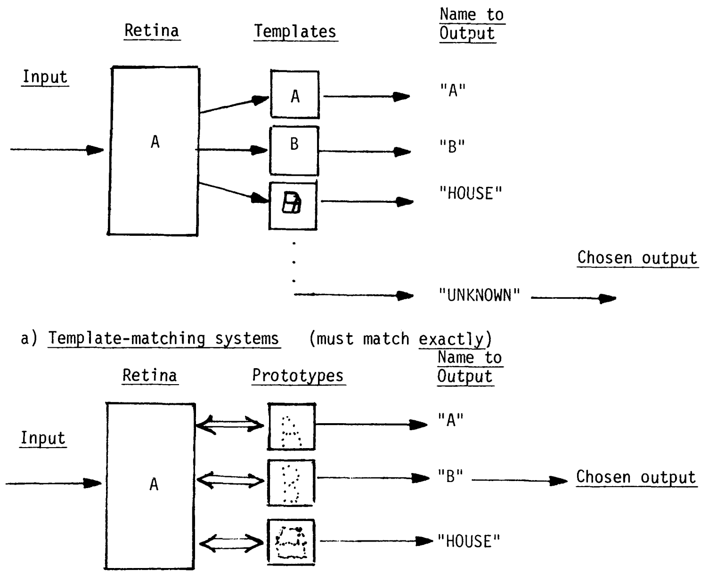
"Perception" refers to the gathering of relevant information: information that is usually (but not always) about objects, and objects that are important enough to have been ennobled with names. As William James suggests, "the consciousness of particular material things present to sense is nowadays called perception" (1950, Vol. 2, p. 76). Thus pattern recognition is probably the central purpose of perception. Often the object must be further described, especially to record its unusual qualities and its missing or aberrant parts. Often a whole scene—that is, the interrelations among several objects—must be described.

Let us briefly survey the approaches that have been taken to perception. The perception systems devised thus far have often had their greatest success in recognizing or describing particular types of objects—e.g. letters, or chromosomes, or polyhedra. (See Duda and Hart, 1973, and Uhr, 1973a, 1974 for recent surveys.) These systems have a common basic structure that may help illuminate what we mean by the terms "sense," "perceive," "sense-data," and "know." The various systems are described below.

(A) A *TEMPLATE* system (Figure 1a) stores a complete and detailed representation of each possible scene that might be input to it (e.g., Hannan, 1962). It matches each new input scene with stored "templates" until it finds an exact match, and then outputs the string of symbols (which might be a name like "B" or "TABLE" or "DACHSHUND" or "DOG" or "JOE") associated with the matching template.

(B) An *IMAGE* system (Figure 1b) stores a set of "typical" representations of possible scenes. These might be prototype templates, or they might be "probability contour maps" or other ways of describing "typical" or "prototypical" or "average" scenes (e.g., Baran & Estrin, 1960, Highleyman, 1962). The system computes the "similarity" between a new input scene and each of the stored "images." ("Similarity" is an obscure and complex concept and many measures have been used. One measure correlates the values stored in each cell of the probability contour map, stored for each possible image, with the values at corresponding cells in the input.) The system outputs the name associated with the image judged most similar to the input scene.

(C) Serial *DISCRIMINATION* nets (Figure 2b) apply a series of tests to the input (e.g., Unger, 1959; Naylor, 1971). Each test



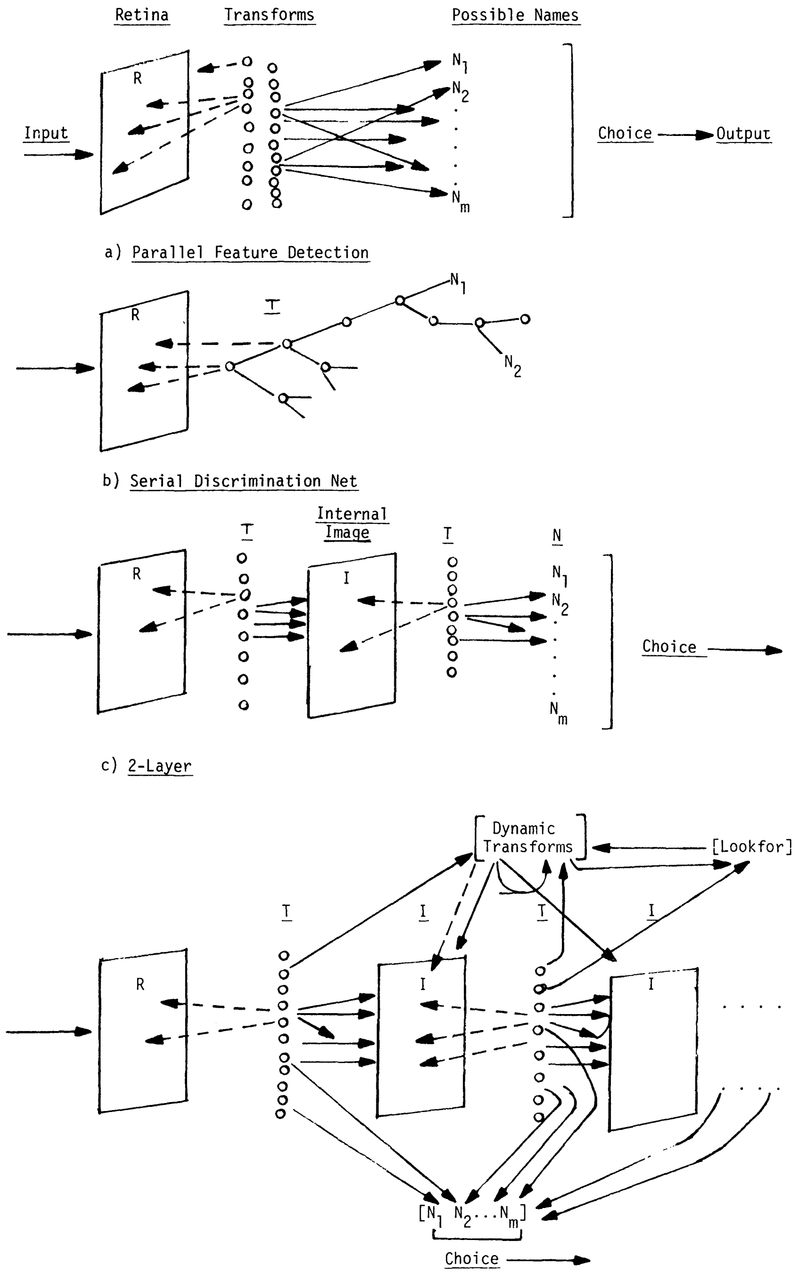
a) Template-matching systems (must match exactly)

b) Systems that measure "similarity" to stored prototype "images"

Figure 1. Perceivers that use templates and prototypes

determines which test to apply next, until the name to output is finally implied. This is the technique typically used for programs for "concept formation" (e.g., Hunt, 1962; Kochen, 1961; Towster, 1970). But it has had little success in pattern recognition, because it depends upon perfect tests that never make mistakes, whereas real-world patterns are so variable that their properties cannot be so easily captured. This weakness can be remedied by having a whole set of transforms make a probabilistic choice, or using a powerful algorithm, at each node. But that turns such a system into a structural perceiver, as discussed below.

(D) *FEATURE DETECTION* systems look for features and merge the possible name or names associated with each successfully found feature into a list of possibilities (Figure 2a) (e.g., Doyle, 1960; Munson, 1968). The most highly implied possibility is chosen and output. The system can use one, a few,



d) A Parallel-Serial Structural

System where Names as well as internal images can be implied at any layer, along with things to lookfor and dynamically implied transforms to apply, in glancing about.

Figure 2. Some structures of perception systems

or many features. Each feature can imply one, a few, or many possible names. Weights might, but need not, be associated with the implied names. A variety of functions have been explored to combine weights and choose among possibilities. Many different kinds of features have been used; e.g., edges, cavities, curves, angles, contours, loops, enclosures, area, center of gravity, dispersion. Different systems use different combinations of these features, chosen carefully to handle the types of patterns (e.g., printed letters, chromosomes, X-rays, aerial photos) the system must recognize.

(E) A feature can also be a whole set, a whole configuration, of features. And features can be embedded in a variety of larger structures, where the parallel structure of a feature-detector program is only the simplest. Thus *STRUCTURES* of features can be built, and looked for, either within each single transform, or by building larger structures of transforms, or both. I shall mention three roughly different kinds of structures.

(E.1) Serial *ALGORITHMIC* structures can be used (see Figure 3) by building a typical computer program that applies whatever functions, and makes whatever interspersed decisions, the programmer feels will best handle the problem (e.g., Brice & Fennema, 1970; Winston, 1975). Such a system has, roughly, the serial structure of a discrimination net, except that each node is a complex subroutine that embodies a complex set of tests. Often, hidden in this subroutine, will be a set of parallel processes.

(E.2) Sometimes processes are organized into large subroutines that are applied in *STAGES* (e.g., Reddy, Erman, Fennell & Neely, 1973). For example, speech recognizers may look for formats, phones, words, syntactic structures, and semantic interactions, in that order. Vision systems may similarly look for local edges, long strokes, angles, objects, and collections of objects.

(E.3) *CONFIGURATIONAL* characterizers can be used, where each looks for a whole set of features (e.g., Uhr & Vossler, 1963; Zobrist, 1971). And these can be structured into larger configurations, whether layered, hierarchical, heterarchical, or any other architecture (e.g., Uhr, 1972, 1973b, 1976b;

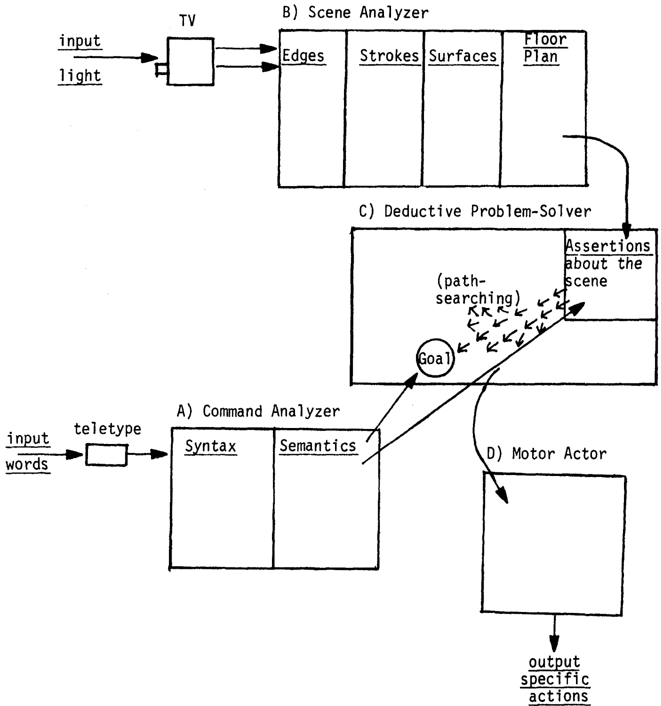


Figure 3. A typical robot system

Williams, 1975; Riseman & Hanson, 1974). Whereas the feature-detector builds one list of possibilities from which an overall choice is made, a configurational system builds many lists, both intermediate and final (see Figures 2c, 2d, and 4 for a few examples), and makes choices from each of these lists. And a wide variety of different inner- and outer-directed processes thus become possible, increasing the likelihood of power and success, and also the problems in finding good systems.

Simple templates are obviously impossible with real-world patterns that vary in unknown nonlinear ways. Far too many templates would be needed for all the variants, and the system would never be able to handle a slightly different new variant without being given a template for it. Image systems do not seem to work very well, but that may be because we do not have a very good grasp of how to compute "similarity" between complex patterns. I suspect that the right way is to break the pattern down into strokes and other features. But that turns the image system into a feature-detecting system.

The systems that use more complex structures probably do the best job. But the particular set of transforms used is probably the most important factor in the success of the system. It is not at all clear what types of structure are most powerful. Most researchers today seem to prefer deterministic algorithms, although I think that probabilistic configurations are more like the brain's network of neurons and give much more promise of flexibility and adaptability. Unfortunately, there are as yet no experimental comparisons to settle the issue. Inner-directed "glancing about" becomes increasingly important with complex scenes. Serial algorithms can be carefully pre-programmed to do this for specific small sets of known-in-advance patterns. But parallel-serial configurations should do better over a wide variety of unanticipated patterns, although very likely at the cost of occasional errors that could have been avoided if anticipated.

Wholistic Cognitive Systems: Piecemeal and Integrated

> Where the sign and what it suggests are both concretes which have been coupled together on previous occasions, the inference is common to both brutes and men, being really nothing more than association by contiguity. . . . Our "perceptions," or recognitions of what objects are before us, are inferences of this kind. . . .
>
> In reasoning we pick out essential qualities.
>
> Let us make this ability to deal with NOVEL data the technical differentia of reasoning. . . . Reasoning may then be very well defined as the substitution of parts and their implications for consequences or wholes (James, 1950, Vol. 2, pp. 326-30).

> This idea [central to Buddhist logic] is that our knowledge has two heterogeneous sources, Sensibility and Understanding. Sensibility is a direct reflex of reality. The understanding creates concepts which are but indirect reflexes of reality. Pure sensibility is only the very first moment of a fresh sensation, the moment x. In the measure in which this freshness fades away, the intellect begins to "understand." Understanding is judgment. Judgment is x = A where x is sensibility and A is understanding. Inference, or syllogism, is an extended

judgment, $x = A + A^1$. The x is now the subject of the minor premise. It continues to represent sensibility. The $A + A^1$ connection is the connection of the Reason with the Consequence. This reason . . . is divided in only two varieties, the reason of Identity and the reason of Causation (Stcherbatsky, 1962, p. 545).

I shall briefly describe some of the things that today's wholistic systems can do, with emphasis on the integrated SEER systems I have been developing.

TOWARD MODELING ORDINARY EVERYDAY INTELLIGENCE

Most of us human beings spend most of our time doing very ordinary things, things that, I suggest, are of the essence of intelligence. We constantly make complex decisions that take into account an enormous variety of relevant information. These decisions are designed to maximize our satisfactions, achieve a variety of goals, and avoid many anticipated and unanticipitaed dangers and pitfalls. But they are about ordinary and obvious things, and on the surface they may appear too mundane to glorify with the term "intelligence."

These include such things as deciding when, where, and what to eat; deciding how to get the food to our mouths (e.g., go to a restaurant, phone a friend, scream "ma-ma," or make it ourselves). We similarly decide what to do during a vacation, or during the evening; what to talk about with a friend; how to respond to a comment. Even ordinary perception uses a mixture of remembered and deduced information, as when we plan how to carve a turkey, or deal with a traffic cop.

These activities sound far simpler than proving a theorem, or playing a middling game of chess, or extracting cube roots. But they are far more difficult to program, and we are only beginning to get a grasp on the problem. We must get the computer to assess relevance, to take only small relevant subsets of large bodies of information into account, to decide in a flexible manner what type of process to effect next (given a continuing contextual interaction from the rich variety of pertinent information that the system attempts to gather as it surveys its external environment and its internal memory stores).

"ARTIFICIAL INTELLIGENCE" ROBOTS

Several robots have been programmed—notably at MIT (Winston, 1972), Stanford (Feldman, et al., 1971), and SRI (Nilsson, 1969)—to (a) input, parse, and "understand" a command teletyped in simple English, (b) sense and describe a room containing several cubes, wedges, platforms, and pyramids, as viewed by a television camera, (c) deduce how to carry out the command on the perceived objects, and (d) actually compose and effect the necessary motor actions. Each of these four major processes is handled by a separate (very large) program, with a minimum of information passed from one program to another.

Commands that have been successfully understood and carried out include the following (see Fikes & Nilsson, 1971):

ROBOT GATHER BOXES (the robot must deduce the sequence: MOVETO BOXA; PUSH BOXA BOXB; MOVETO BOXC; PUSH BOXC BOXB)

ROBOT TOUCH PYRAMID (the pyramid rests on a platform. If a wedge is pushed so its high edge is next to the platform, the robot can roll up the wedge and onto the platform, turn, and roll to the pyramid. Therefore the robot must deduce the sequence: MOVETO WEDGE; PUSH WEDGE PLATFORM: MOVETO LOC(I) (up the wedge and onto the platform); TURN 90^{o} RIGHT: MOVETO PYRAMID.)

A scene analysis program converts the television image into local edges, long straight edges, angles, contours, objects, and, finally, a floor plan of the recognized objects in the room. Some of this information is extracted and converted into a set of logical assertions that are passed on to the problem-solver, as grist for its deductive mill. The typed command is analyzed by a separate language "understanding" program, and used to describe the needed solution and start the problem-solver on its task of finding a solution-path.

Thus an extremely complex set of programs applies a pre-programmed set of stages to the various aspects of the problem. Information is first obtained from the perceived scene and the understood command and then passed to the problem-solver. Everything that the problem-solver might need to know must be anticipated in advance. The problem-solver does not call on the perceptual system to look for new objects that might serve in a conjectured solution.

The sequence of actions to effect a solution is worked out in advance, and then effected without any perceptual feedback.

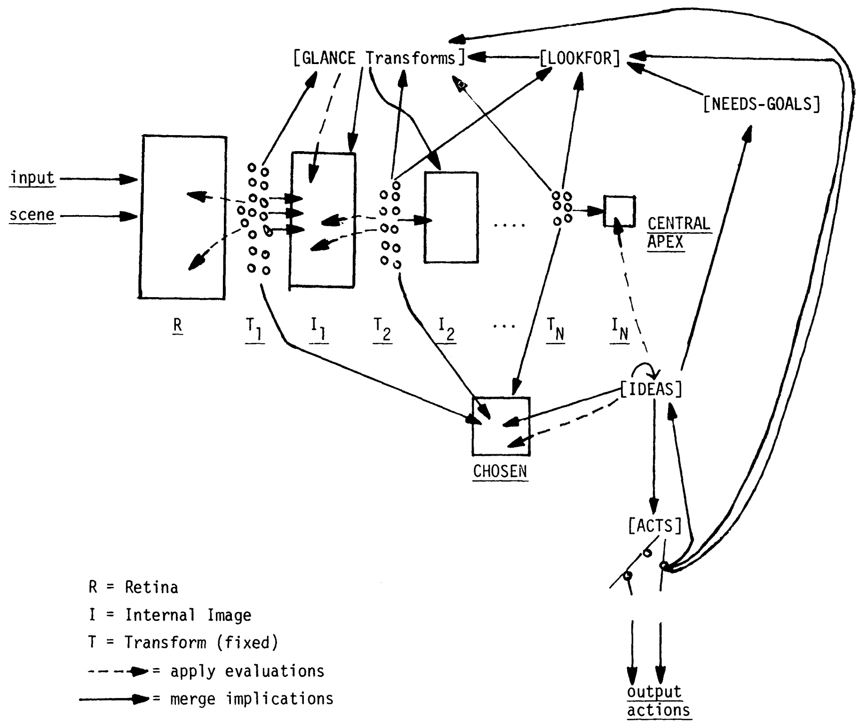
WHOLISTIC INTEGRATED "SEER" SYSTEMS

> According to the analysis of the psyche rendered by the Sāṅkhya, and taken for granted in the disciplines of Yoga, man is "active" (kartar) through the five "organs of action" and "receptive" (bhoktar) through the five "organs of perception." These two sets of five are the vehicles, respectively, of his spontaneity and receptivity. They are known as the "faculties working outward" (Bahyendriya) and function as so many gates and doors, while "intellect" (manas), "egoity" (ahankara), and "judgment" (buddhi) stand as the doorkeepers. The latter three, taken together, constitute the so-called "inner organ" (antahkarana); they are the powers that open and close the gates—inspecting, controlling, and registering whatever is carried through.
>
> The body is described as a town or kingly palace in which the king dwells inactive (according to the Oriental style) amidst the activities of his staff. For the human mind, with its contents and wisdom, is conditioned, in every specific case, by the peculiar balance of the gunas [activities] within the character and disposition of the given individual. His ideas, beliefs, and insights, and even the things that he sees around him, are, finally, but the functions or reflexes of his particular manner of not-knowing-better (Zimmer, 1956, p. 317).

In my own work I have been trying to develop simpler, more general and better integrated, wholistic systems. One of these—the SEER system—is described later; capital letters refer to processes or lists used. (See Uhr, 1975 a, b, 1976 a, b for fuller descriptions.)

Briefly, the SEER system looks for hierarchical Structures of Configurations (see Figure 4). It successively transforms a scene input to its "retina," extracting, abstracting, coalescing, compounding, until all implied possibilities are finally merged back into a single "CENTRAL" cell upon which more central cognitive transforms (called "IDEAS") continue to act. The perceptual subsystem has the overall architecture of a many-layered cone that gives a parallel-serial hierarchical structure to its set of transforms. Transforms are implied from within, by desired expectations, actions, and objects, as well as from without, by the external scene and relevant information extracted from it so far. These dynamically implied transforms are all merged into the GLANCE list, which is applied along with fixed transforms.

The fixed transforms reside at all locations at all levels of depth



The single general type of transform means that any process can call on, and imply information into, any other process.

Variant systems merge all choices into the CENTRAL APEX, others into the separate CHOSEN list. The IDEAS are applied to either, or both, as appropriate.

Figure 4. A wholistic integrated "SEER" system

of the cone (depth comes in layers, but this is only for convenience, and to model living systems—it need not). Each successful transform merges its implied images and names into the corresponding cell at the next layer, finally achieving the grand merge into the CENTRAL apex. Implied transforms to apply and things to look for are merged into the appropriate lists. Choices among possibilities are made in the CENTRAL apex; but choices can also be made in any of the other cells, since among the things that a transform can imply is the decision to choose, in a specified cell, among a specified class of possibilities.

Thus there are many nodes where things are looked for and processes are applied (we might consider calling these nodes of "awareness") and many cells at which decisions are made. (When would we want to consider these nodes and cells as loci for "consciousness" or "conscious choice?")

I am ignoring a number of other important features, and much detail, describing only what seems relevant to the issues examined in this chapter.

The cognitive transforms that look at the CENTRAL list are of two general types: (a) links and pointers into a network of information stored in memory (that is, for remembering), and (b) deductions, computations, and other stored information and/or processes of the sort we normally call deductive problem-solving.

The single most highly weighted transform in the IDEAS list is applied to the CENTRAL apex. If it succeeds (that is, enough of the tests it specifies are passed so that its threshold for success is reached), its implied associations and deductions are merged into CENTRAL, implied transforms to apply are merged into IDEAS and also into the list of GLANCE transforms to apply to the external scene, and implied acts are merged into an ACTS list. Thus perception, remembering, and deduction all intermingle, helping and calling upon one another. When an ACT is chosen (see Uhr, 1975a, b for details of how this is effected in a system that handles a static scene, and 1975d in an extended system that begins to interact with scenes of objects that move about and change over time), the system generates a specific sequence of actions needed to effect that act, binding each action with specific objects that have been perceived. This in turn calls for more perception and/or cognition, by implying further things to LOOKFOR, which in turn imply further GLANCE transforms to apply.

A number of variant systems have been coded, to explore the possible interactions among the various processes. One of these uses a list of HYPOTHESES to give more coherence and direction to the act. Now needs, goals, and expectations imply "hypotheses" that, if acted upon, might lead to their satisfaction. The system chooses the most highly implied hypothesis, which in turn implies acts upon types of objects, which the system looks for, or thinks about in order to arrive at alternative courses of action. The hypothesis also lists expected intermediate and final consequences, intermediate and final consequences, including feedback and need-satisfactions. These now serve to keep the act moving in the desired direction toward the goal, and to set the system to look for confirming or disconfirming evidence to use in assessing the hypothesis.

Perceptual transforms must also handle language. Real-world organisms have no special input channel, like the robots' teletype, for referential information. And that is inevitable, unless we establish artificially simplified relations between the knower and its environment. For signs, symbols, words, commands, suggestions, descriptions, and any kind of language are part of the single environment, and must be input through sensory channels, like eyes and ears, in mixed scenes of words, objects, and other things. Thus, for example, local edges will compound and grow into longer edges; then into several joined edges, or contours; then into larger wholes like a table top, window, tree, or (the letter) D; then into larger wholes like table, house, forest, (the word) DOG. Then DOG might imply such things as "animal" and "bark," and also some perceptual characterizers that will be applied to the successive images of the scene to try to find the object-dog that has been suggested by this recognized word-dog. If "LOOK FOR THE DOG" or "POINT TO THE DOG" is recognized as a larger structure over DOG and the other similarly recognized related words, then the perceptual characterizers needed to recognize and find a dog will be implied with very strong weights.

It would be hard—I suspect impossible—to draw a line in the SEER system where perception ends and remembering and ideation begin. For such a system compounds and associates to higher-level wholes and more abstract classes, qualities, and other concepts. Nor can we reasonably separate language processing from perceptual processing. It therefore seems crucial to use a single general type of transform to handle all aspects of the system's processes. Consequently, the same general transform type is used for memory associations, deductive processes, and also perceptual and linguistic processes.

The system must further decide what type of act to effect—whether to *name* a recognized object, *describe* the input scene, *answer* a question, *solve* a posed problem, *find* and *touch* a particular object, or *find*, *touch*, and *move* one object to another. These it does in very primitive and simple-minded ways. But the important point is that it must be able to choose among a variety of acts, in response to a variety of cognitive problems. It must itself decide which is the appropriate *type* of thing to do. It must also decide *when to decide* this, since there is no fixed sequence of steps to its processes. It must be able to use a mixture of its different types of processes in order to amass the relevant information, decide that it has done so, and therefore choose and effect the appropriate consequent actions.

Finally, these processes are not simply a function of the perceived scene of objects and commands. Rather, the system also has internal needs and goals which imply acts that might serve them and objects that might be needed for these acts, either as tools (e.g., a stick to reach), or as objects (e.g., the banana to eat). These needs and goals also imply classes of things, and particular things, to LOOKFOR, and characterizing transforms to GLANCE at the scene to effect that perceptual process. The transforms implied by needs and goals merge with transforms implied by transforms that have been implied so far, e.g., by partially recognized things. Thus the presses of needs, goals, and expectations from within and from prior percepts serve to direct processing, along with the presses of input from the outer world.

What Seems Necessary to Intelligent Systems?

In answering this question let us look first at perceptual systems, in which several basic structures are present in almost all programs, and then at the wholistic cognitive systems.

ASPECTS OF PERCEPTION

The raw image of the scene must be input to the system, to be stored as a digital iconlike representation in a first "retinal" input buffer. Then characterizing transforms must be applied. In the simplest whole-template systems, the first transform that succeeds will imply the chosen name. In simple parallel systems, characterizing feature-detectors will immediately imply possible names to assign to the input. In the simplest serial systems each transform will imply which transform to apply next. In more sophisticated parallel-serial systems each transform will imply several possible kinds of things, including internal images, new transforms to apply, possible names to choose, and triggers to choose.

In all cases we find a structure consisting of (a) the retinal input image, and (b) the chosen names (or set of names, or description). This structure suggests that we call the retinal input the "sensed image" and the final choice of name or description the "percept."

Almost any system that perceives with any degree of power also transforms the raw sensed input into successively more abstract internal images and makes one or more choices among alternative possibilities. We might want to equate the whole set of internal images (probably including the retinal image) with "sense-data." And the places where choices are made might be considered as very low-level and primitive loci of "awareness." Those loci where names or other elements of the descriptions are chosen might be considered loci where perception is achieved.

In a system that successively transforms the image by smoothing, filling gaps, enhancing edges, and building larger wholes, we have an example of the active constructive processes that tend toward object constancy. If the system further describes a recognized image by outputting elements of an internally stored description of an object-class, as well as a description of the characteristics actually perceived in this particular sensed instance, we begin to get the flavor of the higher-level kinds of construction found in human perception, where major distortions or missing parts are regularized or filled in during perception and thus not noticed by the perceiver.

"KNOWING," BELIEVING," AND "KNOWING THAT"

In a wholistic system the output of the perceiver is not a name or description printed by a teletype, but a piece of internal information used by other cognitive processes within the system to access or deduce related information. We might consider saying that the larger system "knows" what the perceiving subsystem tells it; or, if weights or probabilities are associated with the parts of the output of perception, that it "believes" with a certain degree of certainty.

There is some inclination to say that the perceiving subsystem itself "knows" what it perceives, to the extent that the term "know" carries connotations of an "ego" or "locus of consciousness" identical with the knower. One might for this reason want to equate "knowing" with the larger system into which the perceptual subsystem outputs. Thus in the SEER systems we might say that the IDEAS list of transforms, which looks at the CENTRAL list of chosen names and descriptive information, "knows."

Would we be inclined to say that the system "knows" if it rarely, or even never, outputs the correct answer to a question or the correct name for a pattern? Or would we ask for a high percentage of correct responses over a whole sequence of problems? Or would we insist that all responses be correct (but we would not do that with humans)?

These criteria would force us to say that any simple program that accesses any kind of data base "knows" the accessed information. For example, the system might access the addresses of anybody who filed an income tax return, doing it in the most stupid and "unknowing" ways, e.g., by matching the name in question with every name in memory until a match was found.

We might insist that the thing known be in some sense worth knowing, and that the knowing of it be a major achievement, and that the process of coming to know it be done in the "right way." These are extremely stringent requirements. And to traditional psychologists they will appear impossible to satisfy; for how can we look inside a living mind/brain to see whether it does things "in the right way?" But in fact this set of requirements is too easily satisfied, e.g., by programs for arithmetic. Such programs do indeed add, multiply, or extract roots in the right way, according to any of the procedures that we humans use (because we learned them in school). Arithmetic is worth knowing; the answers are indeed computed and generated, not stored; and a sequence of 10-digit multiplications is a major achievement.

These considerations seem to force us to accept a definition of (primitive) knowing that, like our definition of perceiving, will include many clearly unintelligent, uninteresting systems. And it forces us to consider higher-level aspects of the system, ones that we might want to insist must be present for more complex, powerful, and interesting kinds of knowing.

Interesting issues arise in trying to distinguish between "believing" and "knowing" in terms of the certainty of such systems choices. Suppose we stipulate that a system knows only if its choices are 100 percent certain, or certain beyond some unacceptable margin of error. Then we are forced to say that probabilistic systems believe rather than know—unless we find in or give to such systems routines that convert beliefs into certainty, in which case the system would, I think, be self-delusional in an important sense.

A second more straightforward alternative is to identify the knower with a special subsystem. This amounts to saying that the system does not know its answer, or know that its answer is right, unless the system contains some subject (ego) separate from the subsystem that actually achieved that answer, that accepts the answer as in its judgment right, certifying and outputting it. The difficulty with this second alternative is that the question-answerer or arithmetic unit has many more guarantees of the correctness of its answer than does the "ego" (which might merely be a program that prints out payroll checks) that later looks at the answer to check that it is right. For what could the ego do? It might itself check the answer—recomputing possibly in some different way; thus it would then be, or use, just another arithmetic unit. Or it might simply know that its arithmetic unit is correct in the way that a computer knows each of its processors is correct; then we might better call this "knowing" blind absolute trust.

This is not to say that "knowing that" or a "self that knows" are not valid constructs. I think in fact they are necessary, and we already begin to see them exemplified in primitive form in systems that are required to make higher-level choices and flexibly decide what kinds of processes to effect. Such systems are always deciding, in effect, that "more must be known or done about this or that." The places where such decisions are made—especially if they are few, or singular—begin to have some of the features of a "knowing self" or ego which "knows that things are the case."

But it is interestingly paradoxical that this supposedly "higher" type of ego that "knows that" is not identical with the processor that knows or achieves the knowledge. Such an ego inevitably makes fallible judgments, judgments it is in a poorer position to make than the "lower-level" routines (very likely under its control) that actually did the dirty work of knowing. The ego is, therefore, something of an "illusion," to borrow an idea from Indian philosophy. (Or is it the judgment ("buddhi") that lies beneath the western ego of the persona?)

When might we also say that a system "knows that it knows?" Possibly the use of an expectation to guide its perceptual search for feedback about an hypothesis upon which it has acted is the germ.

ONE "ENDURING EGO" AND "CENTRAL CONSCIOUSNESS?"

These terms suggest a kind of ego and consciousness that are not easily exemplified in information-processing systems. An easy way to introduce their germ is to put the system under the control of an executive who decides at the highest level what to do next, like the executive in a separated problem-solving system who uses its set of heuristics to decide what step to take next in searching for a solution-path. The executive thus has the same structure as any other node where lists of things are compared and choices are made, and it is hard to see why it should be singled out as "the central consciousness." Dennett (this volume) points out how difficult it is to examine one's consciousness and suggests, "we have access—conscious access—to the *results* of mental processes, but not to the processes themselves" (p. 217). I am inclined to call this consciousness of which we are aware "central consciousness," as opposed to the far larger network of unconscious consciousnesses that permeates the mind/brain with points of judgments and choices.

Thus this central consciousness is hard to pinpoint and may be relatively unimportant in the network as a whole. Most thinking may well go on in unconscious consciousnesses. But there are many simple examples of conscious consciousness, as when we try to remember a phone number, do mental arithmetic, memorize a person's name, recall a name, or recall and reconstruct the layout of furniture in a familiar room. In such cases we are conscious of a structured set of information and we make a conscious effort to process this information. Even though most of the attendant processes still go on unconsciously—so that we are not conscious of the details, much less the neuronal events, that lead to the recognition or the recalling of a name—a central consciousness seems to be present, and seems more than just the slot where the answer appears. It seems reasonable to think of all choice nodes as loci of local consciousness. But some of these are more central than others, and if we construct our system with a single executive we get an obvious candidate for central consciousness.

> The mind is at every stage a theatre of simultaneous possibilities. Consciousness consists in the comparison of these with each other, the selection of some, and the suppression of the rest by the reinforcing and inhibiting agency of attention. The highest and most elaborated mental products are filtered from the data chosen by the faculty next beneath, out of the mass offered by the faculty below that, which mass was sifted from a still larger amount of yet simpler material, and so on (James, 1950, Vol. 2, p. 288).

But I do not think this central consciousness will have much of the flavor of human consciousness without a good bit more—something that would give the system some subjective feeling (which could indeed have illusory aspects) that it is conscious, making a conscious effort, is consciously aware, is directing and choosing its processes consciously.

We tend to equate all these various aspects of consciousness with a central unified process, and one pretty much in control, as executive. But James's shifting unpredictable stream of consciousness is a better model. In this model the external press of the environment and the internal presses of more or less enduring needs, goals, plans, and expectations impinge upon and control all decisions. Although a system with a single executive may resemble the social structures we have found easiest to implement in business and political organizations, it is not necessarily the most efficient or productive. The succession of choices among many loci of consciousness may weave a unifying thread and thus replace the single executive. Nature may well evolve systems with several executives, or systems with no executives, in which different subsystems simply interact. In the parallel-serial probabilistic systems that nature seems to use, the fact that at some points choices are made may make it appear that the system has an executive. But there may well be only choice nodes, including the for each moment relatively final choices.

LOCI FOR "UNCONSCIOUS CONSCIOUSNESSES"

An obvious place to look for consciousness is at the point of heuristic serial choice and application of the transforms that have been merged into the IDEAS list to the CENTRAL store of implied and chosen things. At an even lower level, any locus of choice might be considered a local node of consciousness. Fairly global organization is given to a system when the chosen ACT controls what is to be done (whether to perceive, remember, or problem-solve; which problem to work on). Still more global organization is provided when a list of HYPOTHESES is used to handle not only information-processing but also the expectations that guide that processing—motor acts, perceptual consequences, and deductive processes, and also expectations of consequent feedback.

Thus we have several possible loci of possible conscious awareness: (1) at each transform's application; (2) at each and any locus of choice; (3) at the CENTRAL list, as looked at by the IDEAS list; (4) at the ACT that is being executed and therefore directs processing; and (5) at the HYPOTHESIS that controls all processes, binding them together by local expectations and using overall expectations to look forward to feedback. (6) Conscious awareness can also be attributed to a learning-controled system that has higher-level hypotheses to the effect "if this is learned, [I] will better be able to cope."

These loci are, of course, extremely conjectural (I think it is quite premature to take achieved systems seriously as in any way "conscious"). But they do posit precise and completely open lists and processes; and so it seems instructive to look at them in the larger context of the total program, see what roles they play, and begin to ask, in a precise way, what still needs doing, and how else consciousness could be modeled.

EGO, AND THE CONTINUITY OF SELF

> But to ask that the observer should imagine himself as standing upon the sun instead of the earth [to see earth revolve around sun] is a mere trifle in comparison with the demand that he should consider the Ego to be nothing at all, and to resolve it into a transitory connexion of changing elements. (Note: Cp. the standpoint of Hume and Lichtenberg. For thousands of years past Buddhism has been approaching this conception from the practical side. Cp. Paul Carus, *the Gospel of Buddha*) (Mach, 1906, p. 356).

Ego, like central consciousness, might also be illusory. Quite possibly a hypothesis controls and guides actions and expectations for a while (maybe seconds, maybe hours or days), and then, or even interspersed, other hypotheses take over. But all these hypotheses will be related, like a loose family, in that they were learned to serve the same set of needs and goals in relation to a common set of experiences with a common environment. And they are all chosen because implied by the same set of transforms. Might that not be enough to give the rather vague feeling of continuity that leads us all to the (illusory?) conclusion that we have a continuing ego? "The ego must be given up. . . . Ego-consciousness can be of many different degrees and composed of a multitude of chance memories" (Mach, 1906, pp. 24-25).

In some sense there seems to be an observer of the above process, a "self-awareness" that has at least some "free will" and some impact on what the whole system does—e.g., whether it plays chess, bakes cakes, or sings—and seems to be observing, even savoring, the whole process. But all we have in our programs is the locus where the choice of hypotheses is made, and the related loci where prior choices that implied those hypotheses were made. Is there then a need for some highest-level processor that can search about to put that kind of information together, much as a special routine may well be needed to compose an "appropriate" description from the mass of potentially interesting descriptive information that perception implies? Such a processor would give the system the illusion of coherence and self-control; but what productive functions would it serve?

FEELINGS AND PROTEIN

> If consciousness belongs to all protoplasm, by what mechanical constitution is this to be accounted for? . . . This question cannot be evaded or pooh-poohed. Protoplasm certainly does feel; and unless we are to accept a weak dualism, the property must be shown to arise from some peculiarity of the mechanical system. . . . It can never be explained, unless we admit that physical events are but degraded or undeveloped forms of psychical events (Peirce, 1971, vol. 6, pp. 172-73).

Primitive feelings, like hunger, pain, and redness, may well be the most troublesome problem of all. How can a program feel the sharp pinches of pain, or model the difference between a color-blind person, who always correctly understands that certain shades of for-him grey are actually red, and the person who feels the sensual power of the redness? We might argue that such primitive feelings are not necessary for intelligence; but it seems likely that their close relation with perception and with need-motivation will indeed make them necessary. In any case they exist, and it would seem strange to have a model of the mind/brain that could not encompass them.

To explain pain we might posit some complex overloading of information, some pervasive very high weight of implication from multiple sensory sources, both external and internal. Pain signals would thus be qualitatively stronger than and would override all other processes, and would seem overwhelmingly salient. Similarly, "redness" might be a rich welter of strong associations to a variety of sensory experiences. This suggests some complex of information, rather like that posited for consciousness and self. But I think the suggestion is even less satisfactory in the present context.

Alternately, we might say with Peirce that "protoplasm certainly does feel" (1931, vol. 6, p. 173), that these primitive feelings are simply qualities of protein. But what might that mean or entail? We could not have just protein sensors, for they would then have to send information to the nonprotein central computer, which would not be able to "feel." And a protein-based general-purpose computer would not have, at the level of its actual information-transforming processes, anything different from an electronic, mechanical, or optical computer, unless we made it a continuous analog computer.

SUMMARY

I think some perceptual processes are illuminated by today's programs. We can, if we wish, equate them with various parts of a recognizer. But it seems more useful to posit ever higher and more global levels of perceiving and knowing, rather than to insist that one function resides exactly here and another exactly there.

It may be premature to attribute awareness, feelings, consciousness, and self to any existing, or any conceivable, program. Until recently I had been inclined against such attributions at this early stage of our model building. But some suggestions for the various loci where choices are made seem intriguing. The most intriguing suggestion seems to me that these complex constructs may lie not at any single locus, but are subjective constructs from the complex set of processes that play back and forth over many nodes of awareness. This suggestion seems the most consonant with what psychology and neuroscience tell us about living mind/brains, and with intuition and introspection.

An alternate suggestion is that a single highest-level executive reflects and controls decisions. But even if such an executive did make all the decisions, what would give it an awareness and self-awareness of that, or a feeling of its own existence as a sentient ego?

"Iconic," "Short-Term," and "Long-Term" Memories

The first "retinal" input buffer seems a good candidate for iconic memory. It results from a transduction that minimally distorts the energy emitted from the scene to be perceived. It is as much a "picture" as a digital computer (or a living eye of discrete rods and cones) can receive and store. But it is important to note that, although the image is an icon of the actual scene, it has already been transformed into a discrete set of symbols. For that is all that a digital computer is capable of handling. Thus the "icon," along with any other possibly "analog" representations, is handled with a discretized approximation. Whether it is treated as an analog will now depend upon the procedures that transform and interpret it.

The internal buffers into which successful transforms merge their implications (including transformed and abstracted images) are all candidates for short-term memory (STM). All are a function of recently fired transforms. All must fade away relatively rapidly (unless reinforced because they continue to be salient) to make room for new information. Thus all the internal buffers at every layer of the perceptual cone—including the central apex and the lists of ideas to apply, things to look for, dynamic transforms to apply, and acts to effect—are short-term memories.

It may seem reasonable to think of perceptual memories as unconscious and to regard the central apex and ideas lists as what we ordinarily mean by short-term memory. But this would imply that we are subjectively aware of these stores and therefore give them special status. It would be more appropriate to call these stores STM that are also loci of attention. In any case it seems quite unreasonable to assume that there is only one STM involved in temporarily storing directions to a party, memorizing a date, or trying to remember a poem.

The perceptual transforms are an especially permanent part of long-term memory, rather than, as usually pictured, something rather different. They may differ from other parts of long-term memory by being more iconic or analogical (as is a feature-detector for an edge or a stroke, or a configurational characterizer of a set of parts composing a complex shape, such as a B, chair, or face). But they must be handled in the brain by the same substrate—a set of neurons. The use of a single general type of transform in SEER systems shows how we can at least begin to handle perceptual transforms in our programmed models.

Passive vs. Active; Outer- vs. Inner-Directed Perception

The typical pattern recognition system is frequently criticized as too passive, too determined by the sensed input. Minsky (1975) has proposed an inner-directed use of "frames" (which, he notes, are quite similar to Bartlett's (1932) "schemata"). Vernon (1952) and Neisser (1976) have pointed out that human perception is usually active; we glance about, walk about, and conduct experiments, in a continuing attempt to achieve perceptual understanding.

But I think this criticism is unfair, since even in the earliest programs, pattern recognizers used inner-directed processes and schemata. Thus MacKay's (1956) early suggestion of "analysis through synthesis," taken up by Eden and Halle (1961) among others, essentially said: "dynamically build a structure that comprehends the input." Indeed, I think one of the best expressions of this suggestion is found in Sāṅkhya psychology of about the third century B.C.:

> The foremost point of the thinking principle, when meeting objects through the senses, assumes their form. Because of this the process of perception is one of perpetual self-transformation. The mind-stuff is compared, therefore, to melted copper, which when poured into a crucible assumes its form precisely (Zimmer, 1956, p. 288).

The inner-directed application of the "frame" (or "schema" or "map") was used in much the way that Minsky suggests by Grimsdale, Sumner, Tunis, and Kilburn (1959) ("turn the scene into a graph; then look for stored graphs in that representation"). It was also used by Marill, et al. (1963) in a more extreme way, since their system's search for features was entirely inner-directed and under control of the internal description of the object.

A further step is taken when new things to look for and transforms to apply are implied by what has been found and implied so far, as in Uhr's "flexible" systems (1973a, ch. 8). Now the system can begin to glance about, to entertain and follow up hypotheses. The SEER systems have needs, acts, cognitive associations and deductions that imply things into these dynamic lists of things to look for and transforms to apply, and thus achieve a rich contextual interaction of internally and externally implied processes. This, I suggest, is what is needed—a mixture of inner- and outer-directed processes, dynamically changing over time as a function of what has been perceived, thought of, and tried out so far.

Thus cognitive processes play an intimate role in perception. And once we put implied actions into the loop—so that, for example, the system moves itself physically to the right of an object in order to see it better, or rub it, rather than simply applying new visual characterizers to the right—perception will exhibit the motor aspects that Neisser emphasizes. The problem, then, becomes one of selecting the sources (outer or/and inner) of information that imply the transforms to apply and the set of processes (including motor actions) involved in the continuing perceptual loop.

Perception and Complexity

Do today's programs for perception really "perceive" the letter "B" or the "CHAIR" or the "DACHSHUND" that they *correctly name?* They certainly *do something interesting, and rather difficult,* something that we did not know how to do until we started coding such computer programs. And they share this capacity only with the higher animals. I do not think it is terribly important whether we call this capacity "perception" or "recognition" or "identification" or "classification." (I must mention that I do not think definitions serve much purpose except to point and orient.)

It seems very hard to distinguish between programs that do and do not perceive. Does a template program perceive? Even when it has only five or two templates or even one template? Does a card reader perceive? (Remember, the brush that finds the hole in the card is just a template.) Does the riverbed perceive the flowing water or the river? I think that on the usual definition of perception—on the gathering of information about external environment—we must attribute perception to all these things. We might try to grade perception, e.g., "trivial," "real," "realer," etc. Or we might try to construct a set of dimensions that underlies perception.

We can profitably examine the complexity of the precept. We can consider the number of alternate possible names that might be assigned to an input, or the number of combinations of names when the input is a scene that might contain several objects. Another dimension of complexity is the number of possible variants of each object—e.g., all the different bananas, or faces, or expressions on your mother's face. This within-object variability might be quite simple (e.g., when the object is rigid and a linear transformation suffices) or terribly complex (e.g., when the object can be distorted as though made of stretching rubber, or of crumbling and stretching rubber, or of a growing and learning system of muscles).

The interactions between within-object variability and across-object variability are crucial and complicate things further. A jug vs. a knife may be a rather easy distinction to make, even though there is an enormous variety of each. But consider distinguishing higher mammal A vs. higher mammal B from pictures of their faces with many different expressions. If the mammals are from a different race or species from the perceiver, the distinction becomes even more difficult, since relevant features are not as easily learned. And what if the mammals are sisters or identical twins? We can extend dimensionality here to include such things as depth, time, color, texture. On the one hand this may increase the complexity of the problem (the moving object must be tracked, the several color components combined). On the other hand it may simplify, since more information is input (the several moments, or colors).

If we increase the repertoire of possibles we further complicate: 10 numbers; 26 letters, thousands of faces; many thousands of words and objects. Such numbers necessitate, I think, a structure for our perceivers that successively builds larger and larger wholes, thereby reducing a single unmanageable task to an interlocking set of smaller tasks.

Thus we get into issues of the size and difficulty of the problems handled and the internal structure of our system. Here we enter into the scientific enterprise of building and evaluating a theory, for our system can best be taken as a theory of the set of scenes of objects that might be set in its view. So we must apply the familiar canons of hypothetico-deductive method: parsimony, generality, power, fruitfulness, elegance. (This raises another unsolved problem: for we do not yet know how to assess these terribly complex programs for their generality, elegance, or power, and we rarely try to evaluate the size, much less complexity, of the problems they attempt to handle, or the degree of success they attain.)

"General-Purpose" Computers, Analogs, and Discrete Approximations

Today's "general-purpose digital computers" are just enormously large and fast embodiments of the "universal Turing machine," which is a very general logistic system *plus* the specification of the actual procedures (the "machine") needed to effect transformations in that system. Church (1936), Kleene (1952), and Turing (1936) independently developed the lambda-calculus, recursive functions and Turing computable functions. These were quickly proved to be equivalent (see Kleene, 1952, Arbib, 1969). They are generally regarded as defining the broadest concept of effective calculation that can be used in developing a firm foundation for mathematics (and for thinking in general in the view of Turing and many scientists involved in getting computers to "think"—see, for example, Minsky, 1967, pp. 108-11).

And, despite a great deal of effort since, nobody has been able to define any broader concept of calculation except in terms of continuous, analog systems (which can be approximated with discrete digital representations).

Any of today's computers can, when given the proper program, carry out any set of describable processes that could be carried out by *any other* computer, whether hardware or software (although if the computer is too small, or slow, or inappropriately designed for the particular program, it may need more time than could ever be made available to it). But everything input to a computer must be expressed in sets of discrete symbols (e.g., a language, but also a graph, or an array like a television raster). And any internal representation, whether we call it "picture," "icon," "image," "schema," "map," "model," "semantic net," or something else with an analog flavor, is always represented as a discrete approximation.

Using the computer, then, entails the assumption that the mind/brain can be described by structures of discrete symbols. This seems a reasonable working assumption. And the variety of complex functions computers can effect is the largest possible set of functions. Therefore they give the psychologist the broadest possible range of potential theories and the greatest possible expressive power as a vehicle for describing his or her theories. And running the program actually tests its consequences.

But it is by no means certain that the computer is adequate for the formidable task of intelligent thinking or, even more troubling, for more homely processes like feeling pain or seeing red. If we feel that these can be handled only by new hardware, as Gunderson (1971) seems to suggest for such "program-resistive" processes, then we are asking for something very different and difficult. We are asking for a broadening of the definition of a computer, since the meaning of "general-purpose" is precisely that any conceivable hardware-embodied computer can be described in a program so that it can be simulated exactly by any general-purpose computer.

If we feel that "analog" computers are needed, we must specify the analog. Today's "analog computers" typically use a set of simple circuits to integrate and differentiate. But to recognize objects or represent images we would need very different kinds of analogs, in tandem with the procedures that manipulate and transform them, and extract and compute information from them. Most research on perception, question-answering, and memory representation consists precisely in the search for and exploration of different, usually at least partially analog, possibilities. We usually find it convenient, and plausible, to use the digital computer to make discrete approximations. So whenever we talk about "analog" processes on today's computers we mean processes that are actually being realized digitally. If we insist that true continuity exists in the real physical world, and that it is vital for intelligence, we are again insisting upon the inadequacy of today's computers. Today (because we are still tooling up) we can still approximate; but tomorrow we may have to join Peirce and use analog configurations of protoplasm to capture the conscious feelings of continuous life.

DIGITAL VS. ANALOG; STORED VS. COMPUTED; DESCRIPTIONS VS. MODELS

We see, then, that it makes little or no difference in what sort of mind-stuff, in what quality of imagery, our thinking goes on. The only images *intrinsically* important are the halting-places, the substantive conclusions, provisional or final, of the thought. Throughout all the rest of the stream, the feelings of relation are everything, and the terms related almost naught. These feelings of relation, these psychic overtones, halos, suffusions, or fringes about the terms, may be the same in very different systems of imagery (James, 1892, p. 169).

Input to the computer inevitably gives a discrete approximation of the sensed scene. Perceptual transforms, at least at the earlier layers, almost certainly must be analogical, e.g., edge, stroke, and feature-detectors. Transforms can imply abstracted analogs and/or symbols, or strings of symbols including words in a natural language. When such symbol strings are arranged linearly, as in sentences, it is probably best to think of them as nonanalogical, nonpictoral, non-image-like descriptions or propositions. But if symbols are arranged in list-structures or in graphs and are connected by relational symbols (like "above" or "near-left" or "60^o") that are interpreted by procedures that compute these relations, I think it is best to say that these relations, and their resulting structure, and these procedures, and their resulting transformations, give an analogical aspect to the representation. But we cannot exclude even linear strings, because any arbitrary graph or n-dimensional array can be expressed in a linear string, using symbols like parentheses and commas to embed and indicate structure.

The analog flavor emerges, I think, when the representation has some model-like aspects; in the sense that not everything is spelled out and stored explicitly, but an enormous, even a potentially infinite, amount of information can be derived in a reasonably efficient way. For example, a model of the robot's room and its position in that room allows us to compute its location relative to any object, or any point, in that room. But it will be impossibly cumbersome to deduce these relative positions from a small set of logical propositions about the objects in the room.

Nor is a picture or icon an analog, in the sense of allowing for a dense set of useful derivations. The raw picture can be used only as a rigid template, to be matched exactly. A picture of a particular rectangle is a very bad analog concept of "rectangle," for it would be of no help with any other rectangle. Rather, we need an abstract structural description that allows all rectangles to be handled. This will—as in the analytic geometry representation—almost certainly be a mixture of symbols and relations that give an (analogical) structure from which an enormous set of possibilities can conveniently and powerfully be derived.

It seems reasonable to think of terms like "image," "schema," "map," and "model" as vague overlapping suggestions of the analog aspects of the representation. But the actual analog qualities reside not in whether a memory contains symbols or pictures, but in the structure of the representation, the richness and power of the procedures that can manipulate and transform it, and the resulting potential richness of the information that can be derived. Thus is makes eminent sense for Peirce to hold the most extreme conception of mind as a continuous analog and simultaneously develop the deepest of "semeiotic" theories of this mind as a discrete symbolic "existential graph": "the mind is a sign, developing according to the laws of inference" (1931, vol. 5, p. 188).

WHAT DOES REALITY BECOME WHEN COMPUTERS BEGIN TO THINK?

> Bodies do not produce sensations, but complexes of elements (complexes of sensations) make up bodies. If, to the physicist, bodies appear real, abiding existences, whilst "elements" are regarded merely as their evanescent, transitory appearance, the physicist forgets, in the assumption of such a view, that all bodies are but thought-symbols for complexes of elements (complexes of sensations) (Mach, 1906, p. 29).

A computer program to model a world that contains an intelligent organism would have to specify the things in that world, including the organisms that perceive them, and the rules of interaction between things and organisms. As with all computer programs, these specifications must be made in strings of symbols that define spaces, networks, graphs, or other structures.

If we can describe and model an intelligent organism with a computer program, the rest of the model, of its microcosm world, should be easier to construct—since it consists of merely a few additional intelligent organisms, plus other, much simpler, structures. Such a project will probably never succeed: either because no computer will be large enough to contain a sufficiently large piece of the world to present an interesting environment to the intelligent organism that forms a part of it or because the computer must be a continuous analog.

But it is instructive to examine this situation, from the point of view of the simulated organism's attempts to examine how it perceives, what it knows, and what is the reality out there. We know the reality because we have programmed it all, and can examine it all with detachment. Assume that we someday achieve intelligent programs that satisfy all our criteria for "thinking" and "intelligence," and that assert such things as "I know" and "that object is real." Then we shall have an existence proof of a perceived and constructed reality whose "ultimate stuff" *we* know to be merely a description, a construct of symbols. Our programs believe in this "reality" just as we believe in our reality, and we give them just as much credit for intelligence as we give each other. Such a constructed reality would then be the *only bona fide* reality whose structure we *knew*.

Once again we move toward Peirce:

> The universe is a vast representamen, a great symbol of God's purpose, working out its conclusions in living realities. Now every symbol must have, organically attached to it, its Indices of Reactions and its Icons of Qualities; and such part as these reactions and these qualities play in an argument that, they of course, play in the universe—that Universe being precisely an argument. In the little bit that you or I can make out of this huge demonstration, our perceptual judgments are the premises *for us* and these perceptual judgments have icons as their predicates, in which *icons* Qualities are immediately presented. But what is first for us is not first in nature (1931, vol. 5, pp. 75-76).

Summary and Some Troublesome Comments

This paper explores aspects of perception and cognition of special interest to philosophers and to psychologists by examining the relevant structures and processes of actually programmed models.

The many programs for perception are briefly described and categorized, to elicit several common structures that help give precise meaning to the concepts of "sensation," "percept," and "sense-data."

Wholistic cognitive systems are examined by a brief look at today's robots and a more detailed description of the author's attempt to develop better-integrated, contextually interactive, probabilistic SEER systems that perform a variety of intellectual tasks in a flexible way that intermingles the separate cognitive processes.

Perception is examined from an empirical point of view, by studying programs that give increasing power and generality over various dimensions of complexity. Cognitive systems are then examined for the light that their necessary structures and processes throw on such basic concepts as "perceiving," "thinking," "knowing," "understanding," "ego," and "self." It is suggested that systems are needed whose behavior is a function of both the external environment's presses and internal needs, goals, and ideas.

Possible loci of "iconic," "short-term," and "long-term" memories are posited. It is suggested that the interrelated issues of "analog" vs. "symbolic" representations, "computed" vs. "stored" information, and "models" vs. "descriptions" boil down to a question of the degree to which relevant but unanticipated information can conveniently be derived from the internal representation.

Mathematical logicians have been successful in defining a very broad class of "effective procedures" that Turing machines (and that means today's "general-purpose" computers) can embody: such machines will effect any procedure, no matter how complex, that is described by a "program" input to them. This development shows that computers are the most general vehicle for information-processing man has been able to devise. Therefore, writing a computer program is the most general, powerful, and convenient way to model the mind/brain.

But it is not easy to judge how difficult is this time-honored task or how far we have come. Nor do we have any firm results or guidelines relating to the type of program needed—whether it be given predigested knowledge or learn; describe discursively or model analogically; have the flavor of deductive problem-solving or inductive perception; be parallel, serial, or parallel-serial; be deterministic or probabilistic. These are issues of efficiency, power, and generality, to be settled by the hypothetico-deductive method, as we test, evaluate, and rebuild our models—something we have not yet learned to do.

I have tried to argue that certain structures and processes are necessary for thinking. But are they sufficient? Or are some of the essential characters (like the feelings of redness, pain, and love), or even the protagonist (like the conscious self), still missing? We may not even know enough to send out a casting notice.

I have suggested that a unified consciousness, the ego, and (somehow in a more troubling way) feelings might be semi-illusions built up from the complex structure of experience. But how can a system know its external and internal reality? One insight comes from the ancient Indian philosophies of Sāṅkhya, the Yoga-sūtras of Patanjali (see Zimmer, 1956) and Buddhist logic (see Stcherbatsky, 1962). Perception, thinking, ego, and personality reflect the superficial surface agitations of everyday life. Beneath these lies the still, aware self, at one with all being, all consciousness.

Could we attain this deeper level of awareness in our programs? Does this suggest some new kinds of organization with a reflective meta-level that does not in any way control or direct, but rather assesses dispassionately if it is allowed to? Does it suggest that we must program in, must start with, a pervasive universal consciousness? Or is Peirce correct in saying that "matter is effete mind, inveterate habits becoming physical laws" (1931, vol. 6, p. 30)? Computers, whatever their material embodiment, are part of the universal mind-stuff. So we should be able to get them to embody conscious minds that are aware with feelings and thoughts, possibly simply by freeing them to be continuous, analog, probabilistic, and imprecise (protoplasm?). Indeed they already offer us examples of a perceiver that knows and apprehends external objects.

Notes

1. For *SEE* and to *ER*r is human; *SE*mantic learn*ER*; *S*ensed *E*nvironment *E*ncoder, *R*ecognizer and *R*esponder; and, above all, for short.

2. The liberal quotes from James, Mach, and Peirce in this paper are attempts to show that today's programmed models, the problems they attack, and the issues they raise are the children of (a) a central tradition of western psychology that was sidetracked by attempts to push oversimple associationism too far (e.g., the Mills, Pavlov); (b) weird and wondrous flights into introspection turned fantasy (e.g., Wundt, Freud); and (c) a know-nothing behaviorist reaction (e.g., Watson, Skinner) that misapplies the scientific method,

pragmatism, and physicalism and has shamed psychology into ignoring its most central questions. This is an oversimplification (I am not a historian of philosophy or psychology). But I want to emphasize how congenial are the thoughts of a tradition that has its beginnings in Indian as well as Greek philosophy, and how liberating.

Computer programs allow us to return to a precise and sensible examination of complex internal processes. How refreshing, and contemporary, is talk of the stream of consciousness as the processes of judgment and choice, of reality as consisting in constructions over relations, of existential graphs as (the medium for) thought.

References

Arbib, M. A. *Theories of abstract automata*. Englewood-Cliffs, N.J.: Prentice-Hall, 1969.

Baran, P., & Estrin, G. An adaptive character reader. *IRE WESCON Convention Record*, 1960, 4 (pt. 4), 29-36.

Bartlett, F. C. *Remembering. A study in experimental and social psychology*. Cambridge: Cambridge University Press, 1932.

Brice, C. R., & Fennema, C. L. Scene analysis using regions. *Artificial Intelligence*, 1970, 1, 205-226.

Church, A. An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 1936, 58, 345-363.

Dennett, D. C. Toward a cognitive theory of consciousness (this volume).

Doyle, W. Recognition of sloppy, hand-printed characters. *Proceedings of the IFIP WJCC*, 1960, 17, 133-142.

Duda, R. O., & Hart, P. E. *Pattern classification and scene analysis*. New York: Wiley, 1973.

Eden, M., & Halle, M. The characterization of cursive writing. In Colin Cherry (Ed.), *4th London Symposium on Information Theory*. Washington: Butterworth, 1961, 287-299.

Feldman, J. A., et al. The Stanford hand-eye project. *Proceedings of the 2nd International Joint Conference on Artificial Intelligence*, 1971, 521-526.

Fikes, R. E., & Nilsson, N. J. STRIPS: a new approach to the application of theorem proving to problem solving. *Proceedings of the 2nd International Joint Conference on Artificial Intelligence*, 1971, 608-620.

Grimsdale, R. L., Sumner, F. H., Tunis, C. J., & Kilburn, T. A system for the automatic recognition of patterns. *Proceedings of the Institute of Electrical Engineers*, 1959, 106 (pt. B), 210-221.

Gunderson, K. *Mentality and machines*. New York: Anchor, 1971.

Hannan, W. J. The RCA Multi-font reading machine. In G. L. Fischer, Jr., et al. (Eds.), *Optical character recognition*. Washington: Spartan, 1962.

Highleyman, W. H. Linear decision functions with application to pattern recognition. *Proceedings of the IRE*, 1962, 50, 1501-1514.

Hunt, E. B. *Concept formation. An information processing approach*. New York: Wiley, 1962.

James, W. *Psychology (briefer course)*. New York: Holt, 1892.

James, W. *The principles of psychology*. New York: Dover, 1950 (2 vols.) (orig. 1890).

Kleene, S. C. *Introduction to metamathematics*. Princeton: Van Nostrand, 1952.

Kochen, M. An experimental program for the selection of disjunctive hypotheses. *Proceedings of the Western Joint Computer Conference*, 1961, 19, 571-578.

Mach, E. *The analysis of sensations*. New York: Dover, 1959 (orig. 1906).

MacKay, D. M. The epistemological problem for automata. In C. E. Shannon and J. McCarthy (Eds.), *Automata studies*. Princeton: Princeton University Press, 1956.

Marill, T., et al. Cyclops-I: A second-generation recognition system. *AFIPS FJCC Conference Proceedings*, 1963, 24, 27-34.

Minsky, M. *Computation: Finite and infinite machines*. Englewood-Cliffs, N.J.: Prentice-Hall, 1967.

Minsky, M. A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision*. New York: McGraw-Hill, 1975.

Munson, J. H. Experiments in the recognition of hand-printed text. *Proceedings of the Spring Joint Computer Conference*, 1968, 33, 279-290.

Naylor, W. C. Machine imitation. Unpublished Ph.D. thesis, University of Wisconsin, 1971.

Neisser, U. *Cognition and reality*. San Francisco: W. H. Freeman and Co., 1976.

Neisser, U. Perceiving, anticipating, and imagining (this volume).

Nilsson, N. J. A mobile automaton: An application of A. I. techniques. *Proceedings of the 1st International Joint Conference on Artifical Intelligence*, 1969, 509-520.

Peirce, C. S. *Collected papers*. Cambridge, Mass.: Harvard University Press, 1931-1958 (8 vols.) (orig. 1859-1908).

Reddy, D. R., Erman, L. D., Fennell, R. D., & Neely, R. B. The hearsay speech understanding system. *Proceedings of the 3rd International Conference on Artificial Intelligence*, 1973, 185-193.

Riseman, E. M., & Hanson, A. R. Design of a semantically directed vision processor. *Technical Report 74C-1*, Computer Science Department, University of Massachusetts, January 1974.

Stcherbatsky, T. *Buddhist logic*. New York: Dover, 1962.

Towster, E. Studies in concept formation. Unpublished Ph.D. thesis, University of Wisconsin, 1970.

Turing, A. M. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* (ser. 2), 1936, 42, 230-265.

Uhr, L. Layered "recognition cone" networks that preprocess, classify and describe. *IEEE Transactions Computers*, 1972, 21, 758-768.

Uhr, L. *Pattern recognition, learning and thought*. Englewood-Cliffs, N.J.: Prentice-Hall, 1973(a).

Uhr, L. Describing, using "recognition cones." *Proceedings of the 1st International Joint Conference on Pattern Recognition*, 1973(b).

Uhr, L. Form perception and scene description: Toward a theoretical and experimental science of complex structures. *Computer Science Department Technical Report 231*, 1974.

Uhr, L. Toward integrated cognitive systems, which *must* make fuzzy decisions about fuzzy problems. In L. Zadeh et al. (Eds.), *Fuzzy sets*. New York: Academic Press, 1975(a).

Uhr, L. A wholistic cognitive system (SEER-2) for integrated perception, action and thought. *Proceedings of the 3rd MSACC Conference*, 1975(b).

Uhr, L. A wholistic integrated cognitive system (SEER-T2) that interacts with its environment over time. *Proceedings of the 4th MSACC Conference*, 1976(a).

Uhr, L. "Recognition cones" that perceive and describe scenes that move and change over time. *Proceedings of the 3rd International Joint Conference on Pattern Recognition*, 1976(b).

Uhr, L., & Vossler, C. A pattern recognition program that generates, evaluates and adjusts its own operators. *Proceedings of the Western Joint Computers Conference*, 1961, 19, 555-570. (Reprinted, with additional results, in E. Feigenbaum and J. Feldman (Eds.), *Computers and thought*. New York: McGraw-Hill, 1963.)

Unger, S. H. Pattern detection and recognition. *Proceedings of the IRE*, 1959, 47, 1737-1752.

Vernon, M. D. *A further study of visual perception*. Cambridge: Cambridge University Press, 1952.

Williams, H. A net-structure for learning and describing patterns in terms of subpatterns. *Proceedings of the 3rd MSACC Conference*, 1975, 3, 99-108.

Winston, P. H. The MIT robot. In B. Melzer and D. Michie (Eds.), *Machine intelligence* 7. Edinburgh: Edinburgh University Press, 1972.

Winston, P. H. (Ed.), *The psychology of computer vision*. New York: McGraw-Hill, 1975.

Zimmer, H. *Philosophies of India*. New York: Meridian, 1956.

Zobrist, A. L. The organization of extracted features for pattern recognition. *Pattern Recognition*, 1971, 3, 23-30.