GRANT
IN-61-CR

THE ORGANIZATION OF AN AUTONOMOUS LEARNING SYSTEM

Pentti Kanerva

Research Institute for Advanced Computer Science
NASA Ames Research Center
Mail Stop 230-5
Moffett Field, CA 94035

RIACS Technical Report 88.14

May 1988

N89-13976

(NASA-CR-184556) THE ORGANIZATION OF AN AUTONOMOUS LEARNING SYSTEM (NASA) 24 p CSCL 09B

Unclas
G3/61 0175198

Research Institute for Advanced Computer Science

THE ORGANIZATION OF AN AUTONOMOUS LEARNING SYSTEM

Pentti Kanerva

Research Institute for Advanced Computer Science
Mail Stop 230-5
NASA Ames Research Center
Moffett Field, CA 94035

RIACS Technical Report TR-88.14
May 1988

ABSTRACT

This paper looks into the overall organization of systems that learn from experience, human beings and animals being prime examples of such systems. How is their information processing organized? They build an internal model of the world and base their actions on the model. The model is dynamic and predictive, and it includes the system's own actions and their effects.

In my modeling of such systems, a large pattern of features represents a moment of the system's experience. Some of the features are provided by the system's senses, some control the system's motors, and the rest have no immediate external significance. A sequence of such patterns then represents the system's experience over time. By storing such sequences appropriately in memory, the system builds a world model based on experience.

In addition to the essential function of memory, fundamental roles are played by a sensory system that makes raw information about the world suitable for memory storage and by a motor system that affects the world. The relation of sensory and motor systems to the memory is discussed, together with how favorable actions can be learned and unfavorable actions can be avoided. Results in classical learning theory are explained in terms of the model, more advanced forms of learning are discussed, and the relevance of the model to the frame problem of robotics is examined.

Work reported herein was supported in part by Cooperative Agreement NCC 2-408 between the National Aeronautics and Space Administration (NASA) and the Universities Space Research Association (USRA) and in part by a gift from the System Development Foundation to the Center for the Study of Language and Information (CSLI), Stanford University. To appear as the final chapter in my book Sparse Distributed Memory

THE ORGANIZATION OF AN AUTONOMOUS LEARNING SYSTEM

Pentti Kanerva

This paper is about systems that function independently, that interact with their environments and record their interactions, and that therefore have the potential for learning and adaptation. How do such systems work?

In trying to answer this question, we are guided by examples from nature. We can look at animals and ask what kind of internal organization sustains their autonomous, adaptive behavior. Specifically, if the system has a sparse distributed memory for recording its past, what besides the memory does it need, and what is the overall organization of the system like?

Memory for Patterns and Pattern Sequences

Let us first review the memory and see what functions it can sustain. Then the other necessary functions must be accomplished by other parts of the system.

The sparse distributed memory (Kanerva, 1984) works with long vectors of bits. These vectors can be thought of as patterns of binary features. The mathematics generalizes readily to patterns of multivalued features, the most important thing being that the number of features be large. From here on we assume that the features need not be binary. What we have, then, is a memory that can be addressed by large patterns of multivalued features and that can store these very same patterns.

Because a pattern can be used both as an address and as a datum, a sequence of patterns can be stored as a pointer chain. The first pattern in the sequence is used as the address in storing the second pattern, the second as the address in storing the third, and so forth. Any pattern in the sequence can then serve as a retrieval cue that will initiate the retrieval of the rest of the sequence.

Addressing the memory need not be exact. A previously stored pattern can be retrieved not only with the pattern's original storage address but also with addresses similar to it. In general, the address patterns that have been used as write addresses attract, meaning that reading within the critical distance of such an address retrieves a pattern that is closer to the written pattern, on the average, than the read address is to the write address.

This attractor property is fundamental to pattern recognition and sequence recall. To use the memory for recognizing a set of patterns, each pattern is stored with the pattern itself as the address; to use it for recalling sequences of patterns, each sequence is stored as a pointer chain. Reading from the memory is the same in either case: The pattern just read is used as the next read address. Since write addresses attract, the initial read address need not be exact. If it is well within the critical distance of some previous write address, 3-6 iterations will usually suffice to read patterns exactly as written. In other words, successive reading brings us closer and closer to, and actually finds, a stored pattern or sequence.

The memory groups patterns automatically, providing for two kinds of generalization or abstraction. One kind is the attraction by stored patterns and sequences: To read from the memory, we need not know the exact address patterns that were used in writing into the memory. The other kind is when many similar patterns have been used as write addresses. Then the individual patterns written with those addresses cannot be recovered exactly. What is recovered, instead, is a statistical average of the patterns written in that neighborhood of addresses. This generalization is in terms of the features that make up the patterns. The features that are common to all or most of the patterns in the neighborhood will stand out as an encoding for a cluster of patterns.

For example, the memory might be used for recognizing visual patterns. An object viewed from slightly different angles and distances will then produce a set of similar patterns. This being a pattern-recognition task, each pattern is stored with itself as the address. Consequently, many similar addresses will be used in writing into the memory, and they will select many common locations. Reading at any of these write addresses or at nearby addresses is then unlikely to yield a stored pattern exactly. Instead, the memory will produce patterns representing the object in an abstract sense rather than patterns representing any specific views of it. Some features of these aggregate patterns will be prominent; others will be unimportant. Mathematically, the object occupies a region of the pattern space with poorly defined boundaries.

The predictive power of the memory is based on its ability to retrieve sequences and to generalize. If a system's past is represented as a sequence of patterns and if this sequence has been stored in memory, the pattern representing the present moment can be used as an address to retrieve the consequences of similar moments in the past.

Modeling the World

As we—intelligent beings in general—interact with the world, we become better and better at dealing with the world. We say that we learn from experience. Our experiences are stored so that we can predict what is likely to happen and to choose appropriate action, for example, to avoid danger or to seek reward. Many things appear to be learned by nothing more than repeated exposure to them.

We can think of learning as model building. We build an internal model of the world and then operate with the model. What can we say about this model on the basis of how we behave and how our behavior changes with experience?

1. The modeling is so basic to our nature that we are hardly aware of it. It might even be said that this modeling is our way of understanding the world. We understand what is happening only to the extent that we are able to predict what is going to happen, and the internal model is our means of predicting. Again, we are mostly unaware that any predicting is even going on; we just do it because of the way we are built.

2. The modeling mechanism constructs objects and individuals. A person, a tree, a river are constantly changing, and our views of them are different at different times, yet we perceive them as "that person," or "that tree" (or "that species of tree"), or "that river."

3. Operating with the model is a little like operating with a scale model. Not only does the model have individuals and objects; it also mimics their actions and interactions. The more experience we have had, the more faithfully are the dynamics of the world reproduced by the model. This manifests itself in our habitual formation of expectations. For example, having experienced lightning followed by thunder many times, we come to expect thunder whenever we see a bright flash of lightning. Psychological experiments on classical (Pavlovian) conditioning show that proper juxtaposition in time is all that is needed for such expectations to form. The model simply captures statistical regularities of the world, as mediated by the senses, and is able to reproduce them later.

4. Our world model includes ourselves as a part. For example, we can prepare ourselves for a situation by imagining ourselves in the situation. When we do that, we get an idea of how we are likely to feel or act in the situation.

5. There is oneness to our subjective experience, whether that experience is dominated by the outside world or by our internal model of it. In normal, day-by-day life we are constantly in touch with the outside world through our senses. For us, the world is the way our senses report it to be. When we build our internal model of the world, the report of the senses is all that we have to go by. If the recording is faithful, the model can recreate a subjective experience that has been created by the world.

Ordinarily there is sufficient difference between the quality of the experience produced by the world (as mediated by the senses) and that produced by the internal model to let us keep the two apart. For example, we are quite confident that there is water in the pool when we look down from a diving board and see water. On the other hand, even if we can imagine ourselves flying by merely spreading our arms, we are not likely to jump off a cliff. Thus, we tend to recognize some experiences as real and others as imagined. This, however, is not always the case, as dreams and hallucinations illustrate. They are produced almost entirely by the internal model, but to us they can be very real, capable of producing physical signs of pleasure or fear, for example. In extreme cases we may be unable to tell whether the thing actually happened or whether we just "made it up." The point here is that the (subjective) experience produced by the world is of the same quality as that produced by the internal model of the world; there is no fundamental difference between the two from the subject's point of view.

6. Our internal and external "pictures" merge without our being aware of it. We scan our surroundings for overall cues and fill in much of the detail from the internal model. However, when something unusual happens, we begin to pay attention. We are alerted by the discrepancy between the external report of what is happening and the internal report of what should be happening on the basis of past experience.

Driving along a thoroughly familiar road is a good example. We know its turns and intersections so well that we hardly pay attention to details that usually stay unchanged; we rely on the internal model for such details. When some detail changes, as when a new stop sign appears overnight, it is the regular travelers who are the more likely ones to run it on their first few trips past the spot. They usually become aware of the new sign just after running it, and they experience startle.

7. The internal model affects our perception profoundly, again without our being aware of it. This is demonstrated by eyewitness accounts of crimes and accidents, particularly when the witness is prejudiced toward one of the parties involved (Loftus, 1979). The <u>pre</u>judgments are the product of the internal model. In general, perception involves the relating of the present sensory input to past input, which requires memory.

<u>Storing the World Model in Sparse Distributed Memory</u>

If intelligent behavior is based on modeling, what are the modeling mechanisms? I will postulate that memory stores and maintains the model and allows its use. Therefore, the memory must store a record of the system's past in a way that allows the system to predict what is about to happen, to plan action, and to act according to a plan.

For the purposes of the following discussion, let us say that at any given moment the individual is in some subjective mental state. A flow of these states, represented here by a sequence of states, then describes the individual's (subjective) experience over time. The world itself can likewise be described by a sequence of states, but the state space for the world is immense in comparison with that for an individual's experience.

I have emphasized above that a person's experience is influenced strongly by the world as reported by the senses, and that it can be influenced equally by the internal model--by what is retrieved from memory. The simplest way to build the world model, then, is to store the report of the senses in memory and to retrieve it later from there. If it is retrieved faithfully and allowed to feed into the subjective experience in the same way as the senses feed into it, there is no way for the individual to distinguish an experience created by the internal model from one created by the outside world.

To store the world model in a sparse distributed memory, we need to represent an individual's sensory information at a moment as a long vector of features and let a sequence of such vectors represent the passage of time. The memory works well with such sequences, and above all else it stores and recalls them naturally.

We can now begin to look at the overall organization of a system that models the world and that maintains the model in a sparse distributed memory. Since information supplied by the senses and information supplied by the memory can produce the same subjective experience, it is reasonable to assume that some common part of the architecture is responsible for the system's subjective experience about the world, and that both the senses and the memory feed into it. I will call this part of the architecture the system's focus. The system's subjective experience about the world over time is then represented by a sequence of patterns in the focus. By storing this sequence in memory, the memory can later recreate it in the focus. Figure 1 shows the relation of the senses and the memory to the focus.

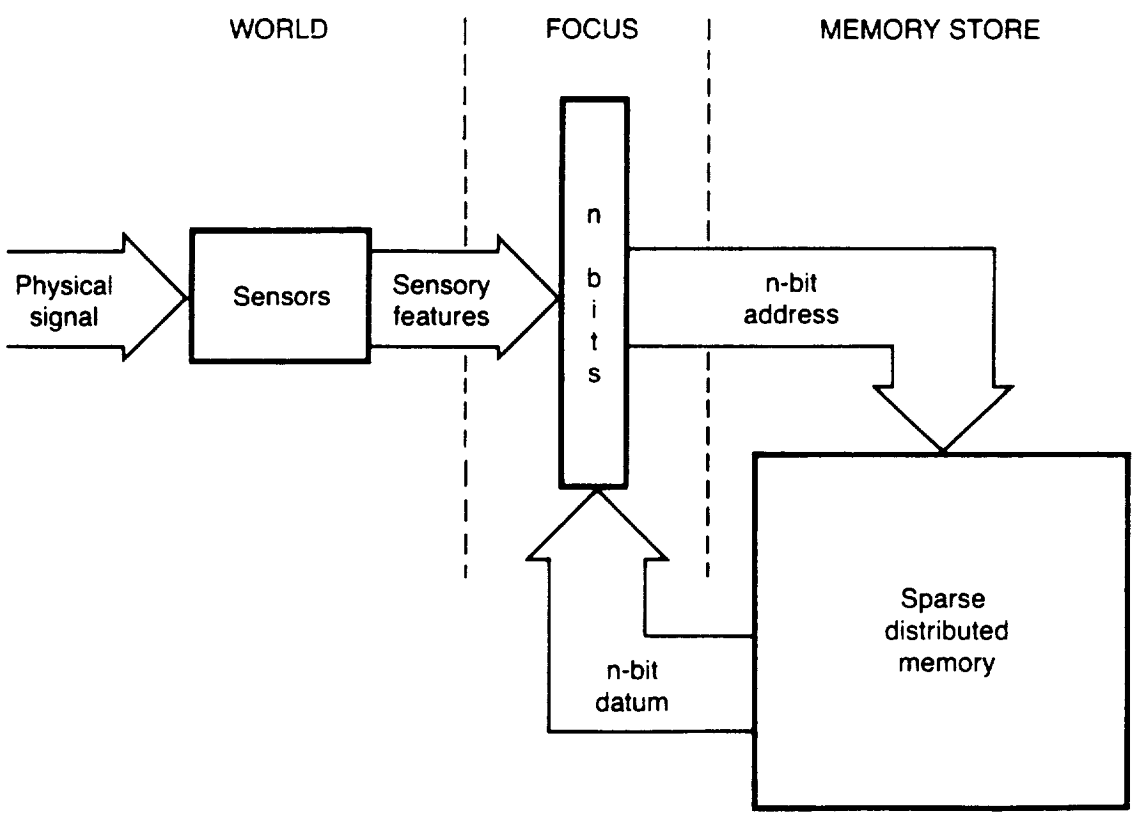

FIGURE 1. Senses, memory, and focus.

Because sequences are stored as pointer chains, the patterns of a sequence are used both as addresses and as data. In computer terms, the focus is a combined address-datum register, meaning that the memory is addressed by the focus, the contents of the focus are written into the memory, and the data from the memory feed into the focus. Thus, when the present resembles the past, the senses create a sequence in the focus that resembles a stored sequence. When this sensory sequence is used to address the memory, the memory responds with what the consequences have been in the past. Comparing those past consequences against what happens this time gives the system a criterion for updating its world model.

The world model is updated by writing into the memory as follows. The pattern held in the focus at time t is used to address the memory, activating a set of memory locations. The response read from those locations is the memory's prediction of the sensory input at time $t + 1$. If the prediction agrees with the sensory input, there is no need to adjust the memory; the read pattern simply becomes the contents of the focus at time $t + 1$. If the two disagree, however, a third, "correct" pattern is computed from them, and it becomes the contents of the focus at time $t + 1$; however, before it is used to address the memory (at time $t + 1$), it is written in the locations from which the "erroneous" output was just read (i.e., in the locations selected at time t). In the simplest case, this third (correct) pattern is just the sensory input at time $t + 1$.

In a more sophisticated updating of the world model, the memory is modified by writing error-correction patterns into it. The corrections for individual pattern components are based on the sum pattern in addition to the final, thresholded output pattern. If the output pattern is in error, the sum pattern can be used to find out by how much each bit counter in the selected locations has to be corrected for the final output to be right. The components of the correction pattern will then not be binary but will range over a larger set of values. As the correction patterns are written in memory over time, the memory builds a better and better model of the world, constrained only by the senses' ability to discriminate and the memory's capacity to store information.

Including Action in the World Model

So far we have seen how an autonomous learning system (an individual) can build an internal model of the world from the report of the senses. Besides observing the world and learning about it, the system also acts and learns from its interaction with the world. To act, the system needs motors (effectors); to learn, it must model its own actions.

The above discussion of a system's internal model of the world postulated the need for something like the focus and that the system's private, subjective experience is based on the contents of the focus. In trying to decide how to include the system's actions in its world model, let us start with the most public aspect of the system's operation, its observable actions.

The observable actions of humans and animals result from the contraction and relaxation of selected muscles. The muscles are controlled by neural signals that originate mostly in the brain, where the signals can be regarded as sequences of patterns over time, akin to the sensory signals. Learning to perform actions then means learning to reproduce sequences of patterns that drive the muscles. This suggests that the system's own actions can be included in the world model by storing motor sequences in memory in addition to sensory sequences. Since the way in and out of the memory is through the focus, the system's motors should be driven from the focus, and since the system's subjective experience is based on the information in the focus, deliberate action becomes part of the system's subjective experience without the need for additional mechanisms. This is fundamentally important to my theory of autonomous learning systems.

The organization of such a system is shown in Figure 2. A simple, idealized way to think about it is to assume that some components of the focus (well over 50 percent of them) correspond to and can be controlled by the system's sensors, and others (say, 10-20 percent) drive the system's motors, in addition to which the focus could have components with no immediate external significance. Naturally, all components of the focus can also be controlled by the memory. Retrieving well-behaved sequences from the memory to the motor part of the focus would then cause the corresponding actions to be executed by the system.

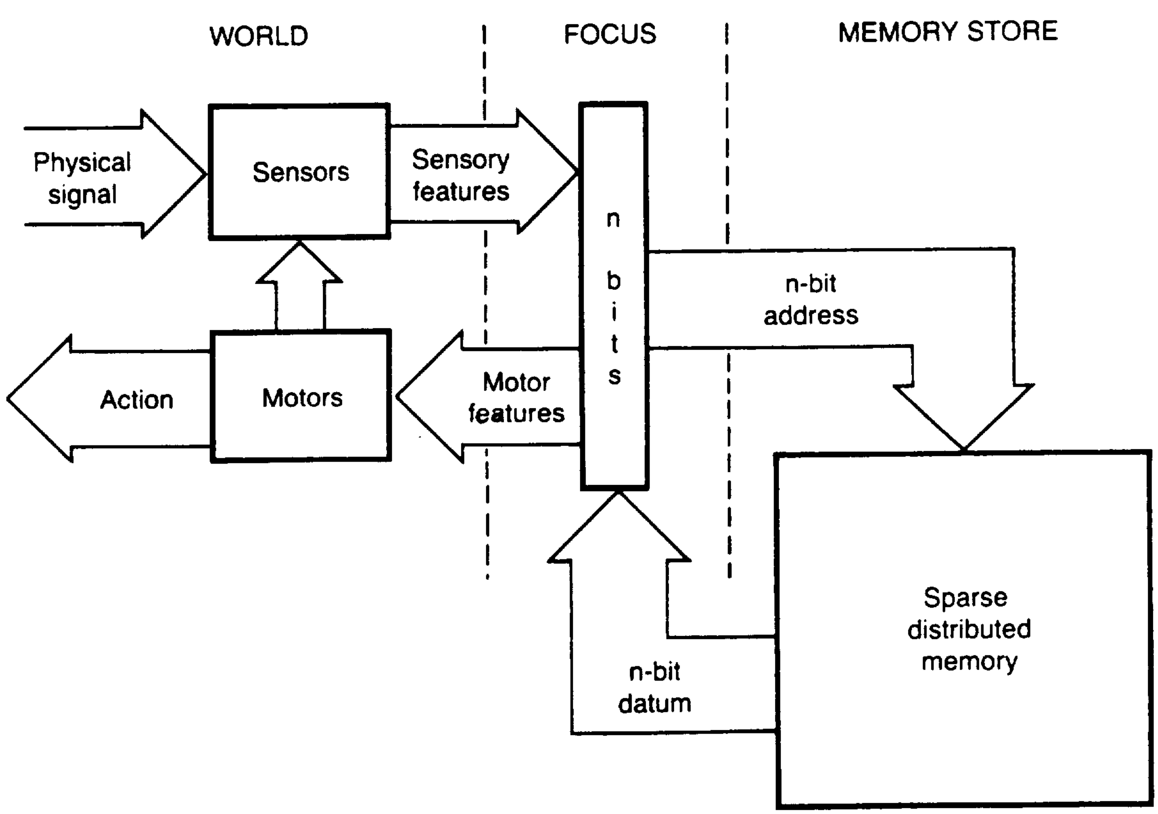

FIGURE 2. Organization of an autonomous system.

This organization makes it easy to describe simple forms of cued behavior. Let us assume that the stimulus sequence ⟨A,B,C⟩ is to elicit the response sequence ⟨X,Y,Z⟩, with A triggering X after one time step and with the two sequences running in lockstep from then on. The pattern sequence that needs to be generated in the focus can then be written as ⟨Aw,BX,CY,dZ⟩, where the first letter in each pair corresponds to the sensory-input section and the second to the motor-output section of the focus, and where the lower-case letters w and d stand for parts of patterns unspecified by the problem statement. A sensory input can be thought of as occupying 80 percent of the components of the focus, and a motor output as occupying the remaining 20 percent. Assume that the sequence ⟨AW,BX,CY,DZ⟩ has been written in memory (the previously unspecified w and d have specific values W and D in the sequence that has been stored), and that A is present (that is, presented to the focus through the senses). Then Aw, which is similar to AW, will be used as an address, and therefore BX is likely to be retrieved from the memory into the focus. This means that the action caused by X (action X, for short) will be performed at the time at which B is expected to be observed. If the sensory report agrees with B, then BX will be used as the next memory address and CY will be retrieved, causing the action Y. If at that time the report of the senses agrees with C, then CY will be used to read DZ, which completes the execution of the action sequence ⟨X,Y,Z⟩.

This example raises several questions: (1) Will the sequence be recalled and the actions performed every time the stimulus A is present? (2) Will the action sequence always be completed once it has started? (3) How might a system be trained for the sequence? The mathematical properties of the memory provide the following answers:

1. If stimulus A controls more than 80 percent of the focus (the critical distance in the examples of Chapter 8 is 209 bits out of 1,000), then presenting A will initiate a sequence of reads that tracks the stored sequence, no matter what the unspecified part w of the initial pattern Aw is. However, if A controls significantly less than 80 percent of the focus, or if the cue is not exactly A but a similar pattern A′, then Aw or A′w may not be sufficiently close to the original write address AW to cause BX (or something close to BX) to be retrieved. To read BX, it is then important that the action part w be similar to W.

By equating intentions and subjective states of receptiveness with actions, we come to a rather interesting interpretation of the above: Sometimes a system will respond properly to a cue only if it is waiting for the cue. For an example, assume that the action W means that the system is paying attention and is waiting for a cue, and that w means that the system is performing some other action. If A or A′ is then presented, the memory will be addressed with AW or A′W, and BX will be retrieved (see the preceding paragraph), whereas Aw or A′w could be too far from AW to cause BX to be retrieved. This means that the system's response to a cue depends on its state at the time the cue is presented. Other cues may be needed to get it in that desired, receptive state. The state might be described as the system's willingness to cooperate.

2. The second question concerns the completion of an action sequence. In terms of our example, the sequence ⟨AW,BX,CY,DZ⟩ has been written in memory, and BX has been read successfully from memory. If input from the senses is now suppressed, the focus will be controlled entirely by the memory, and the rest of the sequence will be recalled and the action completed.

Let us assume, however, that the senses are not blocked off, and that they feed the sequence ⟨A,B,K,L⟩ into the focus instead of the expected ⟨A,B,C,D⟩, where K and L are quite different from C and D. Then BX will retrieve CY, meaning that Y is executed and C is expected to be sensed. But since the senses report K, the next contents of the focus will be not CY but HY (where H is some combination of C and K that, in general, is quite different from C). Consequently, HY is too far from CY for anything like DZ to be retrieved, and this causes the last action, Z, to fail.

This failure can be interpreted in several ways. The simplest is to think of the system as monitoring its environment and ceasing to act when the proper cues are no longer present. We might say then that the response is driven by the stimulus, or that the action is maintained by the environment. The interesting thing is that the action can affect the environment. We can think of the system as monitoring the effects of its own actions, and that when the effects no longer confirm the system's expectations (e.g., when K is observed when C is expected) the action stops, whereas the system could have completed it--however inappropriately--had it not been monitoring its environment.

This example demonstrates how the system's own actions and their effects can be a part of the system's internal model of the world. As the system acts, and since the action is a part of the pattern that addresses the memory, the pattern retrieved from the memory includes an expectation of the action's results--that is, what usually happened on previous occasions right after the action was performed. The world model, or memory, can then be used not only to monitor the course of actions but also to plan action. To plan, the system must initiate the "thought" in the focus and then block off the present (that is, ignore environmental cues and suppress the execution of actions). The memory will then retrieve into the focus the likely consequences of the contemplated actions.

In this section the use of the words 'stimulus' and 'response' may seem strange to someone accustomed to the literature of psychology, where they are defined from a point of view external to an organism, the stimulus being presented to the sensory system and the response being mediated by the motor system. From the point of view of the memory, however, the entire pattern in the focus, including both sensory (stimulus) and motor (response) components, is one big stimulus, and the memory responds with a pattern that likewise contains both sensory and motor components.

3. The third question is about the learning of sequences of actions, which is essential if a system is to be adaptive. It is discussed in the following section.

Learning to Act

A system's model of the world is built from sequences of patterns in the focus, and the model's goodness is judged by how well it predicts such sequences. When the model predicts incorrectly, it is adjusted.

Regarding sensory experience, the world feeds correct sequences into the focus through the senses, so that the world decides whether a sensory prediction coming from memory is correct. If it is not (that is, if the memory's prediction disagrees with the report of the senses), then the memory is adjusted toward the report of the senses.

Regarding action, the picture is more complicated because no external source is feeding correct action sequences into the focus. The action sequences have to be generated internally, they have to be evaluated as to their desirability, and they have to be stored in memory in a way that makes desirable actions likely to be carried out in the future and undesirable ones likely to be avoided. In advanced learning of actions, nearly correct sequences of actions are fed into the focus from memory by recall of actions of role models. This corresponds to one person's adopting another person's speech patterns, mannerisms, facial expressions, ways of walking, and so forth; it will be discussed below in the section on social learning.

Initial conditions for learning. How does a human being or an animal decide whether a sequence of actions is desirable or undesirable? For some things critical to survival the answer is simple: We are born with preferences and dislikes and with instinctive ways to act; they are built in. For example, the preference for a proper blood-sugar level and body temperature need not be learned, nor does the dislike of hot or cold or of excessive pressure on the skin. To be more exact, they need not be learned by the individual; the learning has been done by the species in millions of years of evolution and is now passed on to the individual as a part of its genetic endowment. Likewise, animals have automatic reflexes, such as the sucking reflex of infant mammals. Given that there are such desirable and undesirable (subjective) states, we can define desirable and undesirable action sequences according to the states to which they lead.

To relate such built-in preferences to our model, we require that *some states (i.e., patterns in the focus) are inherently good and others are inherently bad*, with most states being indifferent. Rational action then means that the system will choose actions that lead toward good states and away from bad ones, and learning to act means that the system will store in memory sequences of actions in a way that increases the likelihood of finding good states and of avoiding bad ones.

Another condition for learning has already been mentioned, namely, that the system must *generate action sequences on its own* and store them in memory for later use. These constitute material for selection. To choose favorable ones among the sequences, the system must also *evaluate action sequences*. In what follows I will discuss several ways of finding good action sequences and of avoiding bad ones.

Let me start by expressing the learning problem mathematically. The system's (subjective) state at a particular time is given by the pattern in the system's focus at that time. The system has a (scalar) preference function defined on its subjective states; it is a function on patterns. The good and bad states occupy regions of the pattern space, with the good regions corresponding to high (positive) values and relative maxima of the preference function and the bad regions corresponding to low (negative) values and relative minima of that function. Were the system able to move in the state space, it would seek the relative maxima.

The indifferent states can acquire value according to whether they are found on paths to desirable or undesirable states. Learning to act can then be looked at as assigning preferences to states that start out as indifferent states. Formally, the built-in preference function maps patterns in the focus to (scalar) preference values. For most patterns the value of the function is near zero (meaning indifferent), and learning means assigning positive and negative values to more and more indifferent patterns. Learning to act then means extending positive and negative preferences to patterns with action components in a way that increases the likelihood of actions leading to desirable states and decreases the likelihood of actions leading to undesirable states.

Learning by trial and error. As was stated above, to learn to act the system must generate action sequences on its own. That is, it must generate patterns in the part of the focus that controls the system's motors. The initial generation of actions could be random, corresponding to the thrashing about of infants. Whatever follows these actions, including the effects of the actions, is then fed into the focus by the system's senses. In this way, action-effect pairs (or, more precisely, sensation-action pairs) will appear in the focus, from which they can be stored in memory.

A very simple way to learn is to observe the present situation, generate a random action, observe the resulting situation, and record it all in memory. As a consequence, the memory builds a model of the world that includes also the effects of the system's own actions. The memory can then be used to predict consequences of proposed actions--that is, to plan. Planning would proceed as follows. If the present situation resembles strongly a past one, the system can propose an action by whatever means it has (e.g., by recall from memory or by random generation). The situation and the proposed action together are then used to recall a resulting situation in the past (see the preceding section). If that situation had an action associated with it in the past, further iterations can be made to plan further into the future. If such iterations result in a favorable situation as determined by the system's preference function, the system has a reason to proceed with the proposed action; if it results in an unfavorable situation, it should try another action. We are assuming that in planning of this kind the system can block off external input after accepting the initial input (the present situation) and that it can suspend the execution of actions until it has accepted some proposed action.

A learning scheme of this kind is reasonable if the repertoire of situations and actions (the system's state space) is small and simple, or if the proportion of desirable states among all possible states is large. Under such circumstances, favorable actions could be found with reasonable speed. The systems that are of interest here, however, have very large and complex state spaces with relatively few desirable states, and consequently this learning method is much too slow to be of practical interest. The situation is familiar from artificial intelligence: Systems based on simple searching cannot cope rapidly with complex situations.

The efficiency of searching and learning can be improved considerably if good paths are remembered and are used later to find inherently good states. The method corresponds to backtracking search, and it works as follows: If the effects of a (possibly random) sequence of actions are good in a situation--that is, if an action sequence leads to a desirable pattern in the system's focus--the sequence leading to that pattern is considered to be good. To make use of that discovery later, the positive preference is extended backward, with decreasing intensity, to the patterns leading to the desirable one, and the sequence of patterns (or situation-action pairs) is written in memory. A positive value of the preference function then comes to mean either that the present pattern is inherently good or that a path from the present pattern to an inherently good one--a sequence of actions--has been found and stored in memory. Extending the preference thus improves the system's ability to detect sequences that are likely to lead to a good outcome. Similar ideas are found in Holland's work on classifier systems, in which credit for a good outcome is apportioned among active classifiers according to a "bucket brigade" algorithm, increasing the probability of a good outcome in the future (Holland, 1986; Holland et al., 1986).

Likewise, negative preference can be extended backward to patterns leading to an undesirable pattern. In addition, the undesirable sequences themselves can be stored in memory--although they need not, because, by definition, it is not important for the system to find states that are inherently bad. If the sequences are not stored, the system will still be able to avoid undesirable states, but it will not be able to retrace the steps to an inherently bad state and hence to determine the reason for avoiding a particular state.

Realizing the preference function. The mechanism for storing patterns in a sparse distributed memory can also be used for storing the preference function, which is a scalar function on patterns: The value of the function can be stored the way a pattern component is. Thus, each memory location would have a counter for the preference function. If the address of the memory location is a favorable pattern, the counter will be positive; if it is an unfavorable pattern, the counter will be negative; and if it is indifferent or as yet undefined, the counter will be close to zero.

In reading from the memory, the counters for the preference function can be pooled in the same way as are the counters for a pattern component, and their sum tells whether the pattern in the focus is favorable (sum greater than zero) or unfavorable (sum less than zero).

Having a built-in preference function then means that the function counters of some locations are nonzero from the start and that such nonzero counters may even be unmodifiable. Extending the preference means taking the present value of the function (especially if it is strongly positive or negative), reducing it toward zero, and writing it into the function counters of locations activated by the most recent patterns of the sequence, together with the writing of the sequence itself.

Speed of learning. The learning methods described so far are basic, in that they allow a system to learn even if it is left alone. However, these methods are slow, and therefore behavior based on such learning is not rich and complicated. This does not mean that animals growing up in isolation cannot have complicated behaviors, only that any such behavior they do have is prewired or preprogrammed genetically. Rigidity is typical of such behavior; the behavior is automatic. A standard stimulus elicits a standard response, no matter how inappropriate to the particular situation it may be (e.g., in experimental situations that imitate nature in some significant ways but differ from it in others).

Learning in Social Settings

Let us take a cursory look at learning theory, to see how some well-known results could be accounted for by the memory model. The common thread is that an individual learns from a trainer or a role model.

In competition for survival, fast learning is advantageous. Learning from others speeds up learning dramatically and makes possible the learning of complex behaviors (which can be quite arbitrary). This is most evident in human learning. Knowing how to swim is useful in almost any society, but it is unlikely that most of us would learn without a teacher or an example. Language is a learned skill that is very complex and in many ways arbitrary. Different languages involve very different vocabularies, different ways of making sounds, different grammars, and different systems of writing, and yet they perform very similar communication tasks. The behavior survives by being learned, practiced, and taught.

Classical conditioning. In classical conditioning (also called Pavlovian learning or supervised learning), an artificial (new) stimulus is substituted for a natural (old) one. The natural or old stimulus is one for which the subject already has a (natural or old) response, and the artificial or new stimulus is one for which the subject has no response. By training, the subject learns to give the old response to the new stimulus. The training goes as follows: The trainer presents a new stimulus (e.g., a bell) followed by an old stimulus (food). The subject responds (salivation). After sufficient repetition, the new stimulus alone will elicit the old response; the old response has become associated with the new stimulus. However, if the two stimuli are presented in the other order, old before new (food before bell), there is no learning; the new stimulus will not elicit the old response.

To relate this to the memory model, notice that the old or natural stimulus is meaningful to the subject at the start of the experiment but the new or artificial is not. In terms of the model, this means that the system's preference has been established for patterns representing the old stimulus but not for patterns representing the new stimulus. Thus, when the subject receives the old stimulus, it also encounters a nonzero value of the preference function. This tells the memory to store the sequence of patterns leading to the old stimulus and to extend the preference to those patterns. If the new stimulus precedes the old stimulus repeatedly, the sequence leading from the new stimulus to the old response becomes established in memory and the preference function likewise becomes established for that sequence. Consequently, the new stimulus alone will produce the old response, and it can even take the place of the old stimulus in training another artificial stimulus for the old response.

What if, instead, the new stimulus is presented after the old stimulus but before the old response? Will the old response become associated with the new stimulus? Psychological experiments have shown that, as a rule, it does not. An explanation, based on the memory model, would be as follows: The association from the old stimulus to the old response has already been formed, and the preference has been extended to the old stimulus, so that there is no sudden change in preference when the old response is given, and thus no learning is initiated by this mechanism.

In summary: Classical conditioning makes use of reward and punishment—that is, things that are inherently good or bad or that have in the past been associated with good or bad things—to teach the subject specific behavior patterns, which can be quite arbitrary. Possible memory mechanisms at work here are the recording of meaningful experiences and the extending of preferences to previously indifferent states.

<u>Learning by imitation</u>. The most complex forms of learned behavior, such as the use of language, are acquired largely by imitating other individuals. What learning mechanisms might be at work there?

So far I have proposed two occasions for a system to learn: when an unexpected event takes place, and when a meaningful event takes place. An event is <u>unexpected</u> if the memory provides a clear prediction for it but the prediction is incorrect. The memory record is then considered to be at fault and is modified so as to improve prediction under similar conditions in the future. The occasion of learning is thus the same as in the failure-driven memory of Schank (1982). An event is <u>meaningful</u> if it fetches a strongly positive or negative value of the system's preference function. The sequence leading to the meaningful event is then stored in memory, and the preference is extended back to the last few patterns of the sequence.

The two occasions for learning can be combined into one if success and failure in predicting are meaningful in themselves. Let us assume that when the memory makes a good prediction the system experiences it as a positive value of the preference function, and that when it makes a bad prediction the system experiences it as a negative value. In both cases the just-preceding sequence of events is written (again) in memory and the preference is extended back to the sequence. A system of this kind will learn from mistakes but will take time to build confidence in what it has thus learned; in general, it prefers and tends toward predictable things.

It seems that an internal reward mechanism of this kind is necessary if a system is to learn by imitation. In addition, a second internal ingredient seems to be necessary: The system must use itself to model the behavior of other systems. Successful modeling is then experienced as a positive thing. This may sound like a fancy way of saying that to learn by imitation one must like imitating, but there is more to it when we relate it to the memory model. First, the system must store an image of the behavior of others; second, it must map this image onto actions of its own; third, it must observe the results of its own actions and compare them against its image of the behavior of others (that is, the system must <u>identify</u> with the role model). Because of its internal reward mechanism, the system works to perfect the match between its own behavior and that of the role model.

It is my assumption that such internal mechanisms, including internal reward and punishment, are behind learning by imitation, which, in turn, is primarily responsible for complicated social learning (external reward and punishment being only secondarily responsible). Through social learning, groups of individuals can develop and maintain behavior patterns that have very little to do with an individual's survival in an indifferent environment. In fact, a group's behavior can produce a new environment that is maintained by the behavior. Diverse civilizations and cultures provide numerous examples of this. They are based largely on the models people have and the modeling they do in their heads. The study of such modeling is a major research task and will not be undertaken here.

Application to the Frame Problem of Robotics

The organization of an autonomous system discussed in this paper has been motivated by observations about the organization of information processing in animals. It should therefore help us think of how to build robots, and it should shed light on outstanding problems in robotics. The <u>frame problem</u> is one such problem that has been discussed widely in the artificial-intelligence community (Pylyshyn, 1987). It deals with the updating of a robot's internal model of the world as the robot interacts with the world--that is, how a robot can keep a tally of the side effects of contemplated actions. The following example illustrates the problem.

A robot lives in a world. To function there, the robot maintains an internal model of the world--a data base. In the data base are represented objects of the world (e.g., the robot, a cart, a telephone, room 1, room 2), properties of the objects (e.g., all rooms are stationary, the cart is movable, the telephone is blue), and relations between objects (e.g., the cart is in room 1, the telephone is on the cart, the telephone's receiver is on the hook). To allow the robot to plan actions, the world model must specify the ways in which things interact when the robot acts on the world—say, when it moves the cart from room 1 to room 2. What, besides the cart (and the robot), will end up in room 2? What entries in the data base, other than the ones for the cart and the robot, must be updated? Naturally, what must be updated are the entries for all the things resting on the cart (i.e., the telephone) except those tied by a short cord to the wall (again, the telephone), since they (and things resting on them) will fall on the floor of room 1 and thus will no longer be on the cart (nor will the receiver be on the hook). The story can be made as complicated as one wishes, and that is the source of the frame problem.

Why is this not problematic for humans or animals? An easy answer is that humans and animals have <u>common sense</u>, which robots lack, and this common sense has been gained through <u>experience</u>. But how does common sense work? How is experience acquired, and how is it used?

Most of this chapter has been about that very issue. The world model in the memory has been built from exposure to the world, that is, from experience. The statistical regularities of the world, including the system's own actions and their effects, are an integral part of the model. That means that not just the main effects of actions but also the side effects are recorded in memory. A system without experience cannot predict at all, and one with a lot of experience can produce comprehensive predictions. Therefore, by virtue of how the world model comes to be and how it works, it provides answers to what else might happen (e.g., when the robot pushes the cart from room 1 to room 2), much as a scale model of a physical object can provide answers about the behavior of the real object.

This gives rise to two comments, one about a scale model in the head and the other about the seriousness of the entire enterprise.

The idea of a scale model in the head may seem bizarre at first. Are there supposed to be tiny cats and dogs and trains and robots and telephones with cords, all in the head? Not at all, but there are <u>patterns</u> of activation of neurons <u>caused by those objects</u>. When the real objects are in front of us, they, too, are available to us only as patterns produced by our senses. These patterns are the objects that the brain deals with—not the objects themselves. The memory record is constructed from these very patterns, and the memory reproduces them in the focus more or less faithfully when properly cued. Thus, what the memory reproduces in the conscious part of the mind is of the same nature as what the senses produced there from the real stuff out in the world.

Let us turn to modeling in the physical world and look at the relationship between a physical object and a physical model of it. Let us assume that we want to find out what happens when two trains crash head-on. We can, of course, run the experiment with real trains and see what happens. The information would be reliable, but getting a large sample would be very expensive. A less expensive alternative is to build scale models of trains, run the experiment on them, and see what happens. Much could be learned from this, although the information would not be fully reliable because scale models do not behave in exactly the same way as their real counterparts.

Similarly with the real world and the world model stored in memory: We can establish a set of initial conditions in the world, let the world turn, and see or experience what the consequences are, or we can imagine a set of initial conditions, let the memory turn, and see or experience what consequences it produces. The more experienced the individual, the more faithful the world model and the better the memory's prediction of the consequences. In that sense, then, there is a scale model in the head: It produces in us experiences of the same nature as does the real world.

But who or what interprets the model; who or what interprets what comes out of the memory? The question can be answered indirectly: It is whoever or whatever interprets the world. A direct answer would be more satisfactory, but for that, instead of asking who or what interprets, it is better to ask how the model, or the world, gets interpreted by the whoever. Furthermore, we need to look at the meaning of the word 'interpret'.

The interpretation of a signal, a situation, or a message by a subject manifests itself in the reaction that the thing evokes in the subject. The reaction can be internal or external, 'internal' meaning subjective experience (pleasure, pain, emotion, association, propensity to act--the things that we call 'mental') and 'external' meaning action (e.g., dodging a fastball). We judge the correctness of an interpretation by how appropriate the subjective experience or the triggered action is to the conditions causing it. If it is inappropriate, we say that the subject does not understand the situation or the message.

Observable action involves the use of the muscles. Some actions are wired in as automatic reflexes; others are learned. The learned ones are of interest to us here. In the section on learning to act, we considered how actions can become associated with external cues. For present purposes, the important thing is that the patterns from which the world model is constructed include components for action. When the memory reproduces patterns in the focus, the action components of these patterns are ready to drive the muscles. The stored world model thus includes the system's own actions, so the system can interpret situations and messages via actions.

Subjective experience is a subtler way for a system to interpret situations and messages, but it too can be explained by the memory model. Consider the propensity to act (e.g., back-seat driving) and planning: Present or imagined cues together with the predictive power of the memory bring to consciousness (focus) possible future actions and consequences. However, the system blocks off commands to the muscles, so that there is no immediately observable action even if interpretation based on the world model is going on within. In that sense, subjective experience is just as real a way to interpret situations and messages as is action.

Finally, there are the most basic forms of interpretation. With animals and humans, some things have meaning in themselves and can thus be interpreted without further learning. The experiencing of certain things as pleasurable or painful, and possibly some emotions, are handed down genetically. We have modeled them with the built-in preference function, which gives basic meaning to the world.

In the paragraphs above we have considered only the very basics of interpretation and meaning, but these basics appear to operate in all intelligent beings. With higher animals, and with humans in particular, social learning is exceedingly important, and the resulting web of interpretations and meanings becomes very complex. Even then, the mechanisms of interpretation and meaning can be few and simple, akin to those discussed above, with the complexity arising from the infinity of ways in which the mechanisms allow new meanings to be derived from old ones.

We can now attempt to say who interprets the world or the world model: The individual does. And what is the individual? A composite of sensors and motors, possessing a built-in preference function for some sensory patterns and capable of building from its own sensations and actions a world model for future reference. The preference function and the world model or memory are the means by which the individual interprets, and the motors allow interpretations to be expressed externally.

The second comment is about the seriousness of this approach overall. Can traditional artificial-intelligence methods be replaced with a memory that somehow produces right answers automatically? Can there be such a memory? First, it is not clear that one method has to replace the other, although it is quite clear that the traditional methods alone are in trouble. Second, the memory is mathematically sound and easily built from neuron-like components, and it does not guarantee right answers any more than biological memories do. Third, the retrieval properties, including the ways in which the memory fails, are lifelike, regardless of the extreme simplicity of the model.

Since nature has solved the frame problem, we should be able to solve it by understanding how information processing is organized in animals. To the extent that the memory model captures that organization, it is relevant to the solution of the problem. The position taken here is that the memory, as I have modeled it, contributes to the solution significantly, and that an equally significant contribution is made by the sensory system that prepares information for the memory. Thus, a major part of the burden is on the sensory system. That part is the topic of the next section.

The Encoding Problem

The sensors of the various modalities collectively receive a mass of stimuli of a specific type, and from it they derive patterns for processing by the nervous system. From these patterns the brain builds its model of the world. As the model learns to reproduce regularities of the world, it allows the system to predict and to contemplate the consequences of its own actions, making it possible for the system to plan.

The raw signal arriving at the sense organs is ill suited for building a predictive model. Even if a number of regularities of the world are present in the signal, they appear in far-from-optimal form and are embedded in noise. A cursory look at raw speech waves, for example, makes one wonder how anything of importance can be extracted from them; the waves for the same word spoken by different people can look very different. A sensory system thus has two functions: to filter out noise and to transform relevant information into a form that is useful in building and using the world model.

Transforming the input signal into a form useful for modeling the world is referred to here as the <u>encoding problem</u>. In my model for an autonomous system, the encoding task falls on the sensory system, which is assisted by memory. For an example, let us take vision in our three-dimensional world inhabited by various kinds of objects. The world model needs encodings of those objects, and the visual system has to produce the encodings. From how the memory works we can derive requirements for a good visual encoding. Since patterns stored in memory attract similar patterns, the memory chunks things with similar encodings, forming objects and individuals from them. On the other hand, the retinal image of an object varies widely according to the distance between the object and the viewer; yet those very different images should produce very similar encodings. The job of the visual system, then, is to express the retinal image in features that are relatively insensitive to scale, among other things. Similarly, to understand the speech of different individuals having vocal cords of different length, the auditory system needs to express the audio signal in features that are relatively insensitive to absolute pitch, among other things. Similar remarks can be made for the other senses.

Often a given input signal can be encoded in several different ways, and yet we seem to have only one interpretation of it at a time. How we perceive the Necker cube is an example of this, as the interpretations of our looking at it from above and from below flip back and forth. This can be attributed to assistance or feedback from memory. If the sensory system can produce an encoding of something familiar, it tends to do so. Note that familiarity implies memory.

Once the objects of the world have been encoded properly, a sparse distributed memory can form a dynamic model of the world from the encoded objects. The model will then let us examine the effects, direct and indirect, of contemplated actions in the same way--that is, with the same machinery--as we, as observers of the world, examine the effects of real actions.

This solution to the frame problem is truly a solution only if we can solve the encoding problem. The likelihood of that depends on how good our model of an autonomous system is, and that in turn depends on how well it captures the essence of how animals are organized.

It seems to me that solving the frame problem will require much work, and much of the work has to be devoted to the understanding of sensory systems. In that work, the models of the memory and of an autonomous learning system can serve as valuable guides.

Related work. Our picture here of an autonomous system and of the encoding of sensory data is extremely simple. Grossberg's (1980) paper on the formation of a stable cognitive code goes into the subject more deeply. That paper and Albus' (1981) book emphasize the hierarchical organization of intelligent systems. The autonomous systems of the present paper are roughly equivalent to a single layer in a hierarchy proposed by Albus.

Anderson (1986) and Anderson and Murphy (1986) have emphasized the crucial role of the encoded form of information—that is, the actual representation itself rather than what an encoding represents. (The importance of representation is also appreciated in the field of artificial intelligence, but usually only high-level representations are considered, as in the problem of covering an 8 x 8 board with 1 x 2 domino pieces after a pair of diagonally opposite corner squares have been removed.) Their work and mine suggest that we need to mind the representations at the very lowest levels, and that representations of at least some higher-level concepts might be derived by mechanically combining the encodings of lower-level concepts. Furthermore, it is of utmost importance that the representations be suited for highly parallel computation; at least this is so for brains, which are made of relatively slow neurons.

Traditional artificial intelligence is modeled on how humans reason and how they describe their thinking and their problem solving. These phenomena are at the highest, most conscious levels of human behavior, and are rather serial in nature. Artificial-intelligence methods perform poorly on tasks (such as pattern recognition) that happen at lower levels and that are, from the subjective point of view, automatic. In my memory model, serial phenomena and pattern recognition have very different statuses. Serial phenomena are modeled by stored associations between patterns—that is, by pointer chains—whereas the memory's power for pattern recognition comes from the metric properties of the pattern space, which the memory exploits. However, even serial recall is based on the recognition of patterns and on the convergence of iterated reading to stored patterns and sequences. Thus, the geometry of the pattern space, or the structure of the symbols, determines many properties of the memory. It seems, then, that artificial-intelligence methods need to be augmented with mathematical and statistical methods of dealing with representations in high-dimensional spaces. Thus, in addition to symbolic structures we need to study the structure of symbols. This point is made emphatically in Hofstadter 1985—see, in particular, Chapter 26, "Waking Up from the Boolean Dream."

Summary and Conclusions

In this paper I have developed a model of memory that captures some basic properties of human long-term memory. Although human memory is much more complicated than my model of it or the models of others, it is essential to understand simple models of the right kind before we can hope to develop more comprehensive models and to understand the full phenomenon of memory. The sparse-distributed-memory model is offered in that spirit.

In my modeling, very large patterns of features encode moments of experience, and sequences of such patterns model sequences of events that occur over time. Because the patterns stored in memory can also be used to address the memory, sequences can be stored as pointer chains. Any pattern in a sequence, or a sufficiently similar pattern, can then be used to retrieve the rest of the sequence. The sequences can be arbitrarily long, because the capacity of the memory can be made arbitrarily large by making the number of storage locations sufficiently large. Simple pointer chains cannot handle crossings of sequences. For sequence crossings, the memory model has multiple folds, each associated with its own delay parameter.

Memory plays but a part (though an important part) in human cognition: It stores a dynamic, predictive model of the world. Another part is the extraction of information from the world and the encoding of it before it is stored. That part is carried out by the senses of sight, hearing, touch, and the other senses with assistance from memory. My treatment of sensory systems has been very general and can be summarized as follows: The memory works with features and creates its internal objects and individuals by chunking together things that are similar in terms of those features. In order for those internal objects to match objects of the world, the system's sensors must transform raw input from the world into features that are relatively invariant over small perturbations of objects. To recall a stored "object," the senses—or the memory—must produce a reasonable approximation of the encoding that was used as an address when the object was stored.

Yet another part of human cognition has to do with action as a way of affecting the world. Actions are carried out by motors or muscles. My modeling of motor systems has been very general and abstract, the main point being that the motors are controlled by sequences of patterns that can be stored in memory.

I have combined these ideas in a simple model of an autonomous learning system. The system has a central place, the <u>focus</u>, that accounts for the system's subjective experience. The entity in the focus is a very large pattern, a high-dimensional vector of features that encodes everything about that moment (that is, any specific things that the system may be attending to, the system's action, and the overall context). The memory is addressed by the focus, the memory's output goes into the focus, the senses feed into the focus, and the muscles are driven from the focus. This architecture is motivated by the oneness of subjective experience; an experience created by the senses can also be created by the memory. The system's modeling of the world is founded on this idea.

A system with such an architecture seems capable of learning how the world works and of learning how its own actions affect the world (including how they affect its own well-being). The well-being is modeled by a built-in preference function that is defined on the states of the focus. In learning to act, the system needs to store favorable action sequences in memory and to assign positive and negative preferences to previously indifferent states. In the most advanced form of learning, namely imitation, the system uses itself to model the behavior of others of its kind.

Besides possibly helping us understand human and animal memory, the present research suggests a way to build a new kind of computer memory: a random-access memory for very long words with approximate addressing. To use such a memory in robots, we have to learn to encode information about the world and about motor action into high-dimensional feature vectors. Major research topics for the future thus include sensory encoding, motor action, and memory storage. These three topics entail very different problems, all of which will have to be solved if we are to build robots that can operate with any reasonable degree of autonomy.

REFERENCES

Albus, J. S. 1981. Brains, behavior, and robotics. Peterborough, N.H.: BYTE Books.

Anderson, J. A. 1986. Cognitive capabilities of a parallel system. In E. Bienenstock, F. Fogelman Soulié, and G. Weisbuch (Eds.), Disordered systems and biological organization (NATO ASI Series F, Vol. 20). Berlin: Springer-Verlag. Pp. 209-226.

Anderson, J. A., and Murphy, G. L. 1986. Psychological concepts in a parallel system. Physica 22D:318-336.

Grossberg, S. 1980. How does a brain build a cognitive code? Psychological Review 87:1-51. Reprinted in S. Grossberg, Studies of mind and brain (Boston Studies in the Philosophy of Science, Vol. 70). Boston: Reidel, 1983.

Hofstadter, D. R. 1985. Metamagical themas. New York: Basic Books.

Holland, J. H. 1986. Escaping brittleness: The possibilities of general-purpose learning algorithms applied to parallel rule-based systems. In R. S. Michalski, J. G. Carbonell, and T. M. Mitchell (Eds.), Machine learning: An artificial intelligence approach, Volume 2. Los Altos, Calif.: Kaufmann. Pp. 593-623.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., and Thagard, P. R. 1986. Induction: Processes of inference, learning, and discovery. Cambridge, Mass.: MIT Press.

Kanerva, P. 1984. Self-propagating search: A unified theory of memory (Report No. CSLI-84-7). Stanford: Center for the Study of Language and Information.

Loftus, E. F. 1979. Eyewitness testimony. Cambridge, Mass.: Harvard University Press.

Pylyshyn, Z. W. (Ed.) 1987. The robot's dilemma: The frame problem of artificial intelligence. Norwood, N.J.: Ablex.

Schank, R. C. 1982. Dynamic memory. Cambridge, England: Cambridge University Press.

RIACS
Mail Stop 230-5
NASA Ames Research Center
Moffett Field, CA 94035
(415) 694-6363

The Research Institute for Advanced Computer Science
is operated by
Universities Space Research Association
The American City Building
Suite 311
Columbia, MD 21044
(301) 730-2656