AN ABSTRACT MACHINE BASED ON CLASSICAL ASSOCIATION PSYCHOLOGY

Richard F. Reiss
Librascope Division
General Precision, Inc.
Glendale, California

Summary

The theories of classical association psychology (circa 1750-1900) attempted to explain human thought processes in terms of certain mechanistic forces operating on discrete entities called "sensations," "images," and "ideas." Although these theories have become unfashionable since the turn of the century, due primarily to their ambiguity and the difficulty of experimental verification, and whereas they may never prove adequate for human psychology, it is possible that they may, nevertheless, provide a fruitful basis for some types of artificial intelligence. One method of exploring ramifications of the classical theories is the formulation of an abstract "machine" which constitutes an interpretation of the theories and whose behavior can be examined in any desired detail. In this paper such a machine is partially constructed, and some of its behavioral features and problems are discussed.

Introduction

The general problem of synthesizing intelligent artifacts is sufficiently broad and difficult to justify a variety of research strategies. One such strategy is to select a psychological theory, however inadequate (as they all are), and examine it for useful ideas and insights. This strategy has been used in the studies on which this paper is based.

Classical association psychology was developed primarily in England and enjoyed considerable popularity for more than a century (circa 1750-1900). Space does not permit even a brief recount of its history, but some of the key works will be mentioned for purposes of reference. Although Aristotle vaguely recognized the minimal association principles (see Hamilton89), the modern development of association theory began with some hypotheses laid down by the British philosophers Hobbes10, Locke13, Berkeley7, and Hume11. They were, however, concerned primarily with epistemology rather than psychology. Hartley9 is generally considered to be the founder of associationism as a psychological doctrine, and it flowered during the 19th century due largely to the efforts of Brown7, James and John Stuart Mill14,15 and, above all, Bain1,2,3,4 and Spencer16. For a detailed history, see Warren19.

Association theory proposed to explain the highest levels of conscious thought processes, perhaps the most elusive of all experimental subjects. By 1900, the revolutionary development of physiological and experimental psychology had created an atmosphere which was very hostile to any theory that could not quantify its concepts and be subjected to precise experimental testing. Associationists were unable to meet these criteria successfully, and the classical theories in their general forms were rapidly abandoned. We say theories, rather than theory, because associationists rarely agreed on more than a few basic concepts. Later we shall simply refer to "the classical theory," by which we mean those concepts that did receive widespread acceptance. Some of these concepts are to be found, frequently disguised by new terminology, in many contemporary schools of psychology. But the old theories with their sweeping assumptions have been effectively dead for 50 years.

The problem of artificial intelligence provides a distinctively new framework for evaluating psychological theories. If one is concerned with the subproblem of mechanizing human thought processes as they actually occur, then of course the validity of a psychological theory is of great importance. The general problem, however, does not require such stringent criteria; the search is for suggestive material, in whatever form.

The association theories are of interest on three counts: they were produced by several brilliant minds struggling through introspection with the problem of mind; they involve the action of a few mechanistic forces operating on discrete entities; and they aim at explaining the higher thought processes. Unfortunately, the associationists made no sustained attempts to quantify their assumptions and search out the detailed consequences. This must be done, then, for the first time (to my knowledge), and the resulting abstract machine represents just one possible interpretation of classical theory. Another interpreter might well produce a machine that differs in many respects. Some alternatives will be noted in the discussion below, but in general the postulates will be introduced without apologies or historical justifications.

From Hartley on, associationism was generally tied to parallel theories of the nervous system. In constructing the abstract machine, this facet of association theory will be almost entirely ignored. Except in the concluding remarks, the problem of physically realizing the abstract machine will likewise be ignored.

The objectives here are simply to show what a machine based on association theory might look like, to demonstrate some of its behavioral characteristics, and to raise problems that appear important to further research along these lines. Our strategy in this project has been to strive...
for a reasonably complete machine without regard for our ability to analyze its behavior; we hope thereby to avoid the premature selection of problems that appear interesting but are actually trivial. The nine postulates (together with subordinate assumptions) which are discussed here fall far short of a complete machine. But combined with remarks on further postulates in the concluding section, they should be adequate for the purposes of this paper.

The postulates will be introduced in three groups. After each group has been stated, some behavioral consequences of that group will be examined in detail and a few generalizations ventured.

**The Four-Postulate Machine**

The first four postulates form a group which produces a basic machine without sensory inputs or motor outputs. The behavioral potentialities of this machine must be examined in some detail before proceeding to further postulates.

**Postulate P1:** There exists a finite set $M$ of "memory tokens" $m_1$, $m_2$, $\ldots$, $m_\mu$.

Initially we make no assumptions whatever regarding the nature, structure, or properties of these tokens although they obviously represent hypothetical entities—such as "images," "ideas," "impressions," "concepts" with which the associationists furnished the mind. The number of memory tokens, $\mu$, is assumed constant; and we do not inquire into the origin of these tokens at present.

**Postulate P2:** Between each pair of memory tokens, $m_i$ and $m_j$, in $M$ there are two "bonds" whose "strengths" are given by "coupling coefficients." The bond from $m_i$ to $m_j$ is represented by the coupling coefficient $c_{ij}$; and the bond from $m_j$ to $m_i$ by $c_{ji}$.

The "bonds" introduced by P2 are first approximations to the associationists' "connections" and "attractions." Although the associationists seem to have generally assumed that connections are symmetrical in the human mind, it was and is a point of controversy; and we here take the more general case where the bonds from $m_i$ to $m_j$ and from $m_j$ to $m_i$ are treated as independent entities. It will be seen later that these bonds must, in turn, be split into two components.

The coupling coefficients may be organized to form a square $(\mu \times \mu)$ matrix where $c_{ij}$ is the element in the $i$th row and $j$th column. This will be called the coupling matrix and will be denoted by $\{c_{ij}\}$. Digraphs will be used to illustrate small token groups. Figure 1a shows a system of five memory tokens, and Figure 1b is the equivalent coupling matrix. It will be noted, in the graph, that bonds with zero coefficients have been left out and that symmetrical bonds (as between $m_3$ and $m_4$) are represented by a single two-headed arrow. The statement that $"m_1$ is not bonded to $m_j"$ will mean that $c_{1j} = 0$. All coupling coefficients are assumed to range over the positive real numbers, with a lower limit of zero and no upper limit specified at present.

**Postulate P3:** There is an "attention register" which can hold a set $A$ of memory tokens. The maximum number of tokens in $A$ is $\lambda$, and this integer will be called the "length" of the attention register.

The introduction of the attention register completes the hardware of the four-postulate machine. It will be assumed that $\lambda \ll \mu$ and that $\lambda$ is, therefore, a proper subset of $M$. The intent of the adjective "attention," if not immediately apparent, will become clear in the course of discussion. In this paper, $\lambda$ will be considered an independent variable, or parameter, of the machine.

Before stating the fourth postulate, it will be necessary to introduce a special function and a related set. For every memory token $m_i$ we define the adduction function $f_i$. If $m_i$ is a member of $A$, i.e., is in the attention register, then $f_i > 0$. If $m_i$ is not a member of $A$, then the value of $f_i$ is given as

$$f_i = \sum_{j} c_{ij}$$

where $j$ ranges over the subscripts of tokens in $A$. Thus if $m_i$ is not in $A$, $f_i$ is the sum of the coupling coefficients of bonds from all tokens in $A$ to token $m_i$. In terms of the coupling matrix, $f_i$ is the sum of all those elements in the $i$th column, which are also in rows corresponding to the tokens in $A$.

The adduction set $F$ is defined as follows: memory token $m_i$ is a member of $F$ if, and only if, there exists no memory token $m_j$, which $f_j > f_i$. In other words, the adduction set is composed of the token(s) having the largest adduction function value(s). Be it noted that $F$ cannot be empty; indeed, if all coupling coefficients are zero, $F = M$.

**Postulate P4:** Every at seconds a memory token is chosen at random from the adduction set $F$ and entered into the attention register, i.e., set $A$. If the attention register is already filled, then the "oldest resident" will be simultaneously ejected.

This postulate introduces action into the abstract machine. The state of the machine at any moment is defined by the coupling matrix and a list of memory tokens in the attention register. Every at seconds the machine changes state by entering a memory token into the attention register and ejecting the token that has been longest resident in the register. It will be assumed that all members of the adduction set $F$, at any moment, have equal chances of being selected for entry into set $A$; i.e., if there are $n$ members of $F$, the probability that any particular member will be selected is $1/n$. If $F$ contains only one member,
then the next token to enter \( A \) is uniquely determined.

The sequencing of tokens through the attention register may or may not be a stochastic process, depending upon the nature of the coupling matrix. For example, if there is exactly one non-zero element in each row and column, then the machine's behavior for all future times is precisely determined by the tokens initially placed in the attention register. On the other hand, it is easy to formulate any number of coupling matrices which produce stochastic behavior. It should be noted that the coupling matrix is not a Markov matrix; the coupling coefficients are not transition probabilities.

**Discussion of the Four-Postulate Machine**

The first four postulates produce a class of machines which clarifies certain problems and potentialities inherent in classical association theory. The sequencing of memory tokens through the attention register might be interpreted as a first, crude approximation to the flux of ideas, images, etc. into and out of the conscious state. The associationists generally believed that thought processes are conscious processes and that introspection is, therefore, a powerful tool for the analysis of thought. As Bain put it, "...Introspection is still our main resort—the alpha and omega of psychological inquiry: it is alone supreme, everything else subsidiary. Its compass is ten times all the other methods put together, and fifty times the utmost range of psycho-physics alone." By the time Bain died (1903), evidence was rapidly accumulating in favor of the thesis that subconscious processes influence thinking to a remarkable extent. This development, compounded by failing attempts to make introspection a precise analytical tool and various other trends (see Boring's history) led to a widespread rejection of introspection. It seems the associationists' theory had become so closely connected to the introspective technique that, when the technique was discredited, the theory was discredited too. Actually the failures of introspective technique show only that the association theory is difficult to test, not that it is invalid. It is by no means necessary to assume that all, or even most, association processes occur at a conscious level. In fact association theory has survived (in a simpler form) in theories of conditioning and learning which ignore consciousness altogether.

Since we are not concerned here with the validation of a psychological theory, and since consciousness is not in any event a necessary constituent of association processes (although the classical associationists failed to recognize this fact), we shall not postulate any relations between consciousness and the states of our abstract machine. As a consequence of P4, a memory token must enter the attention register in order to influence the behavior of the machine. Nothing more is to be read into that postulate unless it amuses the reader to do so.

Postulates P2, P3, and P4, taken together, are the result of a difficult decision that must be made in any interpretation of the classical theory. At times the associationists spoke of the connections between mental entities as though they were little more than indicators of possible, or preferred, sequences among which some higher agency chooses a particular path. Thus Brown, for example, insisted on using "suggestion" rather than "association." On the other hand, the physiological bias initiated by Hartley and particularly noticeable in Bain and Spencer, often resulted in associative connections being treated as physical forces or neural links competing in a completely mechanical way. Reading between the lines, the associationists were apparently torn between a desire to create a mechanistic, deterministic theory, resembling contemporary physical and chemical theories, and a desire to preserve the "free will" of an immanent "soul." (It might be noted that a theist engaged in artificial intelligence research today is confronted by a similar dilemma.)

I have been unable to find in the literature a clear, unambiguous discussion of the selection problem inherent in any system of competing associations. But this problem cannot be avoided in the formulation of our abstract machine. The introduction of a deus ex machina, some free will demon which selects tokens, would merely beg the question as to how much might be accomplished by the relatively simple associationist machinery. There seem to me to be two main alternatives. Coupling coefficients might be introduced as transition probabilities with P4 replaced by some rule governing joint probabilities, in which case the machine would be heavily biased toward stochastic processes. The other alternative is exemplified by the postulates used here, with the machine biased toward deterministic behavior. This bias will be magnified by postulates P5 and P6 so that the machine will, in the absence of sensory inputs, tend toward cyclic, stereotyped behavior patterns with few if any random fluctuations. We proceed to examine briefly a few behavioral characteristics of the four-postulate machine.

Assume the machine contains just five tokens, bonded as in Figure 1, and that the attention register can hold only one token, i.e., \( A = 1 \). Suppose that \( m_1 \) is initially in the attention register. The sequence of tokens entering \( A \) will then be \( m_2, m_3, m_4, m_5, \ldots \) an endlessly repeating cycle. Let \( m_3 \) be initially in \( A \). Now the addition function for all tokens has the value zero, and the addition set \( F \) contains all tokens. A token from \( F \) is chosen at random. If \( m_2 \) is selected, then the process is repeated until some other token enters \( A \). The selection of \( m_5 \) will result in \( m_3 \) returning to \( A \). Eventually \( m_1, m_2, \) or \( m_5 \) will enter \( A \), and the endless sequence above will follow.

This example shows two characteristics of the abstract machine. First, a token that has no bonds to other tokens, such as \( m_3 \) (and corresponding to a "terminal node" in graph theory), cannot stop the machine. It simply provides an equal
opportunity for all tokens to enter \( A \). Secondly, the machine can be easily trapped into cyclic behavior by rings of two (such as \( m_2 \) and \( m_4 \)), or more tokens. It should be noted that if the reflexive bond from \( m_4 \) to \( m_4 \) had a coupling coefficient greater than 5, it would not succeed itself since its adduction function is zero when it is in \( A \).

Consider the case where \( \lambda = 2 \), and let \( A = \{m_1\} \) initially. The sequence of tokens entering the attention register will be \( m_3, m_4, m_1, m_2, m_3, \ldots \), another endless cycle. No matter which token, or pair of tokens, is initially in \( A \), the machine will eventually settle down to this three-token cycle. This little example demonstrates that the machine can be trapped into cyclic behavior by networks other than rings, in this case the chain \( m_1, m_2, m_3 \). It is also clear that at least \( \lambda + 1 \) tokens are required for cyclic sequencing.

Before examining further examples of simple coupling patterns, it will be convenient to define "strings," "rings," "chains," and "trees" in terms of the "predecessor-successor" relation. If \( c_{ij} > 0 \), then \( m_j \) is a "predecessor" of \( m_i \) and \( m_i \) is a "successor" of \( m_j \). A "string" is a group of tokens containing a unique "origin" token which has no predecessors and exactly one successor in the group, a unique "terminal" token having no successors and exactly one predecessor in the group, and all other tokens in the group having exactly one predecessor and one successor in the group. A "ring" is simply a closed string, i.e., a string without origin or terminal tokens. A "chain," as the name suggests, has two "end" tokens, each of which is connected to one token in the group, which token is both a successor and predecessor; all other tokens in the group are connected to exactly two others which are also both successors and predecessors. Thus two strings are "embedded" in a chain. A "tree" will be understood to be a group with one origin token having no predecessors and two or more successors, several terminal elements having one predecessor and no successors, and all other tokens in the group having one predecessor and two or more successors. These inelegant definitions will suffice for the rather informal discussions of token structures in this paper.

As a consequence of the adduction function, all of the tokens in the attention register influence the selection of the next token to be entered into \( A \). This corresponds to the process Bain called "compound association." Unfortunately, the associationists did not pursue in detail the implications of such a process. They tended to think in terms of simple strings and trees of tokens (using our terminology), and compound association is not very important in such cases. Consider, for example, the case where all memory tokens are organized into a single string and any one of them is initially in \( A \). Even if \( \lambda \) is large once \( A \) is filled the adduction function will be zero for all tokens except the successor of the last token to enter \( A \). Consequently, the sequencing of tokens through \( A \) will correspond to the ordering of tokens in the string regardless of the value of \( \lambda \). Thought experiments of this sort lead to the general conclusion that the fewer the bonds between memory tokens, the less important is \( \lambda \).

In terms of association psychology, this conclusion could be characterized by the general rule that a mind in which ideas are sparsely associated cannot benefit from an increase in attention span. The behavior of such a mind will tend to be stereotyped in spite of efforts to increase the attention span. Certainly the behavior of the abstract machine becomes independent of \( \lambda \) if memory tokens are organized into simple strings.

However, even simple coupling patterns produce some interesting sequencing phenomena, and we shall briefly examine a few of these. Let us first consider the effects of symmetrical bonds, i.e., cases where \( c_{ij} = c_{ji} \). The four tokens in Figure 2 are connected by symmetrical bonds having the indicated strengths. Suppose that \( \lambda = 1 \) and that \( A = \{m_1\} \) initially. The sequence of tokens through \( A \), which will henceforth be called the \( A \)-sequence, is \( m_1, m_2, m_3, m_2, m_3, \ldots \). If \( A = \{m_2\} \) initially, the \( A \)-sequence will be \( m_1, m_2, m_3, m_2, m_3, \ldots \). Thus we see immediately that symmetrical bonds do not necessarily lead to symmetric sequences, i.e., \( m_1, m_2, m_3 \) versus \( m_3, m_2, m_1 \). Furthermore, an increase in \( \lambda \) can have marked effects. Let \( \lambda = 2 \) and initially \( A = \{m_1\} \). The \( A \)-sequence will be \( m_1, m_2, m_1, m_2, m_1, \ldots \), a type of behavior seen in small groups without symmetrical bonds.

In brief, experiments on a variety of small coupling matrices have failed to show any outstanding effects uniquely correlated with the presence of symmetric bonds or that such effects might emerge in large ensembles of memory tokens.

For the sake of simplicity, symmetrical bonds are assumed in the five-token group of Figure 3, which illustrates a "hub and spoke" pattern of coupling (a special type of tree). The bonds between the hub \( m_1 \) and the outlying tokens have progressively smaller coupling coefficients. The effect of increasing \( \lambda \) is quite regular for such a cluster of memory tokens. Let \( A = \{m_1\} \) initially, and suppose \( \lambda = 1 \). The \( A \)-sequence will be \( m_1, m_2, m_1, m_2, \ldots \). For \( \lambda = 2 \), the \( A \)-sequence is \( m_1, m_2, m_3, m_1, m_2, m_3, m_1, \ldots \); and for \( \lambda = 3 \), it is \( m_1, m_2, m_3, m_1, m_2, m_3, \ldots \). Thus the "hub and spoke" configuration provides a basic mechanism for cyclic scanning of memory tokens, the scope of the scanning process being controlled by \( \lambda \). A simple ring will also produce cyclic scanning, but the scope of scanning cannot be controlled by \( \lambda \).

This is an example of sparsely bonded token clusters which produce marked changes in sequencing as a function of \( \lambda \); it is a counterexample for the general rule stated above.

Although Gestalt psychology has been a major opponent of association theory, it turns out that gestalt phenomena can appear in the abstract machine. Consider the group of memory tokens in Figure 4. Suppose that \( \lambda = 1 \) and that \( m_3 \) is in \( A \),
The "response" of the machine is the entry of \( m_2 \) into \( A \). If \( m_3 \) is in \( A \), then the response will be the entry of \( m_3 \) into \( A \). However, if \( \lambda = 2 \) and both \( m_1 \) and \( m_3 \) are inserted in \( A \), then the immediate response will be the entry of \( m_4 \) into \( A \). Thus the machine's response to the joint presentation of \( m_1 \) and \( m_3 \) in \( A \) is quite different from the "sum" of its responses to \( m_1 \) and \( m_3 \) presented separately. This fundamental characteristic of gestalt processes is therefore obtainable with associationistic mechanisms; there is no necessary conflict between association theory and gestalt phenomena. This case will become more significant after sensory inputs to the attention register are introduced below.

The length \( \lambda \) of the attention register can influence the sequencing through tree structures. Consider the simple tree in Figure 5. Let \( \lambda = 1 \) and \( A = \{m_1\} \) initially. The A-sequence will be \( m_1, m_2, m_3, m_4, m_5, \ldots \). For \( \lambda = 2 \), the A-sequence will be the same; but if \( \lambda = 3 \), a new type of behavior suddenly appears. It might be called "branch jumping." For \( \lambda = 3 \), the A-sequence will be \( m_1, m_2, m_3, m_4, m_5, m_7, m_9, m_1 \), \ldots Instead of continuing out the upper branch, the A-sequence jumps down to the lower branch and continues on that branch. The crucial factor is the retention of \( m_3 \) in \( A \) when the weak bond between \( m_3 \) and \( m_5 \) is reached. At this point, \( f_2 = 1 \) while \( f_3 = 2 \); and therefore \( m_3 \) enters \( A \) instead of \( m_5 \). Thus if \( \lambda > 1 \), the A-sequence is not determined simply by the relative strengths of coupling coefficients at branch points in a tree. Clearly, small changes in \( \lambda \) could effect great changes in the machine's behavior if memory tokens are organized into certain types of trees, another exception to the general rule that \( \lambda \) is an ineffective control parameter for sparsely bonded token systems.

A case of interest is the basic configuration of "intersecting" strings, i.e., where a memory token is the member of two strings. Two simple strings are embedded in the system of Figure 6, \( m_1 \) through \( m_3 \) and \( m_5 \) through \( m_9 \) (including \( m_3 \)). Of course several other strings can be abstracted from this group, but there is no need to identify them. Let us assume that all coupling coefficients have the same value, say, 1.0. Consider first the case where \( \lambda = 1 \) and \( A = \{m_1\} \) initially. The adduction set \( F = \{m_2, m_5\} \), and either token may be chosen to succeed \( m_1 \) in \( A \). In fact, set \( F \) always contains two tokens, regardless of which token is in \( A \). Thus the machine operates in a stochastic mode over this group of tokens, and many A-sequences are possible, e.g., \( m_1, m_2, m_4, \ldots \) or \( m_1, m_3, m_5, m_7, \ldots \). If \( A = m_2 \) initially, the situation is essentially the same.

Now let \( \lambda = 2 \). The behavior of the machine shifts radically from a purely stochastic to a purely deterministic mode! Furthermore the intersection of the several strings does not provide opportunities for "turning corners." To see this, assume that initially \( A = \{m_1, m_3\} \), with \( m_1 \) the "oldest" token (it will therefore be ejected first). The A-sequence is guaranteed to be \( m_1, m_2, m_3, m_4, \ldots \) and on out the horizontal strings. Similarly, if the machine is moving down the vertical strings, the A-sequence will be \( \ldots m_4, m_7, m_3, m_9, \ldots \) and so on.

The bonds from a token to the next but one token, e.g., from \( m_1 \) to \( m_2 \), from \( m_2 \) to \( m_4 \), etc., correspond to what are sometimes called "remote forward associations" in serial learning experiments. While we will not venture a direct interpretation of this abstract machine behavior in terms of serial learning, we will adapt the psychological phraseology to our purposes and refer to these bonds (in such overlapping string structures) as "remote forward bonds."

The example above shows that strings accompanied by remote forward bonds may intersect at many tokens without producing ambiguity in the sequencing. When \( \lambda = 2 \) (or more), it is as if the sequencing process acquires "inertia" which carries it right through intersections without deflection. Thus a certain economy of memory tokens is possible in the abstract machine; a particular token may be coupled in numerous strings without destroying the independence of corresponding A-sequences.

This example also shows why it is dangerous to characterize the machine as either stochastic or deterministic. Given even this simple coupling configuration, the small extension of the attention register from \( \lambda = 1 \) to \( \lambda = 2 \) is sufficient to shift the machine's behavior from one extreme to the other. From the standpoint of classical association psychology, this phenomenon is, of course, very suggestive. The strings with remote forward bonds might well represent habitual sequences of "thoughts" and "ideas." If, then, the attention span of an individual is reduced to a minimum, a condition approached perhaps in "free-association" experiments and dreams, the behavior of the abstract machine suggests that the normal, habitual thought sequences might be broken up and replaced by random, novel sequences.

Having examined a few fundamental behavioral capabilities of the four-postulate machine and engaged in some speculative psychological interpretation, we shall proceed to introduce two more postulates.

The Six-Postulate Machine

Among the numerous types or "laws" of association that have been proposed there are two that gained almost universal acceptance. They were most commonly called "Similarity" and "Contiguity." The "law of Similarity" may be summarized as the tendency of ideas or images to become associated in the mind if they have certain "similarities" in quality, structure, function, etc. For example, the concept of a cat may be succeeded in consciousness by that of a rat because of visual or auditory similarities in the names. The "law of Contiguity" draws attention to the fact that ideas or images which have appeared together or in close succession in the mind tend to become associated.
This principle, translated into a behaviorist framework, underlies most theories of conditioning.

Some associationists argued that the law of similarity could be derived from the law of contiguity, while others argued the opposite. There were debates concerning the nature of the relation of "similarity" (Germane to the modern problems of pattern recognition) and so on. See William James for an excellent critical discussion of such issues; in this paper they will be generally avoided. We shall simply postulate two kinds of bonds suggested by the "laws" of similarity and contiguity, ignore all other types of associative connection, and examine the consequences.

**Postulate P5:** Every coupling coefficient $c_{ij}$ is the sum of an "external coefficient" $a_{ij}$ and an "internal coefficient" $b_{ij}$; that is, $c_{ij} = a_{ij} + b_{ij}$ for all $m_i$ and $m_j$ in $M$.

This is an extension of P2 wherein the bonds previously discussed are split into two subtypes which will be called "external bonds" and "internal bonds." As a consequence, the coupling matrix is the sum of an "external" matrix and an "internal" matrix. This postulate does not require alteration of the definitions of addition functions or sets, nor of postulates P3 and P4.

**Postulate P6:** Initially, the external coefficient $a_{ij} = 0$ for all $m_i$ and $m_j$ in $M$. Thereafter, whenever both $m_i$ and $m_j$ appear in the attention register, $a_{ij}$ is incremented by an amount $\delta_{ij}$. The internal coefficient $b_{ij}$ is constant for all time for all $m_i$ and $m_j$ in $M$.

Thus, at "birth," the machine's external coefficients are zero; and the internal coefficients have values that will be retained throughout the life of the machine. Initially, $c_{ij} = b_{ij}$, and therefore of the coupling matrix, will be increased each time during the lifetime of the machine.

At any time the state of the external matrix reflects the "experience" of the machine up to that time, i.e., the past behavior of the machine, the sequencing of memory tokens through $A$. The internal matrix, on the other hand, reflects certain constant relations between the tokens. The strength of an internal bond, as given by the corresponding internal coefficient, is intended to represent the degree of "similarity" in the internal "structures" of the two tokens involved.

**Theoretical Problems in Artificial Intelligence**

But nothing further is presumed regarding the structures of tokens. In particular it should be noted that these postulates do not imply any transitive relations between internal structures, given three tokens $m_1, m_2, m_3$ and the statement that $b_{12} = b_{23}$, there are no grounds for predicting the value of $b_{13}$.

Although it is tempting to assume that the relations of similarity, however defined, must be symmetrical, we leave that an open question here because a satisfactory discussion would carry us too far from the main line of development. Thus it is assumed that $b_{ij} = b_{ji}$ is possible. Similarly, postulate P6 does not entail the symmetric growth of external bonds. If $m_i$ and $m_j$ are both members of $A$ over a particular time interval, then both $a_{ij}$ and $a_{ji}$ will be increased in value, but it is possible that $\delta_{ij} = \delta_{ji}$. The question of symmetric versus asymmetric growth of external bonds will be taken up in the discussion below.

**Discussion of the Six-Postulate Machine**

As more postulates are introduced, the possible varieties of behavior become ever more numerous and their analysis more difficult. Since we wish to add several more postulates in this paper, it will be necessary to hold the discussion of specific examples to a minimum. In this section two main behavioral issues will be considered.

First, the most apparent effect of P5 and P6 is the progressive elimination of stochastic behavior in the abstract machine as time passes. Consider the fragmentary network of couplings in Figure 7. Assume that the machine has just been born so that $c_{12} = b_{12}$ and the bond coefficients indicated are therefore equal to the internal coefficients. Let $\lambda = 2$ and $A = [m_1, m_2]$ with $m_1$ being the oldest. At this time the addition set $F = [m_3, m_4]$ because $f_3 = f_4 = 1$, and there are no other more strongly coupled successors of $m_2$. The next token to enter $A$ will be chosen at random from $F$; suppose $m_4$ is selected. After $t$, $m_4$ enters $A$ and the external coefficients $a_{24}$ and $a_{42}$ will be increased by some finite increments.

Suppose that at some later time $m_i$ and $m_j$ again appear in $A$ as a result of the coupling pattern for tokens not shown. Now the addition set $F$ contains but one member, $m_4$. As a result of the increase in $a_{24}$ on the previous pass, $f_2 > f_3$, and there will be no random choice this time. The $A$-sequence will again be $m_1, m_2, m_4, \ldots$, and $a_{24}$ will be increased again, giving $m_4$ a still greater advantage over $m_1$. Thus the selection of a token to follow $m_2$ into $A$ was initially a stochastic process but has quickly become deterministic. The immediate and obvious conclusion we draw from this little experiment is that the growth of external bonds eliminates stochastic behavior that may have been implicit in the initial coupling matrix.

But suppose that at a still later time $m_i$ and $m_j$ appear in $A$. If the growth of $a_{24}$ has been sufficiently small, so that $c_{24} < \lambda$, then the
A-sequence will be \( m_1, m_2, m_3, \ldots \), and \( a_{ij} \) will be increased. A repetition of this sequence would result in \( a_{23} = a_{24} \) and therefore \( c_{23} = c_{24} \). Now if \( m_1 \) and \( m_2 \) again appear in \( A \), \( f_3 = f_4 \) and a random choice between \( m_3 \) and \( m_4 \) is required. Thus it is clear that the sequencing of memory tokens in a particular group may be stochastic, then deterministic, then stochastic, and so on indefinitely. A change in \( \lambda \) is not required (as in the four-postulate machine) to shift the machine's behavior from stochastic to deterministic behavior or vice versa. The rule that stochastic behavior is progressively eliminated by the growth of external bonds must hold only for a class of initial coupling matrices. Unfortunately, I don't know what criteria define this class of matrices although small-scale experiments indicate that most simple strings, rings, and trees are members. This problem is interesting but may not be very important; for when sensory inputs are introduced in even a rudimentary form, the situation becomes immensely more complicated. For one thing, it becomes both possible and desirable to assume that initially there are no memory tokens in the machine (and hence no initial coupling matrix), and that the continuous creation of memory tokens via sensory inputs produces a growing coupling matrix.

A second behavioral problem of considerable importance is the general effect of varying \( \lambda \). In the example of Figure 6 it was seen that "remote forward bonds" can produce remarkable changes in behavior as a function of \( \lambda \). As a direct consequence of \( P_6 \), such bonds (external type) are regularly produced when \( \lambda > 2 \). In order to examine this effect, it will be useful to introduce some auxiliary, temporary postulates governing the increments \( \delta_{ij} \) in \( P_6 \).

First we note that by \( P_4 \), memory tokens enter the attention register one at a time, every \( \Delta t \) seconds. Thus the members of \( A \) can always be ordered by the times of entry; and given any two members, one is necessarily "older" than the other. We have been assuming all along that once the attention register has become filled, the entry of a new memory token causes the "oldest" token to be ejected. Now let us postulate that \( a_{ij} \) is increased by \( \delta_{ij} \) every \( \Delta t \) seconds for which both \( m_i \) and \( m_j \) are both in the attention register. Similarly for \( a_{ij} \) and \( \delta_{ji} \), we also postulate (1) that if \( m_i \) is older than \( m_j \), then \( \delta_{ij} > \delta_{ji} \) (which results in asymmetric external bonds) and (2) that the increments \( \delta_{ij} \) and \( \delta_{ji} \) are the same for all pairs of tokens in the machine; let's say \( \delta_1 \) and \( \delta_2 \), respectively.

Consider now three tokens \( m_1, m_2, m_3 \) which, as a result of initial coupling coefficients, enter the attention register in that order. Case 1: \( \lambda = 1 \). Since no two tokens can exist in the attention register simultaneously, the external bonds between these tokens are not altered by sequencing through \( A \). Case 2: \( \lambda = 2 \). Tokens \( m_1 \) and \( m_2 \) will be co-residents for \( \Delta t \) seconds with the result that \( a_{12} \) will be increased by \( \delta_1 \) and \( a_{21} \) by \( \delta_2 \). Similarly for \( a_{23} \) and \( a_{32} \). Case 3: \( \lambda = 3 \). Now \( m_1 \) and \( m_2 \) will be co-residents for 2 \( \Delta t \) seconds, and therefore \( a_{12} \) is increased by 2 \( \delta_1 \) and \( a_{21} \) by 2 \( \delta_2 \). Likewise for \( a_{23} \) and \( a_{32} \). But \( m_1 \) and \( m_2 \) will also be co-residents for \( \Delta t \) seconds so that \( a_{12} \) will be increased by \( \delta_1 \) and \( a_{21} \) by \( \delta_2 \). Here we have the creation of remote forward and backward (external) bonds, the former being stronger than the latter. If \( \lambda \) is raised to larger values, still more remote bonds are formed.

The coupling pattern illustrated by Figure 6 now assumes a peculiar importance: such strings of external bonds are the natural result of permitting the machine to operate with \( \lambda > 2 \), regardless of the initial coupling pattern. Continued operation of the machine with large \( \lambda \)'s will simply reinforce these strings because, as demonstrated earlier, under such circumstances the sequencing process acquires a certain "inertia" which carries it through intersections without deflection. Such strings produce deterministic behavior for \( \lambda > 1 \) except perhaps at intersections of the type in Figure 7. Of course predominantly stochastic sequencing may still be obtained at any time by setting \( \lambda = 1 \), which also terminates further reinforcement of external bonds.

Discussion of the six-postulate machine cannot be concluded without a few remarks on the classical "frequency" and "recency" laws. The former stated, in effect, that the more frequently two mental entities appear together in consciousness (for whatever reasons), the stronger the association between them. Thus repetition is a major basis of learning. Clearly, the growth of external bonds, as postulated in \( P_6 \), produces such an effect. The "law of recency" is somewhat more subtle in its implications. It states that the more recently two mental entities have appeared together, the stronger the association between them. At first glance this "law" appears to require the postulation that associations decay in strength with the passage of time, in which case the law would be fundamental rather than derived from other postulates. However, it is not necessary to postulate decay of associations, at least in some cases, in order to achieve the effects described by the recency law. The example of Figure 7 shows that the growth of competing bonds can, in effect, reduce the strength of a given bond. The postulation of temporal decay in coupling coefficients remains an interesting possibility and would result in a distinct species of abstract machine. Such machines may exhibit unique properties, but we will not pursue the question here.

In closing, we note that \( P_6 \) ignores an important issue, namely, the stipulation of limits on the growth of external coefficients. The associationists certainly assumed that there were upper limits on the strengths of bonds, limits imposed by the physical properties of the nervous system. The same assumption is necessary in the case of any physically realizable machine. If no temporal decay is present, then eventually all bonds that are capable of growth would approach the same strength; and stochastic processes (or...
some substitute decision agent) would tend to dominate the machine's behavior. At least this is an obvious theorem. Of course if the growth rate of coupling coefficients is very small relative to the upper limits, then this problem might be ignored by assuming that the lifetime of the machine is too short to permit many coupling coefficients to reach their limit. However, we will not pursue this problem here for lack of space, evidence, and opinions. We are presently concerned primarily with the early growth phases of the abstract machine's life, rather than old age and senility phases.

The Nine-Postulate Machine

For logical reasons we have constructed the abstract machine from the inside. The six-postulate machine still has no contact with its environment. This is the reverse of the theory construction procedures utilized by the associationists. In developing his epistemology, Locke began with his "tabula rasa" theory which, in opposition to the then popular doctrine of innate ideas, argued that the mind is initially like a blank tablet upon which sensory inputs wrote their record. The mind gradually forms by accumulations and interactions of sensations, images, and what-not derived from sensory inputs. In their theoretical expositions the associationists typically followed this order of development, beginning with postulates concerning sensation and concluding with a description of the mind's internal machinery which ultimately results from sensory inputs. But this order is not necessary, and I find it more satisfying to begin with the internal machinery. It is largely a matter of taste, and the development of the abstract machine in this paper reflects, perhaps, prejudices gained in the computer field where one so often begins with the central processing units and leaves the question of inputs and outputs to the last (sometimes with unfortunate results).

In any event, the next group of postulates finally put the abstract machine in contact with its environment.

Postulate P7: There is a sensory register which at all times contains a set S of N sensory tokens s₁, s₂, s₃, ⋯ s₅.

Postulate P8: At any moment, each sensory token sᵢ in S is bonded to each memory token m in A with a strength given by the "discrimination coefficient" dᵢₗ; and each sensory token sᵢ in S also has a "vivacity coefficient" vᵢ.

Before introducing the ninth postulate it will be necessary to define a special function and related set. They are similar to the adduction function and set defined in connection with postulate P4.

For every sensory token sᵢ in S there is an admission function gᵢ whose value is given by

\[ gᵢ = \sum_j dᵢₗ + vᵢ \]

where j ranges over the subscripts of memory tokens in A. Thus the admission function is the sum of discrimination coefficients between the sensory token and all memory tokens in A, added to the "vivacity" of the sensory token.

The admission set G is defined as follows:

sensory token sᵢ is a member of G if, and only if, there is no other sensory token sᵢₗ such that gₗ > gᵢ and gᵢ > gₗ where "gₗ" is the admission threshold. Unlike the adduction set F, the admission set G may be empty; for the sensory token(s) with the largest admission function may fail to satisfy the criterion imposed by the admission threshold gₗ. We are now prepared to state the ninth postulate.

Postulate P9: Every Δt seconds a sensory token is chosen at random from the admission set G (unless G is empty) and entered into the attention register, causing the oldest memory token in A to be ejected from the attention register. As the sensory token enters A, it becomes a memory token (member of M).

It will be assumed that although sensory and memory tokens may enter the attention register at the same rate, namely, one per Δt seconds, these processes are not in phase. Thus a sensory and a memory token will never enter A simultaneously, and there is no need to modify the ejection rules. It is important to note that nothing has been said about the nature of sensory tokens or their period of residence in the sensory register. We will assume that a particular sensory token may enter S and, if not admitted to A, remain in S for a long period of time. And if it is admitted to A finally, it may be immediately replaced in S by another sensory token identical in structure and properties. This represents the presence of a constant stimulus condition in the environment.

Thus the machine cannot influence the appearance of sensory tokens in S, and its only control over the ejection of sensory tokens is by way of admitting them to A where they become memory tokens, i.e., permanent parts of the machine. It is the environment which determines the nature and sequencing of sensory tokens in the sensory register; the machine is free only to select, from those presented, the individuals that will be admitted. If the machine was given motor apparatus for acting on the environment or reorienting itself relative to the environment, then of course it would gain some measure of control over the appearance of sensory tokens in S. There is insufficient space in this paper to introduce an efficient system by formal postulate, but the subject receives some discussion in the concluding section.

The discrimination coefficients represent bonds from sensory tokens to memory tokens. Their values reflect relations between the structures of tokens, not past experience, and are therefore similar to the internal coefficients previously calculated by assuming that the lifetime of the machine is too short to permit many coupling coefficients to reach their limit. However, we will not pursue this problem here for lack of space, evidence, and opinions.
introduced. The discrimination coefficients constitute a first approximation to the effects of a pattern-recognition process, a measure of the "similarity" between sensory and memory tokens. Since the discrimination bonds are presumed to exist only between sensory tokens in S and memory tokens in A, a token in the picture in A cannot influence the admission of sensory tokens. Thus the attention register still has the pivotal role in the nine-postulate machine.

The vivacity coefficients represent an attempt to reproduce the effects of "vivacity" as postulated by associationists. They assumed that the vivacity of a sensation or percept is a function of stimulus intensity and other factors, such as object-field contrasts. Further refinements of the abstract machine would require postulates governing variables which determine the value of a vivacity coefficient (stimulus intensity, for example), but in this paper the coefficient will be treated as a parameter.

At this point in the machine's development we might well postulate a vivacity coefficient for each memory token and redefine the addiction function accordingly, for the associationists universally assumed that the objects of memory also had varying degrees of "vivacity." To be consistent with the classical theory such a postulate must eventually be introduced, but again the resulting complications are too numerous for treatment within the confines of this paper.

The introduction of the admission threshold \( \theta \) provides a distinctly new type of system parameter comparable in importance to \( \lambda \). Large values of \( \theta \) will tend to isolate the machine from its environment, while low values will permit a steady influx of sensory tokens. If, in further refinements of the abstract machine, \( \theta \) is converted to a system state variable determined by events within and without the machine, then there will be rich opportunities for subtle interactions between the machine and its environment. Eventually a and \( \theta \) (the length of the sensory register) must also be converted to system-state variables with similar consequences, as touched upon in the concluding section.

Discussion of the Nine-Postulate Machine

With the addition of the last three postulates we have two "registers," side by side as it were, with a procession of sensory tokens through one and memory tokens through the other. Depending upon discrimination and vivacity coefficients, there may be periodic transfers from S to A so that the sensory procession may influence the attention procession, but not vice versa. The events in the abstract machine now represent inherent relations between memory tokens, but will immediately begin to form such bonds with whatever other tokens are in A at the time. Since internal bonds represent inherent relations between memory tokens, when a token \( m_1 \) is created, \( b_{11} \) and \( b_{11} \) are immediately fixed for all j for all time.

What sort of memory token structures will grow up as a consequence of sensory inputs to the abstract machine? We shall direct our attention primarily to this basic problem.

As a first example, consider the fragmentary network of memory tokens. The tokens are represented by two kinds of symbols—squares and circles—to suggest similarities and dissimilarities of token structure, hence the internal bonds. Suppose that at some time in the past \( m_3 \) and \( m_4 \) were created from sensory tokens in that order. They are dissimilar in nature but were co-residents of the attention register, and an external bond has formed between them. We shall assume in the following discussion that external bonds are highly asymmetric, i.e., that \( \theta_1 > \theta_2 \), so that for all practical purposes we can assume a single external bond from the older to the
Now suppose that at some later time another sensory token is admitted to A, becoming m1, and due to a great similarity in structure (detected perhaps by a pattern-recognition system) there are large internal coefficients for the pair m1, m2. As a consequence, m2 gains entry to the attention register, say A>4, and let m2 be created At seconds after m1. Now we have m1, m2, m3 in the attention register. Due to previous co-residence m2>0 and due to similarity of structure b34>0. Acting together it is likely that m4 will next be entered into A. Now the attention register contains m1, m2, m3, m4. In addition to the internal bonds between m1, m2 and between m3, m4, external bonds from m1 to m2 and m4, and from m3 to m4 will be formed, while the external bond from m1 to m4 is reinforced; and external bonds between m1, m2 and m3, m4 will form to give extra support to the internal bonds. As a result, the coupling pattern of Figure 8 will appear.

First we note that if, as we have supposed above, newly created memory tokens are able to dominate the selection of memory tokens in the normal A-sequence, then internal coefficients will be acting decisively; and there will be a tendency to form "clusters" of tokens. The pairs m1, m2 and m3, m4 are examples of the simplest possible clusters. Furthermore, if events in the environment are such that two or more types of memory token are repeatedly created in a certain order via the sensory register, e.g., m1 followed by m2, and m3 followed by m4, then strings of clusters will be created. The coupling within a cluster, being a combination of both external and internal bonds, will be stronger than the coupling between clusters, which we assume will be almost entirely the result of external bond formation.

A string of clusters might also be viewed as a bundle of parallel strings with numerous cross-bonds, but in any event, it is clear that the sort of strings and chains assumed in our analysis of the more primitive four and six-postulate machine can be created in profusion by sensory inputs. It should be noted that if A is large, say, \( A > 5 \), then it will be less likely that sensory tokens disrupt the normal sequencing of memory tokens unless the machine's memory tokens are sparsely bonded into simple strings with few remote forward or backward bonds. A very youthful machine will, in fact, tend to have only sparsely bonded strings and hence will be strongly influenced by sensory inputs even if \( A \) is large. But as the machine grows older and the memory tokens become densely bonded, it will be decreasingly influenced by its environment; indeed whenever \( A \) is reasonably large, an old machine will tend to select only those sensory inputs which fit in with its experience (internal and environmental). Thus it appears that as the machine grows older it will display a type of rigidity in its behavior (relative to the environment) which is quite distinct from the tendency to deterministic sequencing discussed earlier.

In Figure 9 we suppose that at some time in the past m3 and m4 were created in sequence and are externally bonded. Now m1 is created and due to a strong internal bond drawn m1 into A. But the creation of m1 is followed by the creation of m2, which is quite different from m3, m4, and m5. An external bond is formed from m2 to m5, and it now becomes strong as the old bond from m2 to m4. We have in this process a basic mechanism for the creation of...
"tree" structures in the coupling matrix. If in the future \( m_2 \) enters \( A \), then there may follow a random choice between \( m_2 \) and \( m_4 \), unless, of course, a sensory token is meanwhile admitted and turns out to have a strong internal bond with \( m_2 \) or \( m_4 \), thereby determining the \( A \)-sequence. On the other hand if, say, \( c_{24} > c_{23} \), the \( A \)-sequence is deterministic unless a sensory token is meanwhile admitted and creates a memory token with a strong internal bond to \( m_3 \). It may be then that \( f_3 \equiv f_4 \), and the \( A \)-sequence becomes stochastic rather than deterministic.

In brief, it is clear that not only can tree patterns of coupling be easily produced, but that further sensory inputs may either cause or prevent stochastic sequencing in such tree patterns. Having shown how branches are formed, the equally important formation of merging strings must be demonstrated.

Consider Figure 10. At some time in the past, \( m_2 \) and \( m_4 \) were created in sequence and are accordingly bonded externally. At a later time, \( m_1 \) is created, followed by the creation of \( m_3 \), which has a strong internal bond to \( m_4 \). Assuming \( A > 2 \), and that \( m_1 \) causes \( m_2 \) to enter \( A \), an external bond from \( m_3 \) to \( m_4 \) is formed. Hence \( m_2 \) becomes a point at which strings merge. In the future, both \( m_1 \) and \( m_2 \) may be followed by the same token, \( m_4 \).

Since the machine can produce strings, branching strings, and merging strings, it appears that any conceivable finite coupling matrix can also be created given appropriate sensory inputs in the past. But this is not a proven theorem. It is an extremely important theorem because, if true, then whenever one finds a coupling matrix which produces a desirable or interesting behavior pattern (with or without further sensory inputs assumed), it is guaranteed that there exists at least one possible history of the machine which would lead to the coupling matrix in question. My intuition tells me that the theorem must be true, but it also seems quite likely that new behavioral principles will emerge if machines with large numbers of memory tokens (say 100,000) can be examined in detail. It is conceivable that really large token ensembles, given the postulates introduced here, have a class of "forbidden" states, i.e., forbidden coupling matrices. I have no proof, and of course the classical associationists never pursued their theory in sufficient detail to even recognize the existence of this problem.

A less serious problem arises from the likely formation of clusters of memory tokens connected by strong internal bonds (as in the first example above, Figure 8). It appears that if the number of tokens in a cluster exceeds \( A \), the machine would tend to be trapped, to cycle endlessly within the cluster. This is another problem not recognized by the associationists. It would be desirable to have just one or two tokens from a cluster enter \( A \), "stand for" the whole cluster as it were, and then have the \( A \)-sequence move on to the next cluster. This seems to have been assumed by the associationists; but it is clear, within the framework of our present machine, that a special kind of postulate must be added in order to produce the desired behavior. It will be touched upon in the concluding section.

There are numerous other effects, such as conditioning and gestalt responses, which can be demonstrated, at least in a rudimentary form, to be possible in the nine-postulate machine. Some of these were discussed briefly in connection with the behavior of the four- and six-postulate machines and can be easily extended to the nine-postulate case.

To summarize, the introduction of sensory inputs in even the crude form of \( P_7, P_8, \) and \( P_9 \) yields a machine which grows with the passage of time. It is not "preprogrammed" in the usual sense, having just two "registers," the ability to create permanent memory tokens from transient sensory inputs, and two kinds of coupling between tokens --- one that suggests physical links (external bonds) and one that suggests the existence of a pattern-recognition apparatus or "resonance" phenomenon (internal bonds). The coupling patterns which grow within the machine reflect the pattern of events in the environment. As the machine grows older, these patterns may become self-reinforcing by selecting from the sensory input only those tokens which match themselves. This tendency may be disrupted by the appearance of sensory tokens with relatively large vivacity coefficients (e.g., due to strong stimuli) or by a drastic reduction in the length of the attention register (in psychological terms, a state approximating that achieved in "free association").

If the sequencing of memory tokens into the attention register proceeds at a somewhat faster rate than the admission of sensory tokens, it will, in effect, be continually "predicting" future events in the environment on the basis of past experience. By this I mean that a sequence of the sort obtainable from the group in Figure 8, for example, can lead to a prediction: the \( A \)-sequence \( m_1 \), \( m_2 \), \( m_3 \) can bring in \( m_4 \), which is a "prediction" of the type of sensory token that will appear in \( S \). But of course the machine has the nasty tendency to select for admission those sensory tokens which satisfy its own predictions --- a habit not entirely unknown in humans and perhaps exemplified by this paper.

Conclusion and Glimpses of Further Postulates

It must be emphasized that the nine-postulate machine constitutes neither a complete nor unique interpretation of classical association psychology. The vagueness of the associationists' doctrine and the variations from one author to the next suggest a variety of interpretations. We have presented here but one partial interpretation.

The analysis of the abstract machine's behavior has been based on small groups of tokens. However, the association theories were intended to
explain the behavior of the human mind, a very large collection of objects to say the least. The major problem posed by the abstract machine is, similarly, its behavior when large numbers of memory tokens — say $10^5$ to $10^7$ — have been accumulated and are interacting with each other and with environmental inputs and outputs. Small-group analysis can at best suggest general conclusions regarding large group effects, and the arguments in this paper are only tentative. We are in dire need of techniques for analyzing massive collections of tokens.

From the standpoint of artificial intelligence, the most interesting aspect of association theory is its purported ability to account for the growth of mechanisms that can "think," a growth that begins with nothing more than a few forces operating on the products of sensory inputs — Locke's "tabula rasa" concept. If the associationists' claims are accurate, then they have, in fact, invented a machine which structures itself without the aid of "pregrogramming." But their claims have never been confirmed experimentally (in humans) nor have the logical consequences of their axioms been explored in any detail. Abstract machines of the sort discussed here provide a vehicle for such explorations providing that sufficiently powerful analytical techniques can also be evolved.

Computer simulation seems to offer, at present, the most promising approach to analyzing the abstract machine, although mathematical analysis has not yet been given a fair testing. Whether or not the complex combinatorial problems, which arise when a reasonably complete machine has grown beyond several hundred tokens, can be successfully attacked by conventional mathematical techniques remains an open question awaiting the attention of suitably talented mathematicians. However, the abstract machine has been constructed in such a way that it can be simulated — though the cost might be forbidding — on a general-purpose digital computer.

In constructing the abstract machine, the most difficult decision is confronted in formulating a mechanism for choosing among tokens which have equal claims to succession. This problem is inherent in any quantified associationist system where objects are competing for attention. The associationists never faced this problem directly, preferring to leave a certain loophole through which a free-will soul might be slipped should the theory threaten to become too successful.

We have here postulated a selection mechanism which utilizes random choices only when two or more tokens have exactly equal claims on the machine's attention. This problem is inherent in any quantified associationist system where objects are competing for attention. The selectionists never faced this problem directly, preferring to leave a certain loophole through which a free-will soul might be slipped should the theory threaten to become too successful.

We have here postulated a selection mechanism which utilizes random choices only when two or more tokens have exactly equal claims on the machine's attention. This problem is inherent in any quantified associationist system where objects are competing for attention. The associationists never faced this problem directly, preferring to leave a certain loophole through which a free-will soul might be slipped should the theory threaten to become too successful.

In constructing the abstract machine, the most difficult decision is confronted in formulating a mechanism for choosing among tokens which have equal claims to succession. This problem is inherent in any quantified associationist system where objects are competing for attention. The associationists never faced this problem directly, preferring to leave a certain loophole through which a free-will soul might be slipped should the theory threaten to become too successful.

We have here postulated a selection mechanism which utilizes random choices only when two or more tokens have exactly equal claims on the machine's attention. This problem is inherent in any quantified associationist system where objects are competing for attention. The associationists never faced this problem directly, preferring to leave a certain loophole through which a free-will soul might be slipped should the theory threaten to become too successful.

We have here postulated a selection mechanism which utilizes random choices only when two or more tokens have exactly equal claims on the machine's attention. This problem is inherent in any quantified associationist system where objects are competing for attention. The associationists never faced this problem directly, preferring to leave a certain loophole through which a free-will soul might be slipped should the theory threaten to become too successful.
tokens within a node be introduced. These coefficients should be partially responsible for determining vivacity coefficients, and for admission functions connected with the multilevel attention register discussed below.

Third, all memory tokens must be assigned vivacity coefficients whose values are functions of those accompanying the sensory tokens from which these tokens were created. The addition function must be appropriately redefined.

Fourth, an entirely new type of coupling must be introduced. In this paper we have considered only token-token coupling, but token-system coupling is required to represent the influence of general 'system states' upon the machine's behavior. For example, we postulate a system state parameter $P$ which ranges between positive and negative limits: call it the "Pleasure-Pain" scale. When memory token $m_1$ is created, its pleasure-pain coefficient $P_{m_1}$ is set equal to the current value of $P$. The addition function is redefined to $E_j = \sum_i c_{ij} + P_{m_i}$ when $P$ and $p_i$ are of the same sign and relatively large, the state of the system as a whole can be responsible for the selection of $m_i$ for entry into $A$. The next step is to convert $P$ to a state variable whose value is determined, partially at least, by certain sensory tokens. Further refinements to this token-system coupling scheme carry one well beyond the classical theory.

Fifth, several levels of attention registers must be postulated. As Bain put it in a poetic outburst: "No associating link can be forged... except in the fire of consciousness; and the rapidity of the operation depends on the intensity of the glow." He was referring here only to those types of bond which can be reinforced; in our machine, external bonds. It was generally recognized that there are degrees of consciousness, and we interpret this as discrete attention registers. Let us say that there are $n$ attention registers of lengths $A_1, A_2, \ldots, A_n$ which hold sets $A_1, A_2, \ldots, A_n$. Each attention register is assigned a "level coefficient" $k$ with $k_1 < k_2 < \ldots < k_n$. These may be thought of as "energy levels," and the addition and admission functions as measures of energy available to the corresponding tokens. The level coefficient of the set $M$ is assumed zero, and the admission threshold becomes the difference between the levels of $S$ and $A_1$. Figure 11 shows an example of the postulated relations between levels of $S$ and $A_1$. Figure 11 shows an example of the postulated relations between levels of $M$ and $S$, and the various $A_i$ sets. The aduction set $F$ must be redefined to include a comparison of addition functions with a set of addition thresholds similar in effect to $\theta_g$. In Figure 11 there are four sets of transitions labeled a, b, c, and d. The first shows the types of transition involved in the simple transfer of a sensory token from $S$ to $A_1$, where it becomes a memory token, and after brief residence in $A_1$, ejection. These are the kinds of transition discussed in connection with the nine-postulate machine. Set b shows the entry of a memory token directly into $A_2$, its drop to $A_1$, after a brief period, then finally its ejection. Transition c is forbidden by postulate. Set d shows how a token may move among the various attention registers for a considerable period of time before ejection.

As suggested by Bain, the increment in external bonds between co-residents within any given attention register is an increasing function of the level of that register. Thus external coefficients for co-residents in $A_2$ would increase at a greater rate than those for co-residents in $A_1$. Clearly multilevel attention registers create rich opportunities for multiple sequencing and the formation of complex coupling patterns.

Having once connected the notion of "energy level" with the level of a register, it is natural to consider conservation postulates. Stout's theorizing leads to some rather exciting possibilities in this regard. A conservation postulate that suggests itself here is $\Sigma_k A_k = \text{constant}$. This in turn suggests a postulate, not unreasonable in a physiological interpretation of association theory, that similar tokens are stored in similar places (to put it crudely) and that the energy supplied to any particular neighborhood is limited. Thus when a token is once raised to a particular energy level (i.e., attention register), the aduction functions of similar tokens (as measured by internal coefficients) are correspondingly reduced, thereby lessening the tendency for endless cycling within clusters of similar tokens --- a problem mentioned earlier.

These examples of further postulates give some indication of the incompleteness of the nine-postulate machine. And if one attempts to extend the abstract machine much beyond the explicit and implicit assumptions of classical theory, the nine-postulate machine constitutes only the barest beginning. Consider the problem posed by the creation and use of linguistic symbols in goal-directed thinking. The postulates presented or suggested here appear quite inadequate to produce such behavior. If linguistic symbols are always assumed to be constructed from stimulus patterns (mainly visual and auditory), then of course the nine-postulate machine, not to mention the more complex machines produced by the suggested additional postulates, would be capable of learning to associate such symbols with clusters of tokens which constitute the internal denominations. But this capability is still far removed from goal-directed thinking.

There are hints in the literature, particularly in Taine, which when followed up may lead to sufficiently powerful postulates that are not too question-begging and are consistent with the general approach of association theory. However, we can only conclude here that the postulates so far considered could produce at best a machine having a level of intelligence comparable perhaps to that of the lower vertebrates. And one might understandably take issue with this tentative conclusion, for the nature of the sensory and memory tokens have not been stipulated in sufficient
detail for comparisons with animal behavior, nor have we carried out methodical experiments on reasonably large abstract machines.

Acknowledgments

We are pleased to acknowledge the support of the U.S. Air Force Office of Scientific Research (Contract No. AF 49(638)-1039) in carrying out the research upon which this paper is based.

References

Figure 1

Figure 2

Figure 3
Figure 4

Figure 5