A COMPUTER SIMULATION MODEL
OF ATTACK LEARNING BEHAVIOR IN THE OCTOPUS

John Crittenden Clymer

January 1973

ABSTRACT

A COMPUTER SIMULATION MODEL
OF ATTACK LEARNING BEHAVIOR IN THE OCTOPUS

by

John Crittenden Clymer

Chairman: Larry K. Flanigan

A quantitative model of attack learning is developed in the dissertation based on J.Z. Young's theory of memory in the octopus brain. This theory suggests a binary neural element (the "mnemon") as the basis for memory, and says that circuits located in the animal's upper lobe structures may function to write into and read out of these mnemons. The model is subsequently used to generate learning curves, which compare favorably with behavioral data obtained from experiments with real animals. The dissertation then looks at effects in the model analogous to operations in which the upper lobe structures are damaged or removed, and at other degradations in performance caused by interference. Certain pattern recognition aspects and problems of design are pointed out, and directions for additional studies are suggested.

ACKNOWLEDGEMENTS

Special appreciation is due the members of my doctoral committee:
Professor J.H. Holland, Assistant Professor S.P. Hubbell,
Professor H.H. Swain, and to the chairman, Associate Professor
L.K. Flanigan, in particular. Each provided me with guidance and
encouragement throughout this work, and with many helpful suggestions
in the preparation of the dissertation. I would especially like to
thank Assistant Professor Hubbell for the extra time he devoted to
me in many such discussions.

I would also like to extend special thanks to Professor J.Z. Young
of University College London, and to the *Stazione Zoologica* in Naples
for allowing me to observe octopus training experiments firsthand during
the summer of 1971. Considering the time pressure of that work and
the limited facilities available, permission to be there during these
experiments was greatly appreciated. Dr. R.O. Stephen of University
College also deserves recognition for his valuable suggestions concerning
the model.

My debt of gratitude must then be acknowledged to Jan McDougall
for her excellent secretarial assistance in producing the manuscript,
and to Monna Whipp who typed many of the chapters. This assistance
in typing, editing, and drawing the figures removed a great burden,
and allowed more time to be devoted to experiments with the model.

Next, I would like to express appreciation to many friends who
provided advice and encouragement along the way, and especially to
my wife, Joanne.

Finally, I must acknowledge my debt to Professor Young for the
inspiration of this work in his writings. Its shortcomings and
inadequacies are my own responsibility, but his work provided impetus
for the undertaking in the first place.

TABLE OF CONTENTS

TABLE OF CONTENTS

(Cont'd)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

(Cont.)

CHAPTER ONE

INTRODUCTION

1. Biological Background

Biological systems can be characterized as mechanisms which maintain themselves by adapting to their environment. In viewing biological systems this way, we are interested in how this adaptation is accomplished, both in the short term life of a single plant or animal, and in the long term life of an entire species. Short term adaptation involves homeostasis and learning. Long term adaptation involves the action of genetic mechanisms to preserve the coded instructions for producing successful individuals.

Animals are provided with a spectrum of adaptive mechanisms ranging from simple reflexes, tropisms, and instinctive behavior, through various degrees of true learning capability. These mechanisms enable the individual animal to survive an increasing range of environmental situations. There are cases in which learning may prove maladaptive, but the ability to react in less stereotyped fashion generally has survival value to an animal faced with a complex environment in which it must find food and avoid danger. This advantage must be paid for by increased dependency of the young during training, however, and longer periods before maturity is reached. There is thus a heavier investment in the survival of the individual as the complexity of the individual increases, but the individual becomes increasingly able to discriminate between and react to "good" situations and "bad" ones, and hence more likely to survive to propagate.

This increased learning capability is related to an increasing complexity of the nervous system, and of the brain, in particular. Perhaps the most challenging problem on the frontiers of modern science

is to understand how brains work. The direct study of brains and nervous systems faces many difficulties. The neurons of which they are composed are exceedingly small, delicate, numerous and connected in complex tangles. They can be viewed directly only after some type of staining operation. Classification systems imposed upon them are largely the subjective constructs of investigators due to the relatively continuous variability observed in the sizes and shapes of the neural cells. The tools at hand of micro-electrode recording, surgical lesions, and chemical analysis are still relatively crude. We have made progress in understanding the action of single nerve cells and of some small assemblies of nerve cells, but the complexity of even very small brains makes the task of understanding their operation very difficult. Their gross anatomy is often relatively simple, but so little is really known about the connectivity and timing relationships between elements that any theory of operation for something as complex as the human brain seems a long way off.

The ideal of relating structure to function can be achieved to some extent in certain areas of the nervous system. In other areas it is next to impossible even to determine what the structure is. The most promising approach seems to involve applying hypotheses of what to look for based on biological knowledge of the animal and its environment. Horridge concludes, for example, that "The problems of *what one should look for* when studying interneurons is thrown into the lap of the biologist [who] will know the special features of the environmental stimuli of interest to his subjects, and he is wise who studies specialized animals to which certain stimulus situations have great survival value."[1] Young makes a similar point: "If we are to understand how a brain works we must first answer the rather obvious question 'what

does it do'"[2]

The approach taken by Young and his associates is to study an animal of intermediate complexity in some detail. For this purpose they have chosen   *Octopus vulgaris* Lamarck, which is plentiful near their Naples laboratory. Young and his colleagues have used the octopus in numerous conditioning experiments, and have traced a good many neural connections through anatomical studies. They have also run comparative behavioral studies with lesions in parts of the nervous system, and have studied the effect on behavior of removing various lobes of the brain.

Despite its fairly simple nervous system, the octopus exhibits a surprisingly complex range of behavior. For many cases of interest, however, it can be considered to react to stimuli in only two ways-- by attacking or by retreating. The animal normally sits in its "home" among the rocks watching for a food object to pass nearby. Its visual system is more rudimentary than that of vertebrate systems, but it can discriminate among a wide range of forms. The horizontal and vertical direction seem to play a special role in this discrimination. Thus it easily distinguishes horizontal from vertical rectangles, but has difficulty in distinguishing a square from a circle. Each of its eight arms can also be considered a "distance receptor", in the sense that the animal can obtain tactile information on objects within its reach. In this regard, it is interesting that there are tactile centers in the central nervous system which appear to be functionally quite similar to the higher centers of the visual system. In summary, "The nervous system [of the octopus] serves to make decisions between a relatively small number of things the animal can do. In particular, a decision must be made whether to attack or retreat, to seize an

object and bite it, or to reject it by pushing and blowing it away.
Of course the nervous system must control other detailed processes ...
But the main central nervous system is concerned with deciding whether
to advance or retreat, and over 90 per cent of the central neurons
are concerned with this decision. It is this relative simplicity of
the action system that makes these animals so suitable for work on
the coding and learning mechanisms."[3]

The importance of this binary decision to the animal, as reflected
in the structure of its nervous system, leads to the belief that some
type of functional model might capture a significant portion of the
information processing going on in these centers, and that such a model
would provide some insight into the difficulties to be encountered in
designing such a system. That is to say, given the task of designing
the decision apparatus for an intelligent automaton to occupy the ecolog-
ical niche of the octopus, what design problems might one encounter,
and what solutions to these problems would perhaps suggest themselves?

As Young puts it, "In particular we are interested in a self-teach-
ing homeostat, that is to say, one whose information and instructions
are not entirely built in by heredity. Such a learning device alters its
performance as a result of experience so as to produce results that are
satisfactory for its self-maintenance. In order to do this it must
record certain features of the input in a code, recording also whether
they were accompanied by conditions that were 'good' or 'bad' for the
homeostat. We shall try to study the principles by which such a system
uses 'rewards' and 'punishments' to control its performance and how it
finds the optimal rate of change of behavior as it learns in any given
environment."[4]

A model of such a system reflects at a functional level some theory

of operation, in the sense that a theory concerning the internal

structure of the octopus nervous system can be reduced to a set of

model equations and programmed for a digital computer. The parameters

of such a model can then be varied so as to produce model behavior

patterns comparable to experimental behavior, and the implications

of the theory can be shown. The model can thus be considered as a

"black box" which produces behavior similar to that of the animal.

Chapter 4 will describe a learning model I have developed based on

Young's "mnemon" theory and what is known about the octopus. The

following section gives some of the background for this model.

2. General Background of the Model

In a very general sense, biological systems can be said to perform

a mapping from the space of all possible stimulus situations to the

space of all possible alternative responses, as illustrated in figure

1.1. That is, the system adjusts itself in response to its perception

of the state of the world.

Figure 1.2 shows a schema for how this process takes place in

the higher animals. The process first involves an encoding of the

environment via several sensory modalities and possible pre-processing

of this sensory information before it reaches the decision making

centers of the brain. For example, Hubel and Wiesel (1962) found in

their work on cats that the visual image projected on the retinal rods

and cones was converted before reaching the cortex into small circular

excitatory fields with concentric inhibitory surroundings, or vice-versa.

After pre-processing, however, they found that in cortex cells these

fields had been converted into several quite different types of

fields, such as line or edge detectors with specific orientations.

Thus in the most general case, peripheral and cerebral pre-processing

SPACE OF
RESPONSES

BIOLOGICAL
SYSTEM

SPACE OF
STIMULUS
SITUATIONS
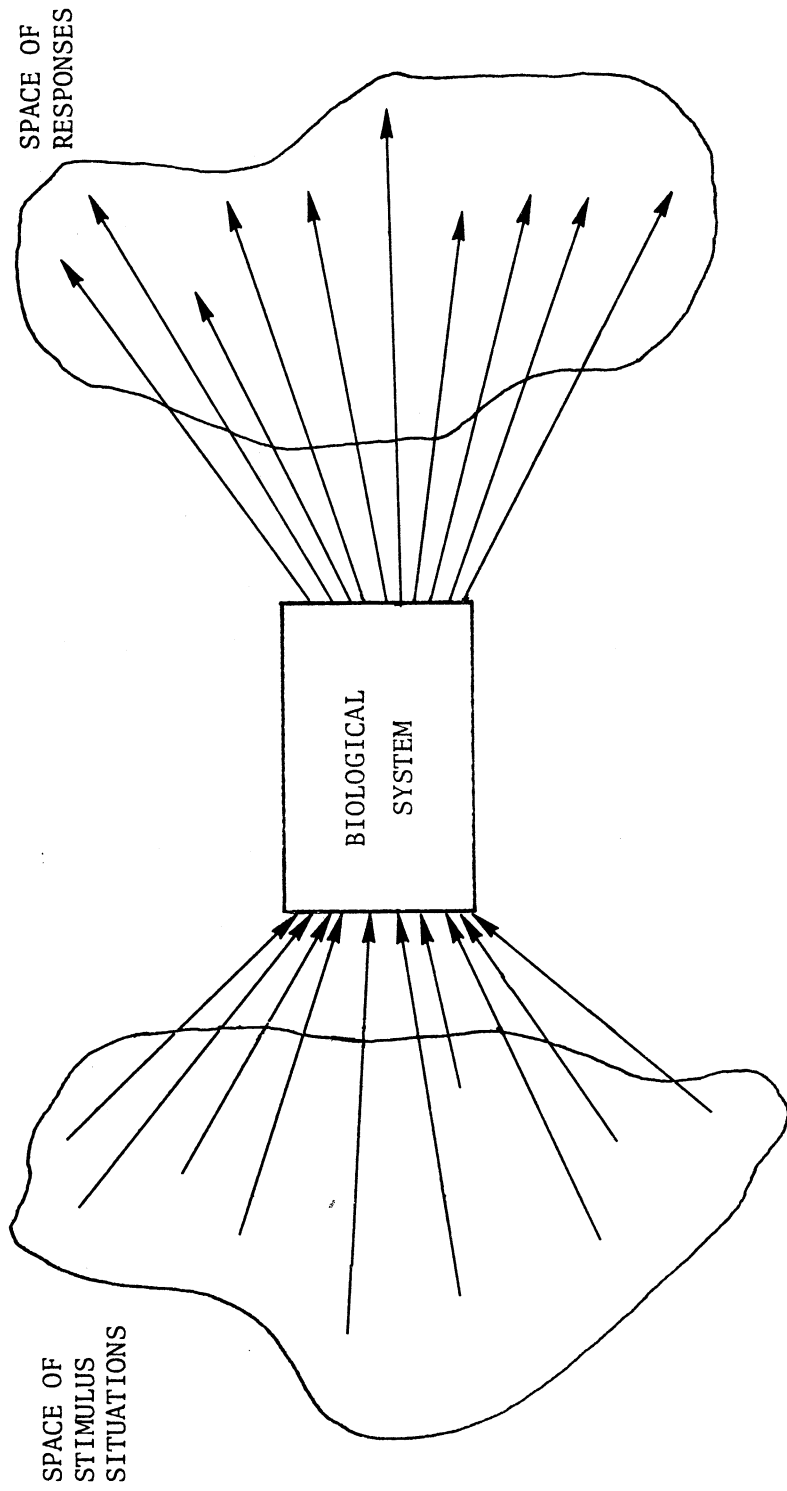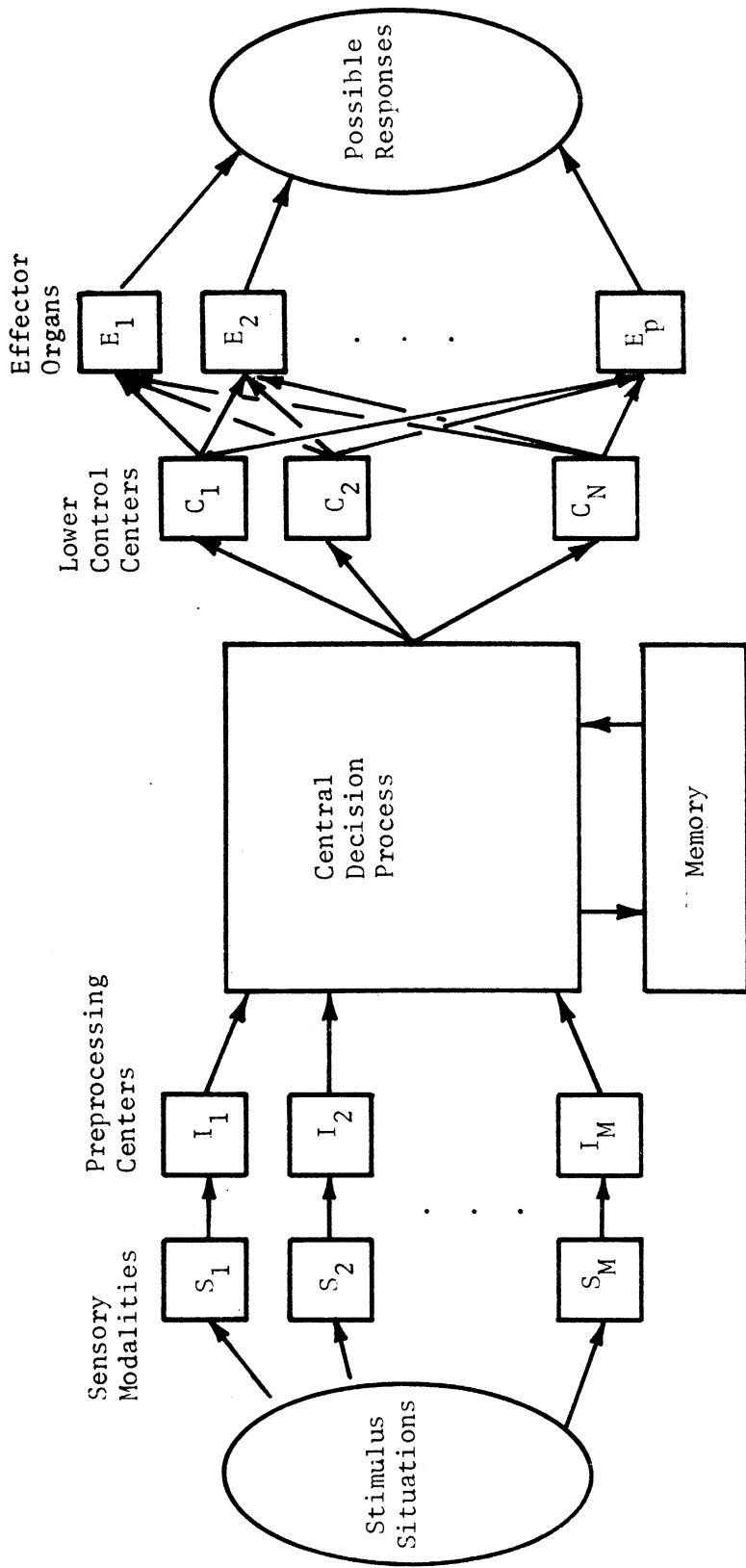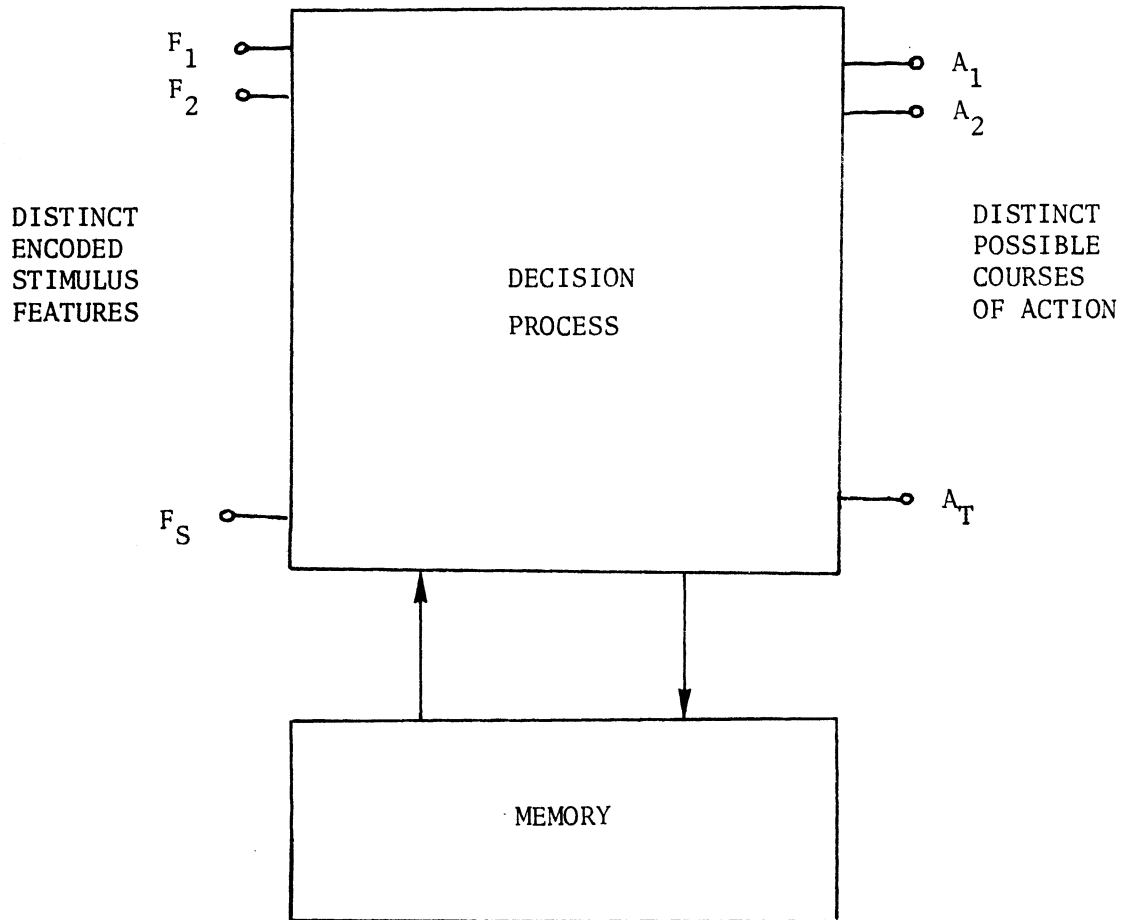
Figure 1.1

Figure 1.2

Figure 1.3

results in two sets of transformations being applied to each modality of sensory information used as input to the decision centers of the brain. With this input from all the sensory modalities and the memory of past experiences, a decision is arrived at by the central process from its repertoire of possible actions. This decision is passed on to the lower control centers where it is elaborated into sequences of instructions to the effector organs themselves and becomes a response.

The model discussed here will concern itself only with the center boxes of figure 1.2. It will be assumed that sensory information has been pre-processed and that output decisions can be converted by lower centers into effective actions. That is, the decision box can receive any combination of inputs from among all the possible stimulus features of the various modalities, and on the basis of this input and memory contents, it chooses one of the courses of actions open to it. (Refer to figure 1.3) The decision process is thus a problem in pattern recognition, but there are additional factors to consider.

In all real systems, memory contents are not static but change with experience. Moreover, evidence indicates that more than one level of memory is at work. After surveying numerous experiments, Horridge distinguishes three periods of time in memory experiments with mammals: "First, an initial phase of seconds or a few minutes, sensitive to electro-convulsive shock and analeptic drugs (strychnine) but not to antibiotic drugs. A second phase, of a few minutes, is inhibited by temporal lobe injections of puromycin in mice, but is not sensitive to actinomycin or acetoxycycloheximide. This suggests that growth or protein synthesis is involved but there is no direct evidence as yet of the growth of nerve terminals. In the third period the memory is in a more permanent form and is insensitive to cooling,

drugs, shocks or anything except further training. This can be inter-
preted as a new distribution of synaptic effects, with neurons acting
as if they have new connections."[5]

Wooldridge also concludes there are different memory mechanisms:
"Thus a combination of subjective impressions and objective observations
has led us to a three-stage concept of memory. The brain automatically
preserves for a brief interval the description of the immediate external
environment provided for by the nerve impulses coming in over the external
neuronal receptors. In the absence of selection and reinforcing activi-
ties, this input information fades away quickly--probably within a few
seconds. Normally, however, partly consciously and partly unconsciously,
an attention-focusing mechanism, which may be the reticular activating
system, performs an initial sorting operation on the incoming data, some
of which is earmarked as being of special interest and thus is preserved
in what we call the medium-term memory. Once reinforced by this attention
focusing process, the memory trace can apparently persist in the
medium-term-memory status for a period of some minutes. In a brain incapaci-
tated by lesion, surgical excision, or electric stimulation, the memory
trace may be incapable of fixation and may be permanently lost after a
few minutes. In a normal brain, however, some kind of process goes on
by means of which the memory is made permanent. Here again, attention
focusing, perhaps by the same mechanism that selected the original
sensory data, seems to cause some recollections to go into permanent
memory in indelible fashion, and others to be stored so tenuously that
if they are not later reinforced by repetition or subsequent recall,
they may drop out of the storage system altogether."[6]

One of the problems solved by short-term memory is that of delayed
reward or punishment. After an attack has been launched, the stimulus

will be viewed from a different perspective and some of the stimulus
components which were originally present may no longer be present. Yet
something must be read into permanent memory for each of these compon-
ents when the goal is eventually reached and leads to food or pain. The
octopus will also chase a prey which has disappeared from view, so it
seems clear that some form of short-term internal representation is
set up. Differences are observed, however, between short-term learning
performance and long-term learning. Performance within learning sessions
improves rapidly and reaches a higher level than can be observed at
the beginning of the next session sometime later. In addition, animals
with lesions in the upper lobe structures can exhibit short-term learn-
ing at 4 minute intervals but not longer-term learning at 8 minute
intervals. Similar effects are observed in humans with certain types
of brain damage. Further, "After a few repeated shocks octopuses will
sometimes not attack crabs for several days. We can thus speak of
setting up some representations that prevent attack, but it is interesting
that these do not seem to destroy or replace the representations ensur-
ing attack that had been previously established. The new one seems to
be distinct from the old and if it is less well established it fades
more quickly, leaving the old tendency to attack unimpaired. This
suggests that the representations are embodied in some way in distinct
sets of cells in the nervous system."[7]

This is also suggested by reversal studies: "... if training
with two rectangles is reversed, signs of the original responses
remain apparent for a long time ... Such reversal experiments suggest
that the effect is produced not by attaching a new condition (shock)
to the old representation, but by establishing, say, a crab with shock.
Being weaker than the original one this representation soon fades and

the other is again revealed intact. ... Experiments in which the cues
were repeatedly reversed provide further insight into the problem ...
As judged by the percentage of errors the performance deteriorated to
a random level."[8] It is important for the learning model to take
account of these interacting memory mechanisms with different time
scales, therefore.

Another aspect of the model which will distinguish it from standard
pattern recognition schemes is the attempt to portray incremental
changes taking place over real time, rather than performing a simple
quantum update at each "event" or encounter with the environment. Thus,
memory is altered over a period of time rather than all at once, and
it is possible to study deficiencies in the mechanism for this reading
into memory. Similarly, the decision to attack involves a process of
build-up, and it is possible to interfere with the mechanism for reading
out of memory. Other interference effects should be seen whenever
events occur too closely together and conflict.

Finally, the model incorporates certain concepts derived from the
anatomy of the animals. Considerable study has been devoted to anatom-
ical relationships in the octopus nervous system and a good many neural
pathways have been traced. Much is thus known about this nervous
system, but a great deal is also unknown. The model is therefore con-
strained at the level of its basic structure, but not entirely defined
by our empirical knowledge. The approach will be to use what is known
in obtaining formal constructions which can be used as a basis for
simulation. A central feature of the model will be Young's concept
of bi-stable memory elements, or "mnemons" which can be switched between
the "attack" and "retreat" modes of behavior. Sensory input is assumed
to be separable into discrete components which connect into mnemons,

and behavior patterns result from recruitment of these activated mnemons. In particular, a threshold is applied to their weighted output in order to determine attack or retreat.

The philosophical viewpoint adopted in this work is that biological systems incorporate a high order of complexity, and models of these systems serve as an aid to our understanding of their operation, just as models can sometimes be used to advantage in studying other large scale systems of considerable complexity. Modeling and simulation have have long been used in technological applications and basic physical research. As our detailed knowledge of local properties in complex biological systems increases, we should also see an increasingly fruitful area for collaboration between computer scientists and biologists.

3. Summary

Nervous systems and brains have evolved which permit a wider range of adaptive behavior patterns in animals and thereby increase the chances of individual survival in complex environments. The brains of even very simple animals present a challenging subject for study, and our understanding of nervous systems is still very rudimentary. One way of viewing the function of the brain is to note that biological systems react to sensory input by selecting from a repertoire of responses. In the octopus this repertoire can be quantized into two principal components, attack or retreat. The octopus thus provides a good subject for an attempt to model some aspects of the operation of a relatively simple nervous system. A basic assumption of the model is that sensory information can be separated into a number of discrete, independent, and distinct components, or features, and that the contribution from memory for each of these features adds linearly to produce an overall response pattern. Memory has three aspects, a short-term aspect, a medium-term

aspect, and a long-term aspect, with short-term memory serving to read into and out of long-term memory through the action of medium-term memory. Reward or punishment, in the form of input from taste or pain fibers, enters short-term memory for transfer into the long-term store, assumed to consist of "mnemons." Each mnemon behaves as a bi-stable element which can be switched by the taste or pain inputs, and which retains some memory of its past states. Attack or retreat results from recruitment of large numbers of mnemons by one behavior pattern or the other. The approach taken throughout is that models developed with the aid of computer specialists can help provide insight into the operation of biological systems, just as they sometimes aid us in understanding other complex or large scale systems.

CHAPTER TWO

THE OCTOPUS

1. Genera1 Observations

The octopus is a cephalopod and exhibits perhaps the highest

degree of intelligence observed in invertebrates. It customarily

makes its home among the rocks or inside old pots found on the sea

bed. Fishermen, in fact, often catch the animals by lowering a string

of pots to the bottom for a few days, and then raising them again to

find that animals have made homes in some of the pots. They feed

chiefly on crabs and mussels, although in a laboratory situation they

will readily accept bits of fish which have been prepared as a reward

for correct responses. It is not known for sure whether this fondness

for crabs is innate or a learned behavior pattern, since *Octopus vulgaris*

is pelagic for a period after hatching and is difficult to raise from

birth.

As cold-blooded animals, octopuses are quite sensitive to water

temperature and become very sluggish in colder water. Because of their

chromatophore covering they can exhibit striking color patterns, and

these color changes often signal their intentions in the laboratory.

The ability to assume different color patterns or to "stipple" their

skin allows them the advantage of being able to camouflage themselves

or appear fearsome to an attacker, however. When frightened, they can

disappear in a cloud of ink.

The normal situation is for the animal to sit in its home until

a food object appears in its visual field, and then to dash out and

grab the crab or other food object with its arms. The crab is passed

under the mantle and carried back to the home to be eaten. Outside

the home one usually finds a scattering of shells from crabs, lamelli-

branchs, and other animals that have been eaten by the octopus. If

the object in the visual field is unfamiliar, the attack will be cautious

and slow. Later attacks will be faster and more sure if the object

yields food. If it yields a painful result, however, the octopus rapidly

learns not to attack it and retreats into the home.

In the laboratory, animals are kept in tanks with circulating

seawater which have a masonite home constructed at one end. They

can be trained to come out and attack figures cut from plastic by

providing a piece of fish for correct responses and a mild electric

shock for attacks on the wrong figure. Usually two sessions of sixteen

trials each are given per day, with positive and negative trials alter-

nating at about five minute intervals. According to Young, "The animal

learns very rapidly. Responses may begin to be correct within a first

session of sixteen trials and are sometimes quite good after the second

session. Retention is excellent for the time it has been tested, up

to five weeks. Learning can be reversed, if necessary repeatedly. In

fact the animal answers well to the questions we wish to ask."[1]

2. Nervous System

The two chief modalities of sensory input in the octopus are

visual and tactile. Chemoreceptors also supply information, but are

of lesser importance. Visual information from the retinas of its

two slit eyes is mapped directly onto the surfaces of optic lobes

through optic nerves. These optic lobe surfaces are structured much

like the vertebrate retina. Inner and outer granule cell layers are

separated by an elaborately layered plexiform zone. Inside the lobe,

the inner granule layer merges into a medulla, which also receives

input from taste and pain fibers and has its output into an optic tract.

Figure 2.1 illustrates the pathways which have been found to

VL1 - Lateral Superior Frontal $(0.08 \times 10^6$ cells$)$

VU1 - Medial Superior Frontal $(1.8 \times 10^6$ cells$)$

VU2 - Vertical $(25.1 \times 10^6$ cells$)$

VL2 - Subvertical $(0.8 \times 10^6$ cells$)$
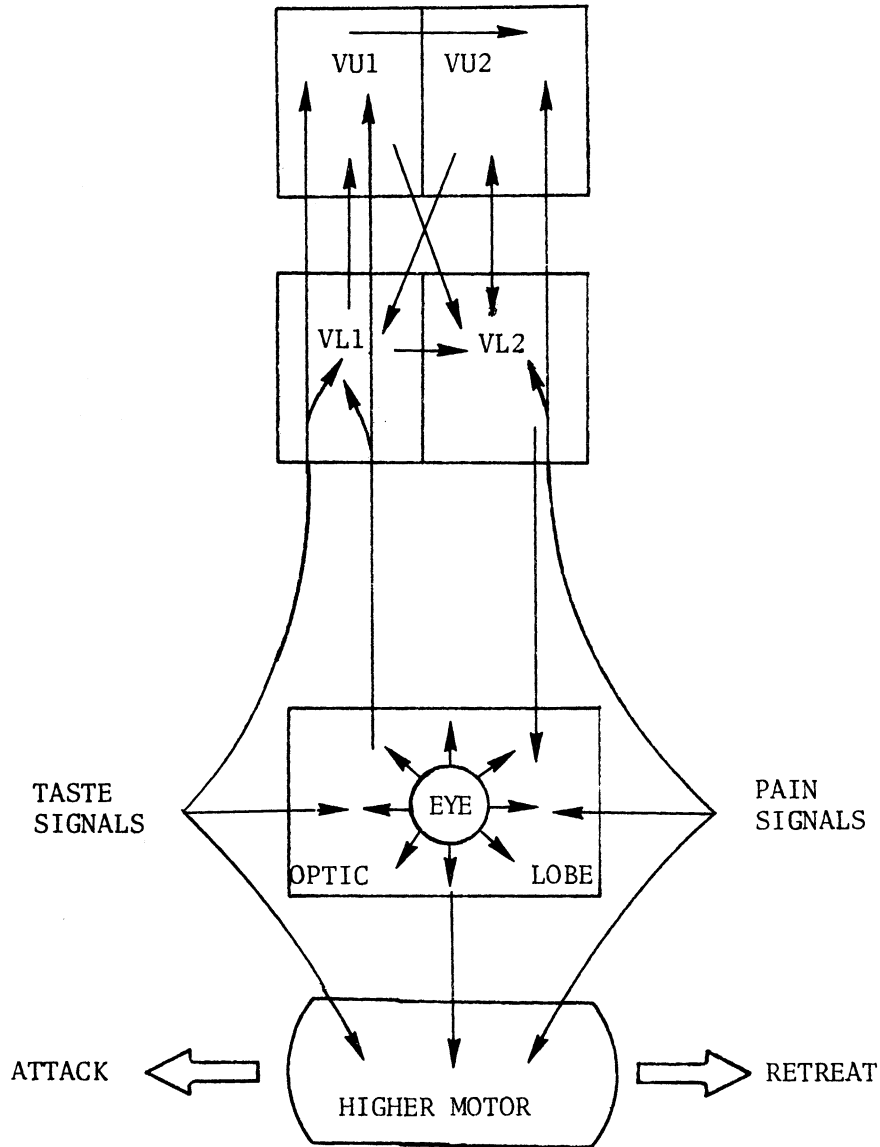
Figure 2.1

exist in the visual system (the path from VU1 to VL2 is questionable,
however). Output from the optic lobes is passed to higher motor centers
where it is elaborated into patterns of action. Optic lobe output
also goes to a set of higher centers (labeled in the figure as VL1,
VL2 for "vertical lower", and VU1, VU2 for "vertical upper"). This
vertical lobe system, which also receives inputs presumed to be taste
and pain fibers, is thought to mix and amplify signals before passing
them back into the optic lobe, so that these upper lobes perhaps function
as reverberatory pathways. In this regard they may resemble the Papez
circuit of the mammalisn limbic system. A similar set of paired centers
is also found in the tactile system. With all four upper lobes removed
the animal still puts out an arm to seize a nearby crab, but will
not follow nor launch an attack with the entire body.

The first pair of upper lobe centers (VL1 and VL2) appear to act
as an amplifying device, in the sense of increasing the effectiveness
of visual impulses in producing an attack. "Fibres from the optic lobes
pass to the lateral superior frontal [VL1]. Here they meet fibres from
the inferior frontal [tactile system], and perhaps other chemotactile
sources and a large component from the vertical [VU2] and subvertical
[VL2] lobes. These and the optic fibres make a dense plexus through
which run the axodendritic trunks of the cells of the lateral superior
frontal lobe itself. These trunks are relatively large and carry many
long receptor branches as they pass through the neuropil. These
axons then proceed medially to the subvertical lobe and end there. From
the subvertical lobe large cells send axons back to the optic lobe."[2]
This structure of the lower circuit thus suggests the function of
spreading and amplifying impulses in a relatively few optic tract axons,
unless pain intervenes in the subvertical lobe. Animals from which
the two upper lobes are removed but the two lower ones are left intact

will attack objects moving at a distance, but often slowly and irregularly compared with a normal animal.

The upper set of lobes lie superficially and are easy to operate on, so more information is available about them. "The median superior frontal [VU1] receives its input mainly (? wholly) by fibres that have passed through the lateral superior frontal [VL1] lobes. Many of the fibres of the optic to superior frontal tracts do not end in the lateral lobe but pass through its neuropil to the median lobe. Here they break up into numerous bundles, which interweave with other bundles that come from the inferior frontal lobe [tactile system] and other chemotactile pathways. The bundles divide and recombine repeatedly ... This interweaving of bundles gives us a clue to the method of functioning. The system allows fibres that were previously together to separate and to have opportunity to interact with other fibres, with which they were not previously associated. ... The median superior frontal fibres carry collateral dendrite twigs as they pass through the neuropil and are thus able to be stimulated by the combination of impulses presented by the incoming bundles."[3]

"The vertical lobe [VU2] has an entirely different structure. It consists of five cylindrical lobules, each with a very thick cortex composed mainly of masses of minute [amacrine] cells ... each having a single trunk that breaks up in the neuropil into a bush of equal branches, no one being distinguishable as an axon. The output of the lobe is produced by a small number of larger cells. Their trunks turn and twist about in the neuropil, carrying numerous dendritic side branches. Finally, these fibres leave the lobe, passing downwards in bundles to end in two destinations (1) the subvertical [VL2] and (2) lateral superior frontal [VL1] lobes."[4]

In addition, large tracts of fibers from the arms and other regions
reach the subvertical (VL2) and some of these pass up into the vertical
lobes (VU2). These are presumed to carry nocifensor (pain) fibres,
either directly from the periphery or after synapse in VL2. In the
vertical lobe these run across the incoming fibres of the superior
frontal to vertical tract.

This upper pair of lobes thus provides an additional reverberating
pathway which can be interrupted by pain. Following damage to these
lobes, animals are slow to learn to attack an unfamiliar figure or to
learn not to attack crabs, which are a familiar food object. They are
slow to learn to discriminate objects shown successively, whether by
vision or touch. They are also unable to transfer learning from one
visual field to the other. After vertical lobe removal attacks become
much slower and more variable. The vertical lobe system is therefore
concerned with producing a strong and stable level of attack.

"Animals without vertical lobes do not appear, on superficial
observation, to differ in any way from normal ones ... The lobes are
evidently not necessary for any simple motor function. This agrees
with the fact that electrical stimulation ... produces no obvious effects
when applied to the superior frontal or vertical lobes. The deficiencies
that follow vertical lobe removal become apparent in some situations
in which the animal has to learn, or to perform an action learned
before operation ... the vertical lobes are especially concerned with
preventing attack, and ... they help in some way to preserve a record
based on events of the immediate past."[5]

There is some question whether the actual memory representations
might be located in the vertical lobe, but experiments indicate that
operated animals which show little discrimination with sequentially

presented stimuli do reasonably well in discriminating stimuli shown simultaneously. This shows that some memory must remain intact following vertical lobe removal. They also do better in discriminating between figures shown sequentially but without reward or punishment. In this case, immediate effects do not overide long-term influences. "One function of the vertical lobe circuit thus seems to be to ensure a proper balance of tendencies to attack and retreat, so that the immediate effect of [food and shock] is not excessive. Only if the tendencies of the animal are balanced in this way will the behavior be controlled by the representations set up in the past, rather than by the effects of immediately preceding food or painful stimuli. We can thus say that the vertical lobe system is necessary for 'reading-out' of memory as well as for 'reading-in'."[6]

A second set of paired centers has also been found in the tactile system. The lateral inferior frontal [TL1] feeds into both the posterior buccal [TL2] and median inferior frontal [TU1]. The latter feeds into the subfrontal lobe [TU2] where fibres reach back into TL1 and TL2. Taste input goes into TL1 and TU1 whereas pain fibres enter TL2 and TU2. The situation is thus quite similar to the paired centers of the visual system.

A very large portion of the nervous system is distributed along the arms. An isolated tentacle will still make writhing movements when meat extracts are placed near it, and will pass a piece of fish along the suckers toward the direction of the mouth, so that considerable computation is actually carried on peripherally. Wells and Wells (1957) have conducted extensive experiments on this tactile system by conditioning blinded animals to accept small, smooth balls of perspex plastic and reject such balls with grooves cut in them, or vice-versa.[7] The

animals rapidly learn to draw in one type of ball and push away the other. The sophistication of these experiments has been further increased by training the two sides of split-brain animals in opposite directions.

One of their early findings was that a period of time was necessary for learning that had taken place with one arm to spread to the other arms. This spreading probably takes place through the inferior frontal lobes of the tactile system. Another finding is that there is little or no position sense in the arms and that tactile discriminations seem to be based on degree of roughness or smoothness.

3. Young's Mnemon Theory

As noted earlier, a point to point mapping is maintained from the retina to the surface of the optic lobe. Young postulates that oriented dendritic fields within the plexiform zone of the optic lobe serve for the extraction of visual "features", and calls cells with such dendrite fields "classifying cells". Tangential sections show a preponderance of dendrites in the horizontal/vertical direction, which is known to play a special role in the visual system of the animals, so that this seems to be a reasonable hypothesis and fits in well with the Hubel-Wiesel studies.

If each classifying cell can be switched between alternate pathways for "attack" and "retreat" by the action of taste and pain signals, then the entire assembly acts as a bistable element. Assuming some form of memory to be retained whenever an element is switched, each visual feature would acquire an associated memory value with experience. Young calls such an assembly a "mnemon." He postulates that the operation of such mnemons could come about through the action of recurrent collaterals from the memory cells of each pathway upon minute amacrines.
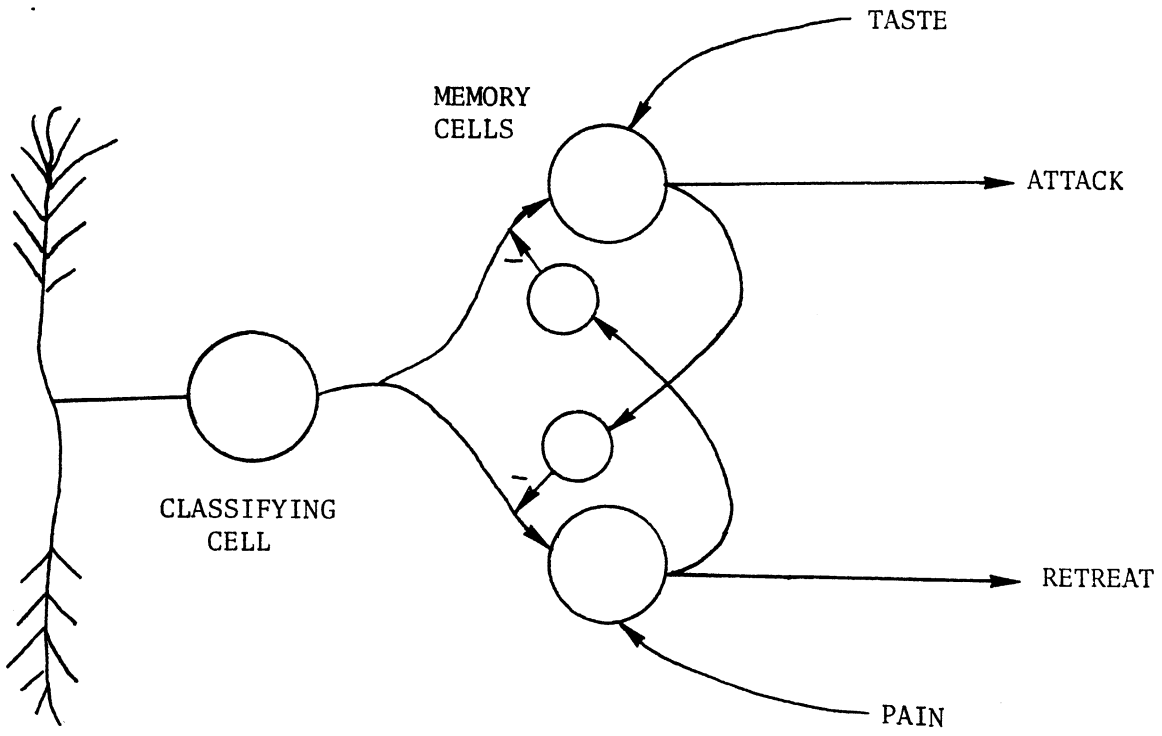
Figure 2.2

These, in turn, could block input to other memory cells by pre-synaptic inhibition. Electron micro-graphs do show the presence of amacrines and pre-synaptic inhibition in the optic lobes, and cells whose axons have several branches are also observed. There is evidence for these components, therefore. Such a mnemon assembly is shown diagramatically in figure 2.2.

Under this hypothesis the optic lobes would be the seat of visual memory and the upper lobes would function to "address" the proper mnemons, so that reading into and reading out of memory could be accomplished. Basically, these higher centers are thus thought to be associated with short-term memory and the process of reading to and from long-term memory. Whenever a set of classifying cells is activated by a stimulus object and an attack is initiated, some mechanism must be responsible for seeing that the taste or pain signals resulting from the attack affect all those mnemons that were active but no others, because the visual input itself will obviously have disappeared by then. This is a critical function in any organism which must learn from a real environment, and the mechanism involved is not known. Reverberation in the upper lobes could perhaps accomplish this addressing by temporarily keeping mnemons sensitized which had recently been active. Young postulates that once a mnemon has been switched, it remains so permanently, but the model described here assumes memory change can be incremental.

4. Summary

The octopus is an animal of intermediate complexity and has characteristics which make it a convenient animal for laboratory study. It can be conditioned to discriminate between a variety of stimulus objects both visually and by touch. Extensive anatomical studies have

been conducted and a great many neural pathways have been traced, with the visual system receiving particular attention. The latter consists of two large optic lobes, which are felt to be the seat of visual memory, and a two-tiered set of paired centers, thought to be associated with addresssing memory and other higher functions. Output from the optic lobes goes to motor centers where it is elaborated into patterns of actions for attacking or retreating. A similar set of centers is also found in the tactile system and is thought to function similarly. Young postulates a bi-stable assembly which he calls a "mnemon" as the mechanism for memory. Such an assembly has a classifying cell which extracts some visual feature and outputs into both the attack and retreat paths through memory cells. These are mutually inhibitory following an attack. Such an assembly could be kept sensitized by feedback through the upper lobes until the signals of results arrive. The next chapter discusses some approaches to modeling such learning systems.

CHAPTER THREE

SOME LEARNING MODELS

1. Initial Approach

The creation of any model involves a choice of which aspects of the problem to portray and which to ignore. The resolution of this question of emphasis hinges on our state of knowledge and interests regarding the problem, and our repertoire of tools for dealing with it. It is here that some element of art must enter the modeling process.

In its simplest terms, a model of learning will incorporate a memory of past experiences and a method for altering that memory. The initial approach here, therefore, involved a short pilot study of such an elemental system. If one assumes each stimulus can be represented as a set of component features and that the memory associated with each stimulus feature is independent and continuously variable, then the memory value of each such feature might be considered a weighted sum of its past associations with rewards and punishments. Model memory was simply taken to be a list of the last ten inputs which had been received (positive values for reward and negative values for punishment), each convolved with a linear weighting function. That is, the most recent input was given a weight of one and the other weights fell linearly to zero for the preceding inputs. This linear form was chosen for simplicity. A separate memory list was kept for each stimulus feature, or orthogonal set of features, and the output value read from memory for each list was just this convolution or weighted sum of previous inputs. A complex stimulus pattern could thus call forth an average of the outputs from several different lists, each contributing in proportion to its presence as a component of the stimulus pattern. Even such an elementary model allowed for considerable complexity, therefore.

The composite value read from memory in this fashion was equated to a velocity and used to determine attack or retreat. Hence a feature of this early model was a simulation of the attack itself. Position and velocity were updated on each time step, so that the simulated animal would decide which stimulus to attend if several were present, and would actually "follow" a prey around within its tank. Afterwards it would return "home". This aspect of the model was dropped when it was decided not to pursue "attention-focusing" mechanisms, but some of the graphical outputs which showed intensity of attack increasing with experience were rather interesting.

2. Two Stage Memory Model

This pilot study of behavior patterns in a learning automaton was succeeded by a more advanced model of learning with somewhat different emphasis, which was developed as a term project for a Computer Science course in Biological Simulation (CCS 680). The learning automaton had shown conditioning and interference, but was somewhat simplistic. One of the first things realized was that the unwieldy memory lists could be replaced by single numbers if exponential convolution functions were used instead of linear convolutions. The results of convolution for this case would be the old memory value multiplied by an exponential factor and added to the new input value. This change reduced storage requirements by an order of magnitude.

A two stage concept of memory was then introduced. This incorporated a short-term component which changed rapidly with current experience, but which tended to decay toward quiescence at the long-term value of memory, and a long-term component which changed slowly but tended to move in the direction of the short-term value. Each encounter with the environment thus eventually produced a slightly altered quiescence value

for memory, but long-term memory values could be temporarily masked by recent experience. This allowed behavior to be entirely different for a short period after an encounter from what it might later be, but required that memory be updated on every time step rather than only upon each encounter. This basic concept has remained present throughout later versions of the model.

Another interesting feature was what might be called an "attention" factor governing memory read-in and read-out. The amount of change to long-term memory resulting from an encounter with the environment depended on the value of this factor. If it were small there would be little change to long-term memory, even though short-term variations were not affected. The amount of read-out was also influenced by this attention factor. It acted as a weighting for each stimulus component, so that the contribution from each depended on the "attention" devoted to it. (The factors might alternatively be interpreted as arising from non-uniformity of the visual field.) Thus both read-in and read-out were under control of these attention factors, which entered the equations as weightings and convolution scale factors.

The original pilot study had included a crude characterization of hunger and digestion, but this characterization was extended and refined somewhat in the two-stage memory model. Hunger tended to increase exponentially toward its maximum, and was incrementally reduced as digestion proceeded of food objects which had previously been attacked. The value of hunger multiplied what was read from memory when a stimulus was present to determine a probability of attack, which was then compared to a random number drawn from a uniform distribution. If the attack or retreat probabilities exceeded this random threshold, the simulated animal would attack or retreat; otherwise it would take no action.

The model was used to generate learning and extinction curves which compared favorably with data taken from real octopus experiments. It was not pursued further, however, as it was decided to make a series of basic revisions which would create a model more closely tied to the mechanics of the mnemon and the anatomy of the octopus nervous system.

3. Single Mnemon Model

Chapter 2 discusses octopus nervous system anatomy and the theory of mnemons in some detail. Figures 2.1 and 2.2 show what is known about connections in the visual learning system and how a mnemon assembly might be constructed. The idealizations of figures 3.1 and 3.2 were arrived at from this information and were used as a basis for modeling the operation of a single mnemon.

One of the first things noted, even in the earlier models, was the greater convenience of normalizing all the variables, so that they can take on only values in the range from zero to one. Thus a normalized firing rate might range from zero for background level to one for the maximum rate of firing which can be sustained. Similarly, hunger levels between their minimum and maximum values can also be normalized to the range from zero to one, and when this is done for all of the variables, the model can be represented in terms of a normalized connection graph or matrix. Each node of the graph represents one of the normalized state variables of the model and the edges give values of the connection factors between them.

Within this general schema, the interactions between nodes can then be written in terms of a set of difference equations as follows. Let $v_j(t)$ represent the value of the normalized variable at node j, and let w_{ij} be the weighting, or connection factor, from node i to
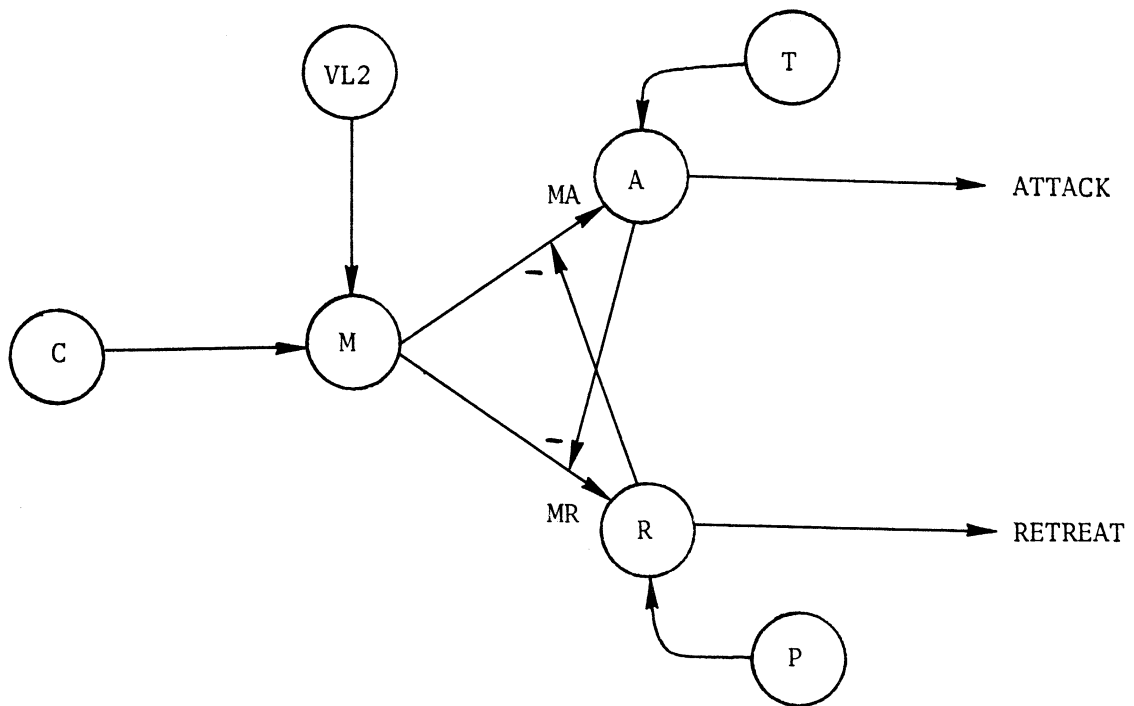
Figure 3.1

(S = A+R)
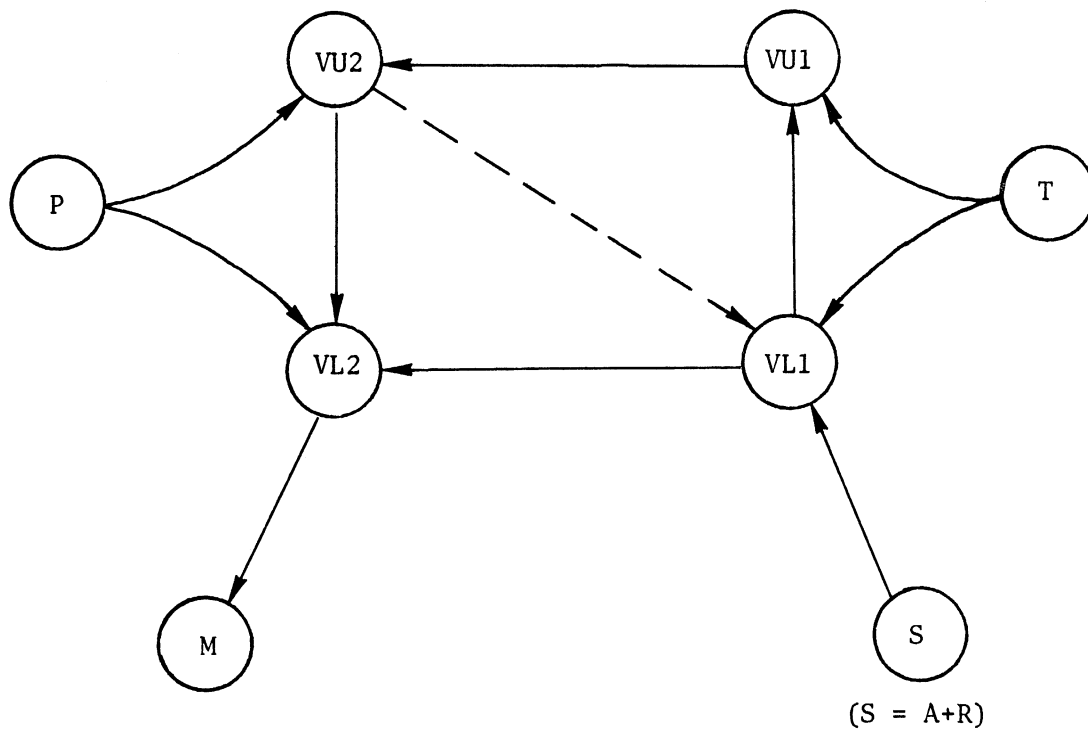
Figure 3.2

to node j. Assume also a rate constant, θ_j, associated with each variable. Then the general form of the difference equation transition function can be written as:

$$(3.1) \qquad v_j(t+1) = <\theta_j \cdot \sum_{i=1}^{n} w_{ij} \cdot v_i(t) + (1-\theta_j) \cdot v_j(t)>$$

where the pointed brackets are used to imply that the value of v_j is kept restricted to the range from zero to one. This shows that the value of v_j on the next time step depends on its current value and the sum of its inputs. The amount of change to v_j depends on the value of the rate factor, θ_j. When θ_j is small, v_j changes very slowly. As θ_j approaches one, on the other hand, v_j tends toward the sum of its inputs.

If every variable of the model could influence every other variable of the model, then n variables would require the specification of n^2 parameters with this formulation. Fortunately, anatomical constraints and other considerations reduce this number drastically, so that "tuning" the model actually involves considering the values of only a few free parameters.

The state variables corresponding to the simplified diagrams of figures 3.1 and 3.2 would be the following: classifying cell input (C), mnemon firing rate (M), attack path (A), retreat path (R), taste (T), pain (P), attack and retreat memory (MA and MR), and a represen-tation for each of the upper lobes (VL1, VL2, VU1, VU2). In fact, though, the model was somewhat more complex, and its actual connection diagram is shown as figure 3.3, Disregarding the details of this figure, the chief point of interest is that mnemon structure and anatomy were included as state variables in the model, and that an attempt was made to identify such state variables with physical quantities (such as firing rate in some cell or lobe) whenever possible.
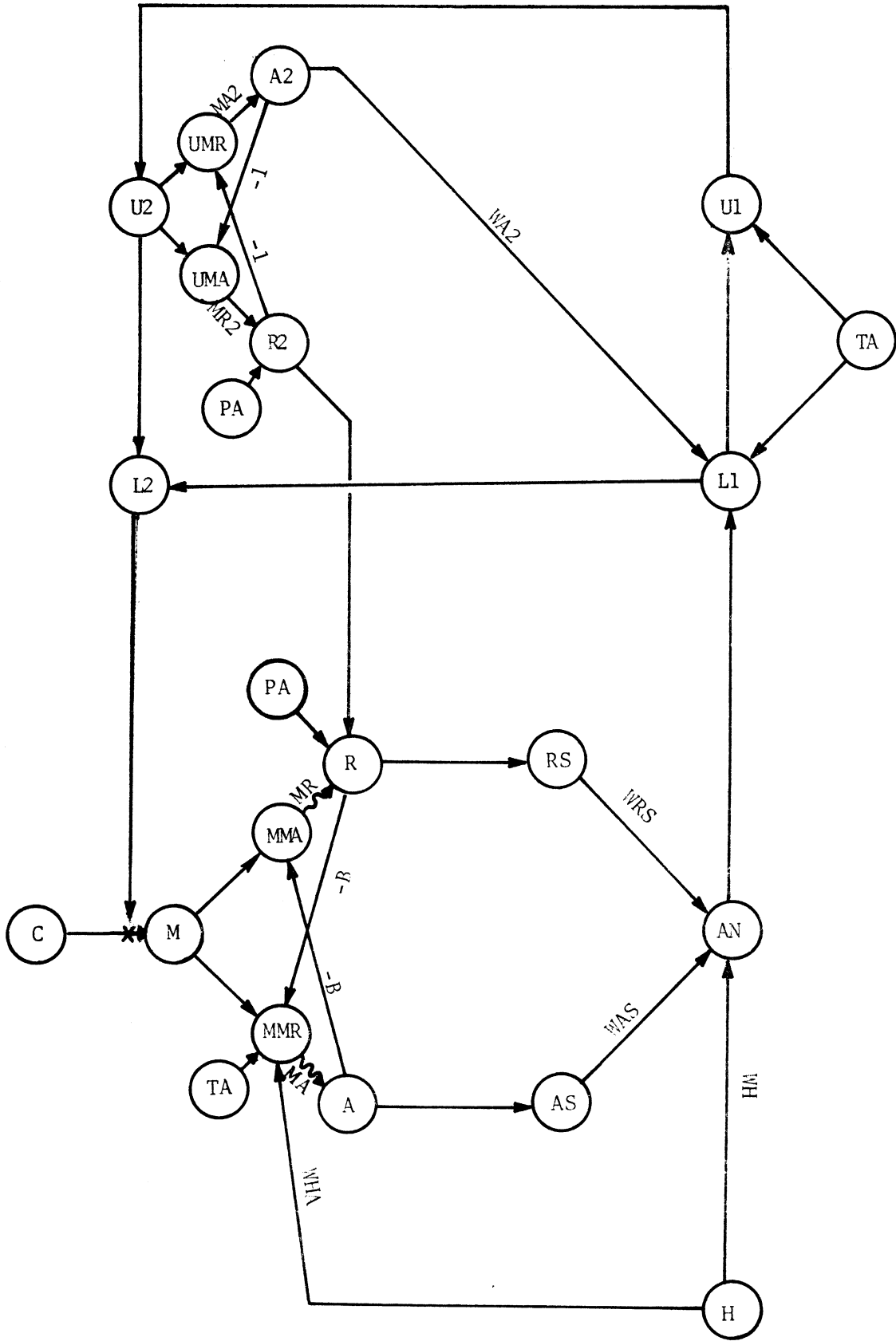
Figure 3.3

This version also incorporated the two-stage concept of memory, with perturbations in short-term values causing changes to long-term memory as they decayed back toward quiescence. The amount of this read-in was dependent on upper lobe activity through the value of VL2, which was assumed to control build-up and decay in the mnemon input. The effect of individual parameters in the model was thus somewhat complex, since each typically entered into more than one process. Rate constants were constrained to realistic half-life values, however.

The model did show octopus-like characteristics in many ways. The latency before attack decreased with experience, for example, and recent experience could mask longer term memory temporarily. In reversal experiments, both attack and retreat memories would increase with training so that the model would eventually tend to a random level of performance, just as real animals do. (There is little evidence for a specific "learning to reverse" in the octopus.) Performance following simulated operations was also similar. Interfering with the upper circuits would prevent read-in and read-out of memory from the mnemon. Trials at short intervals would show apparent learning, due to the short-term aspect of memory, but there would be little or no long-term change to memory. Recent taste or pain input would show a disproportionate effect therefore.

Several features were incorporated into the simulation program itself to increase its efficiency. One was a test for "quiescence". If all the variables changed value on any time step by less than some small fixed amount, then clock time was set forward to the next "event". (This event might be presentation of the next stimulus or removal of the current stimulus, for example.) Slowly changing variables were also updated at this time using an exponential approximation. This

technique, though less sophisticated than using a variable step size, worked very well.

Another feature introduced was a flexible system for specifying input sequences which allowed a set of experiments to be succinctly defined by control cards. One control card would be required to specify all the parameters of a single stimulus presentation, but once specified in this way, it could be used in a repetitive sequence. If exponents denote repetitions and parentheses delimit basic sequences, then this "mini-language" allowed such forms as: $A \ B^2 \ C$, $(A \ B^2)^3 \ C^2$, etc., where A, B, and C are different stimuli. An experiment could also be repeated with different initial values of the internal variables, or with a different set of parameter values, if desired. Additional flexibility was gained by providing the option of punching out final values for all of the variables at the end of a run so that the program could read these in and continue from that point at a later time. These features proved useful during development of the program.

4. <u>Maldonado's Octopus Model</u>

One other model which should be mentioned is the block diagram model of visual-learning circuits in the octopus brain developed by Maldonado[1], shown in figure 3.4. Many features of the models described in sections 1-3 are similar in concept to Maldonado's view of brain operation and draw inspiration from it. Classification units activate memory units which also receive taste/pain signals and feedback from a pair of higher level amplification systems. Output from the memory units is summed and goes into these amplification systems, which also receive taste and pain input. The output of the amplification systems is called the "experimental parameter" (EP), and along with a "hunger parameter" (HP) and "unspecific effect parameter" (UEP) controls the
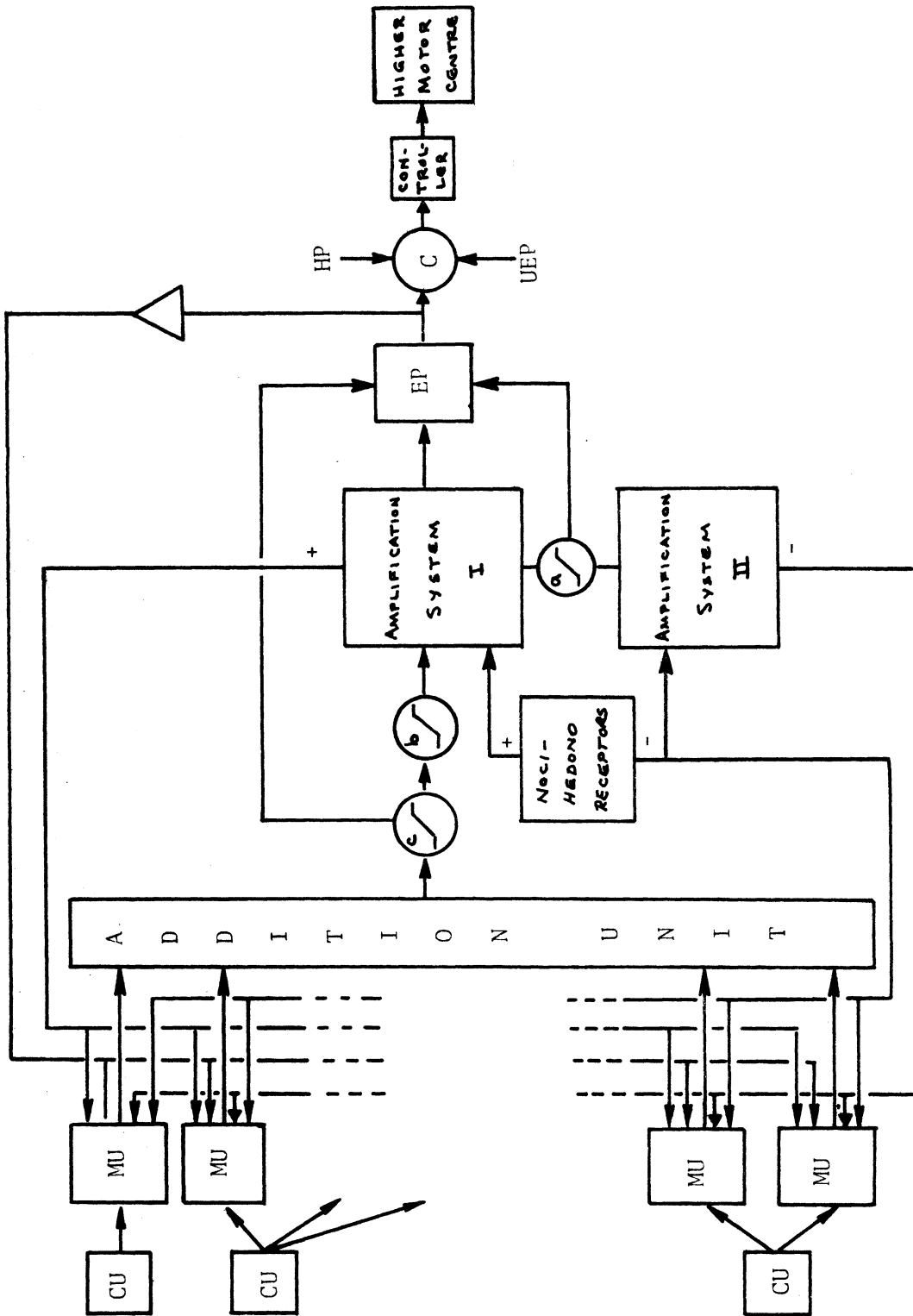
Figure 3.4

command (C) to attack or retreat. This command passes to higher motor centres where it is elaborated into a course of action. Very recent experience has its effect through the unspecific effect parameter, which raises or lowers the general tendency to attack irrespective of which classification units are active. Positive feedback occurs through the value of the experimental parameter returned to the memory units. Maldonado postulated two identical sets of these classification and memory units corresponding to the two optic lobes, with like units on each side being connected through the ventral optic commissure. This would account for transfer of learning from one eye to the other and the necessity for the vertical lobe structures (corresponding to the amplifier circuits) to remain intact in order for this transfer to occur. The amplification system could be activated by signals from either eye, so that results of an attack would alter memory in the corresponding units of both optic lobes, thereby accomplishing transfer across the mid-line.

5. Summary

The development of tractable models of complex phenomena requires that some aspects of the problem be given prominence and others ignored. The most interesting aspect of octopus behavior in this regard is the animal's ability to learn, and the interplay of short-term experience with longer-term experience. An initial pilot study of a learning automaton was conducted, and then followed by a term-project in which a model incorporating separate long and short term values for memory was developed. The next step was a general model in the form of a network of state variables and connection factors, with a difference equation specifying transitions. This was constrained to a reasonably small number of free parameters by knowledge and assumptions regarding

circuitry in the octopus nervous system, and an attempt was made to relate model variables to identifiable portions of the real system insofar as possible. A block diagram model by Maldonado provided inspiration for some of the concepts in these models. Details of the latest model verison are given in the following chapter.

CHAPTER FOUR

DESCRIPTION OF THE MODEL

1. Overview

Chapter 3 described some of the earlier approaches taken toward developing models of learning. During the summer of 1971, I was fortunate enough to spend some time at the *Stazione Zoologica* in Naples and to have an opportunity there to observe behavior of the animals firsthand. Following my return, I began work on a multi-mnemon model which was somewhat simplified and incorporated an extended version of long-term memory. The model thus came to have the three levels of memory described in quotations from Horridge and Wooldridge in section 2 of chapter 1. At the same time, certain simplifications were incorporated which shifted emphasis away from close conformity with the neuro-anatomy in order to produce a multiple mnemon simulation model with reasonable running times on the computer.

The general form of this multi-mnemon model is shown in figure 4.1. In concept it is quite similar to the models of chapter 3. Each mnemon is activated by a particular classifying cell input, and can be switched by its taste/pain inputs. Each incorporates the three levels of attack memory and retreat memory mentioned above, and these control its outputs. The attack outputs of all mnemons are summed to produce an overall attack strength (AS), and the retreat outputs are likewise summed to produce a retreat strength (RS). These strengths are then combined and used to determine whether the model will attack or retreat. Inputs activate the upper lobe structures to generate a value (Q) which is necessary for "reading" into or out of the mnemon memories.

The operation of an individual mnemon is shown in more detail with figure 4.2 and will be discussed next. Input (C) activates an
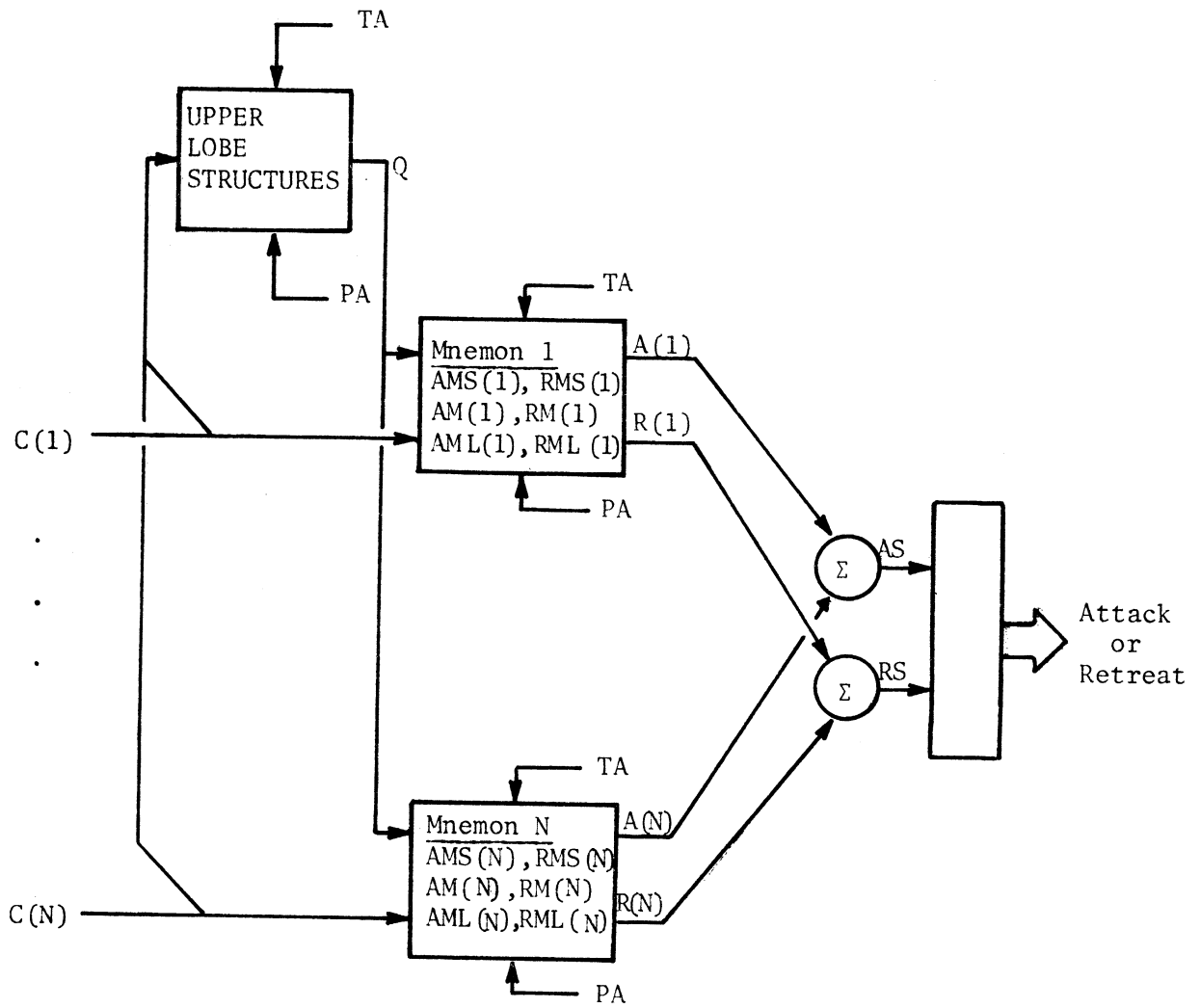
Figure 4.1
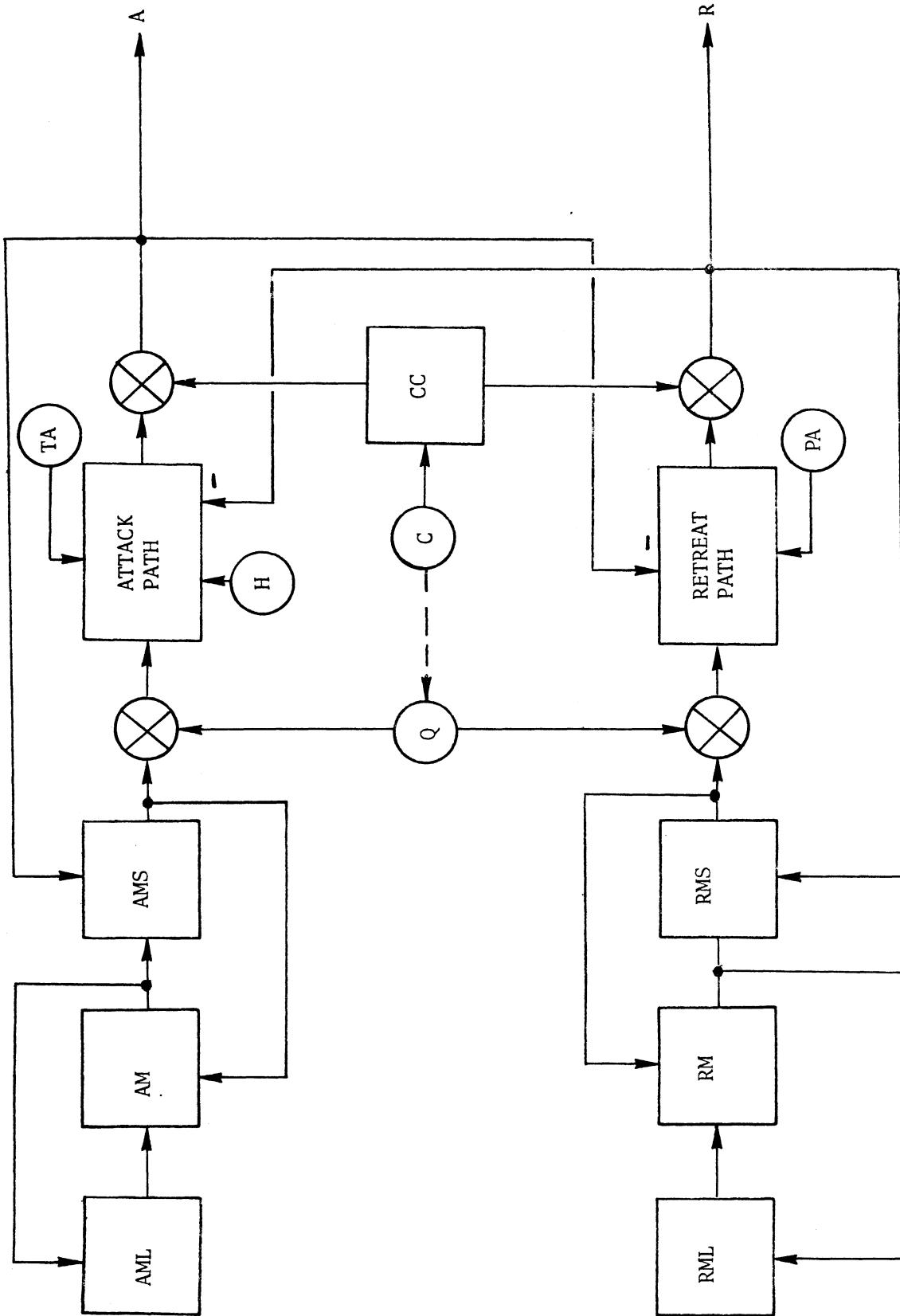
Figure 4.2

internal variable (CC), which then enables the outputs of the attack

(A) and retreat (R) paths. These, in turn, are produced by read-out

from the attack and retreat memories, which are enabled by feedback

from the upper lobes (Q). These attack and retreat paths are mutually

inhibitory, and the attack path also receives a bias from hunger (H).

In addition, there are random factors (RNA and RNR) which enter the

attack and retreat paths.

The value of short-term attack memory (AMS) is influenced by

positive feedback from A, and as AMS increases, it also tends to increase

medium-term memory (AM). Likewise, changes in AM tend to cause the

level of long-term memory (AML) to rise. An analogous situation obtains

for the retreat path, with feedback from each type of memory influencing

the type preceding it. Taste (TA) or pain (PA) inputs to the attack

or retreat paths can override the influence of memory temporarily and

cause the mnemon to switch modes until these inputs decay below the level

of memory influences once again. The exact time course of events thus

depends on memory levels at the time of attack and values of the decay

constants, and can be somewhat complex.

2. General Form of the Equations

All of the equations are similar in form. On each time step (taken

as equivalent to one second of real time), a set of "final values"

toward which each of the normalized state variables would tend in the

absence of other influences is computed:

4.1 $$F(x_j) = f_j(x_1, x_2, \ldots, x_m)$$

The x_j are state variables (such as A, R, Q, etc.), and $F(x_j)$ is the

"final value" of variable x_j, which may in general be some function

dependent on all the other variables.

Once these final values have been found, a "modified difference"

is then computed. Let E_j be the exponential rate constant (previously

denoted as θ_j in chapter 3) associated with state variable x_j, and

let:

$$4.2 \qquad E_j = \begin{cases} ER_j & F(x_j) \geq x_j \\ EF_j & F(x_j) < x_j \end{cases}$$

where ER_j and EF_j are model parameters. That is, variable x_j may have

different rate constants for rise and for fall. The modified difference

is then simply the difference between the final value and the current

value, modified by this rate constant:

$$4.3 \qquad D(x_j) = [F(x_j) - x_j] \cdot E_j$$

Notice that although x_j is normalized to the range from zero to one,

there is no such restriction on the modified difference. The new values

of the state variables must therefore be checked and possibly truncated

following update:

$$4.4 \qquad x_j(t+1) = x_j + D(x_j)$$

$$4.5 \qquad x_j = \begin{cases} 0 & x_j \leq 0 \\ x_j & 0 < x_j \leq 1 \\ 1 & 1 < x_j \end{cases}$$

Truncation creates the danger that some variables may artificially be

"locked" at their extreme values, but this possibility can be checked

during the "tuning" phase of model development.

The basic sequence of operations defined by equations (4.1) - (4.5)

is carried out on each time step, with new values of variables replacing

the old. The exact form of the transition equations is given in detail

in the following section.

j	x_j	Description
1	CC_i	internal input variable
2	AMS_i	short-term attack memory
3	RMS_i	short-term retreat memory
4	A_i	attack path output
5	R_i	retreat path output
6	AM_i	medium-term attack memory
7	RM_i	medium-term retreat memory
8	Q	upper lobe feedback
9	TA	taste
10	PA	pain

other variables:

C_i	classification input
H	hunger level
AML_i	long-term attack memory
RML_i	long-term retreat memory
AS	attack strength
RS	retreat strength

Table 4.1a

i	$F(x_j)$		$f_j(x_1 \ldots x_m)$
1	$F(CC_i)$	$=$	C_i
2	$F(AMS_i)$	$=$	$[AM_i + A_i] \cdot Q$
3	$F(RMS_i)$	$=$	$[RM_i + R_i] \cdot Q$
4	$F(A_i)$	$=$	$[AMS_i + TA + H \cdot HC + RNA \cdot RC - R_i] \cdot CC_i$
5	$F(R_i)$	$=$	$[RMS_i + PA + RNR \cdot RC - A_i] \cdot CC_i$
6	$F(AM_i)$	$=$	$\begin{cases} AMS_i & AMS_i > AM_i \\ AM_i & AMS_i \le AM_i \end{cases}$
7	$F(RM_i)$	$=$	$\begin{cases} RMS_i & RMS_i > RM_i \\ RM_i & RMS_i \le RM_i \end{cases}$
8	$F(Q)$	$=$	$CMAX + TA + PA \qquad (CMAX = \max C_i)$
9	$F(TA)$	$=$	0
10	$F(PA)$	$=$	0

other variables:

C_i	$-$	input
H	$-$	constant during session
AML_i	$-$	updated between sessions
RML_i	$-$	updated between sessions
AS	$-$	sum over A_i
RS	$-$	sum over R_i

Table 4.1b

3. <u>Model Equations</u>

Using the notation developed thus far, let i range over the set
of mnemons and let j index the state variables. A description of these
variables is given in table 4.1a and the actual equations used in the
model are given as table 4.1b. The constants HC and RC which appear in
equations 4 and 5 are weighting factors, or parameters which determine
the relative importance of hunger and random influences in the model.
(The random numbers RNA and RNR are taken as cubes of values drawn
from a distribution which is flat from -1 to +1.)

These final values of table 4.1b are used to compute modified
differences and new values for each of the variables, as discussed in
section 2. Once new values have been computed for all of the state
variables, a test is made for attack/retreat. This involves use of the
sum AS and RS in a threshold function:

4.6a $\quad AS = \sum_i A_i$

4.6b $\quad RS = \sum_i R_i$

4.6c $\quad CS = \sum_i C_i$

4.7 $\quad AS \cdot (1 + AS - RS) > CS \Rightarrow ATTACK$

4.8 $\quad RS \cdot (1 + RS - AS) > CS \Rightarrow RETREAT$

If neither condition obtains (or if no stimulus is present), then
no action is taken on that time step. This will become clearer as the
structure of the program is discussed.

The update of long-term memory takes place between training sessions
and on a different time scale. The equations for this update are as
follows:

4.9 $\qquad AML_i = \begin{cases} AML_i + DF \cdot (AM_i - AML_i) & AM_i > AML_i \\ AML_i & AM_i \leq AML_i \end{cases}$

4.10 $\qquad RML_i = \begin{cases} RML_i + DF \cdot (RM_i - RML_i) & RM_i > RML_i \\ RML_i & RM_i \leq RML_i \end{cases}$

The factor DF is an important parameter of the model which determines
the amount of read-in to long-term memory from the training received
during each session.

As mentioned in chapter 3, a test for quiescence is used in the
program to eliminate unnecessary computation. If the change in every
variable is less than some small amount, ε, on a given time step, then
clock time is advanced to the time of the next event and slowly changing
variables are updated with an exponential approximation. That is, if for
all j:

4.11 $\qquad D(x_j) < EPS => QUIESCENCE$

then, for example:

4.12 $\qquad H(t') = 1 - [1 - H(t)] \cdot e^{-\theta_H(t'-t)}$

and so forth. The value of EPS is a parameter of the model, but not
a very critical one.

4. Description of the Program

This model was used as the basis for a simulation program which
was coded in Fortran and run on an IBM 1800 computer at the *Logic of
Computers* research group. Later it was converted to run on an IBM
360-67 under the University of Michigan Terminal System (MTS). A
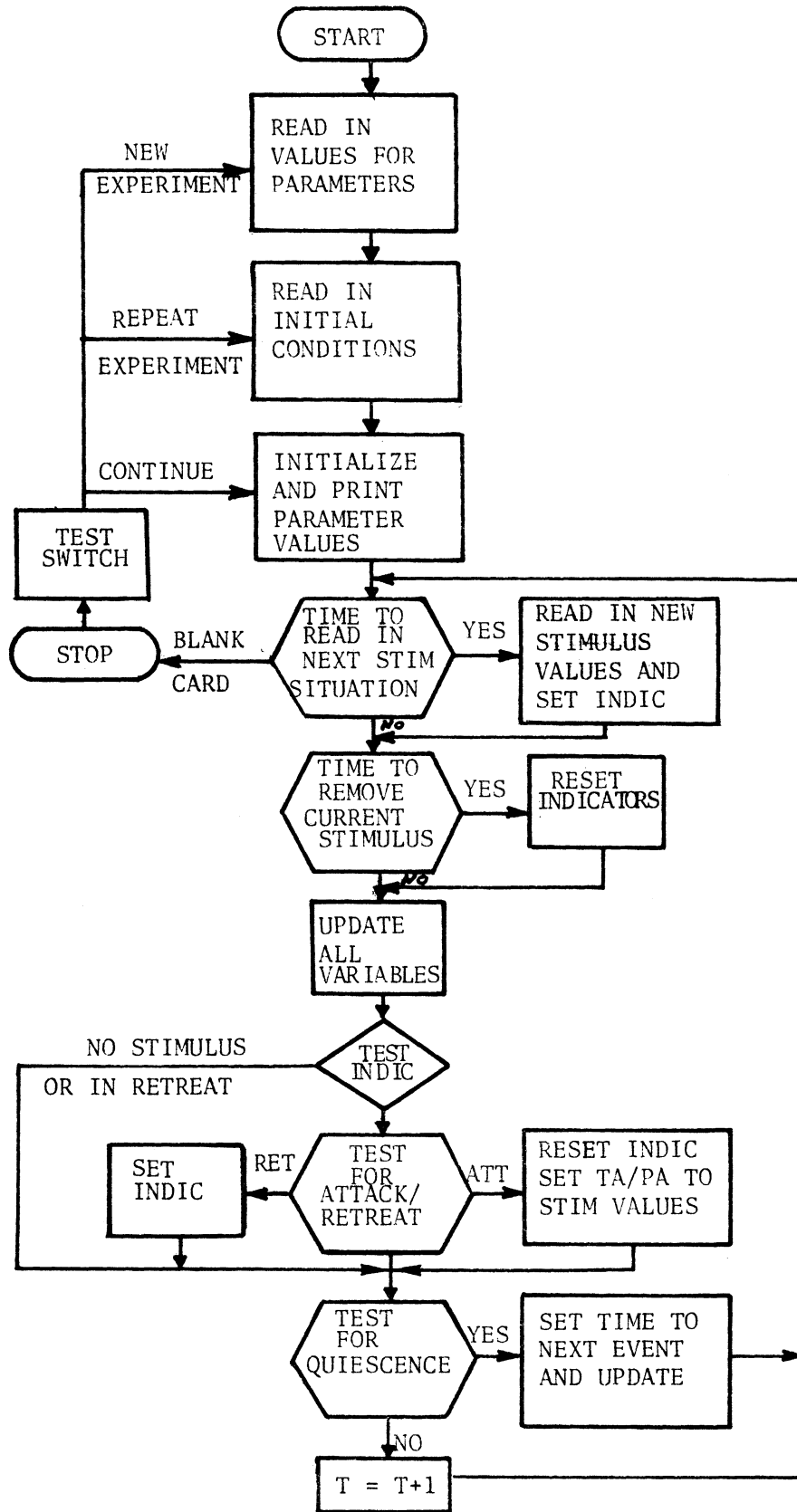simplified flow chart is given in figure 4.3. A discussion of this
figure follows.

Figure 4.3

Free parameters and initial conditions are read from cards. (All state variables are normally initialized to zero unless otherwise specified.) After initialization, the basic sequence of operations begins. Clock time is tested to see whether a new stimulus situation should be read in, or whether there is currently a stimulus present which should be removed. A system of indicators is used to show whether a stimulus is present or not, and if so, whether the model is in retreat state or not. Whenever an "event" occurs, the indicators are altered to reflect the event.

After all the variables have been updated, as discussed in section 3, a test is made for attack or retreat. In the case of attack, indicators are reset to show the absence of a stimulus, and TA or PA is set equal to the stimulus value for that variable. (This TA/PA value then begins to decay toward zero in accordance with equations 9 and 10 of table 4.1.) In the case of retreat, an indicator is set to show this fact. If no stimulus is present or if the model is already in retreat state, the test for attack/retreat is skipped.

Clock time is incremented by one unless there is quiescence, and the basic sequence is then repeated. This loop continues until a blank card is encountered, which terminates the run. A switch controls the entry point for the next run so that it can either begin an entirely new experiment, or repeat the preceding experiment (with different random factors), or simply continue.

Printouts of various types can be obtained at a number of points in the program, and these printouts are under the control of a set of print switches which are entered on a control card at the beginning of the run. (It is also possible for debugging purposes, to alter the switch settings from the computer console during the run.) Any combin-

ation of these fifteen or so different printouts can be obtained, which allows a great deal of flexibility in the microscopic/macroscopic level at which program operation is being viewed. This becomes of special importance due to the wide range of time scales encountered. It is sometimes of interest to view the changes in every variable at each step, and at other times only gross summary data is required.

It should be noted that the actual code for the program encompasses more complexity than that shown in figure 4.3, but this diagram indicates the basic conceptual structure of the program fairly well. Certain subtleties force complications in the program logic which would probably be more confusing than helpful if they were included in the figure.

Another point to be made is that I have striven over the course of the past three years to simplify the structure of the program and the notation of the model to the greatest degree possible, and that each revision provided an opportunity to profit from the mistakes of the previous version. The extent to which the notation is natural and easily understandable and to which the program logic contains a minimum number of epicycles, is a result of this constant process of revision and simplification. In computer programming, as in mathematics, a succinct result often conceals a multiplicity of cumbersome attempts.

5. Summary

The approaches described in chapter 3 were simplified in some respects, and extended to include three layers of memory and multiple mnemons. All of the update equations in the model have a common form which involves computing a final value, a modified difference, and a new normalized value for each state variable. The actual update equations are given in table 4.1 and equations 4.3, 4.4, and 4.5. The

structure of the simulation program is shown in figure 4.3. It involves an initialization and an iterative loop in which tests for attack or retreat are made following update, whenever a stimulus object is present. A flexible set of printout switches allow model behavior to be viewed at a variety of levels and time scales. The degree to which the representation of the model has been simplified and made more powerful reflects the extent to which past mistakes have proved instructive. Operating characteristics of the model and results derived from the simulation are presented in the next two chapters.

CHAPTER FIVE

LEARNING PERFORMANCE

1. Characteristics of the Model

Details of the model equations and simulation program were given
in the last chapter. This chapter takes a look at model operation in
typical learning situations, and compares some of its results with
experimental results. It may be useful in reading this section to refer
back to figure 4.2 and table 4.1 for a summary of the model.

A list of the parameter settings used in obtaining the results
of this section is given as table 5.1. Rate constants not equal to
one in the table are half-lives in seconds. As noted in chapter 4,
the rates of rise and fall may differ. From these parameter settings
it may be seen that one mnemon was involved (N=1), normalized hunger
level was set to one-half (except in figure 5.7), and that random
effects played no part (RC=0) except in figure 5.5. The initial
value of all variables was set to zero unless otherwise noted.

Figures 5.1 through 5.5 show values for several model variables
over time during encounters with the environment. A stimulus object
is presented on the initial time step in each figure by setting stimulus
input to one (C=1). This stimulus then remains present until attacked.
Each unit of time corresponds to one iteration of model operation and
may be taken to represent approximately a second of real time. The
ordinate for each variable shows its normalized range from zero to one.
(A magnified scale is used to show small changes more clearly in the
case of attack memory.)

Figure 5.1 illustrates the simplest case of an attack on a neutral
stimulus. The stimulus is represented by setting classification input
(C) to one until attack occurs (on time step 26). Following attack, C

Parameters

N = 1

H = .5   [H = 0 for figure 5.7]

HC = .1

RC = 0   [RC = .7 for figure 5.5]

Rate Constants

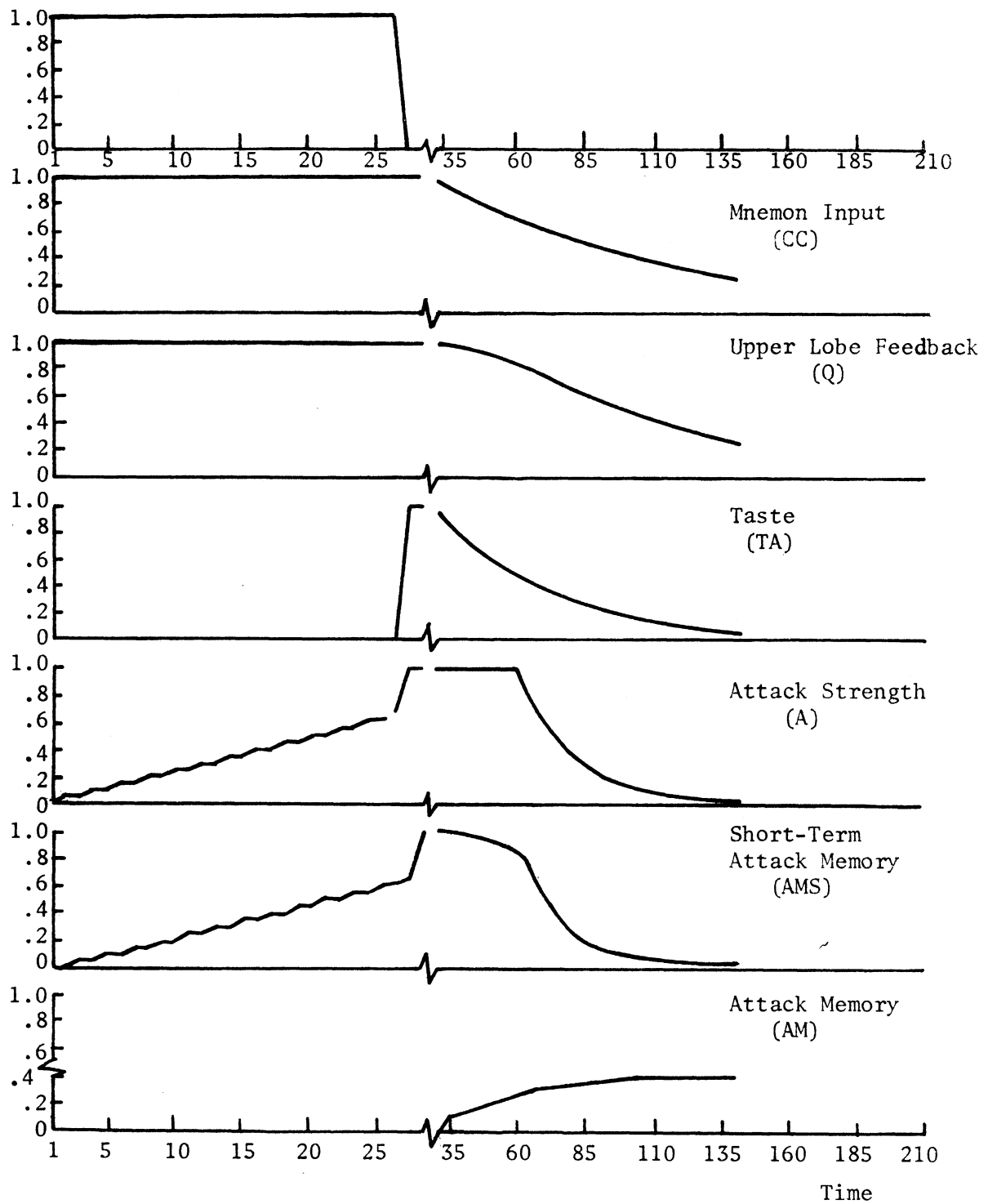|	Rise	Fall
CC	1	60
AMS	1	1
RMS	1	1
A	1	1
R	1	1
AM	900	--
RM	900	--
Q	1	30
TA	1	30
PA	1	60

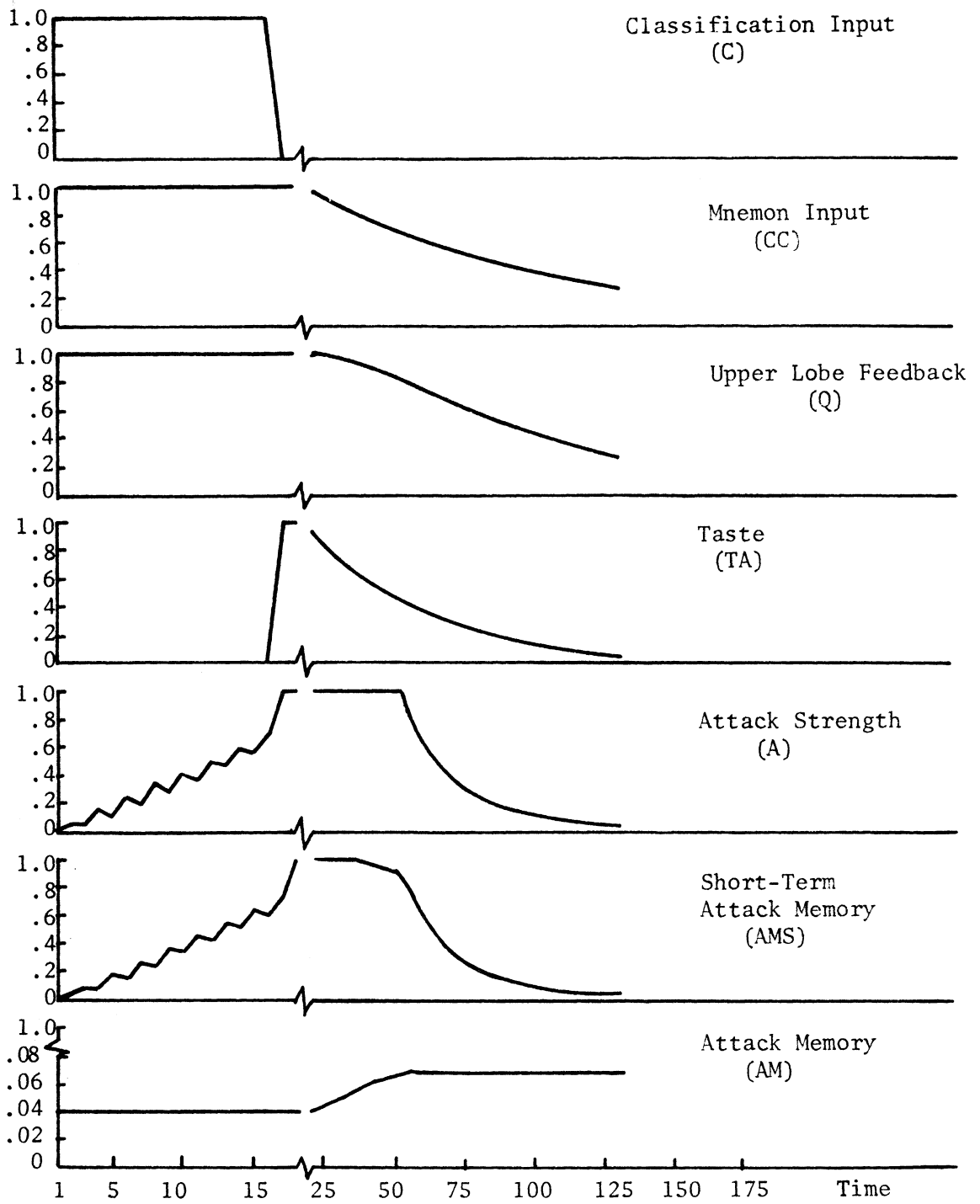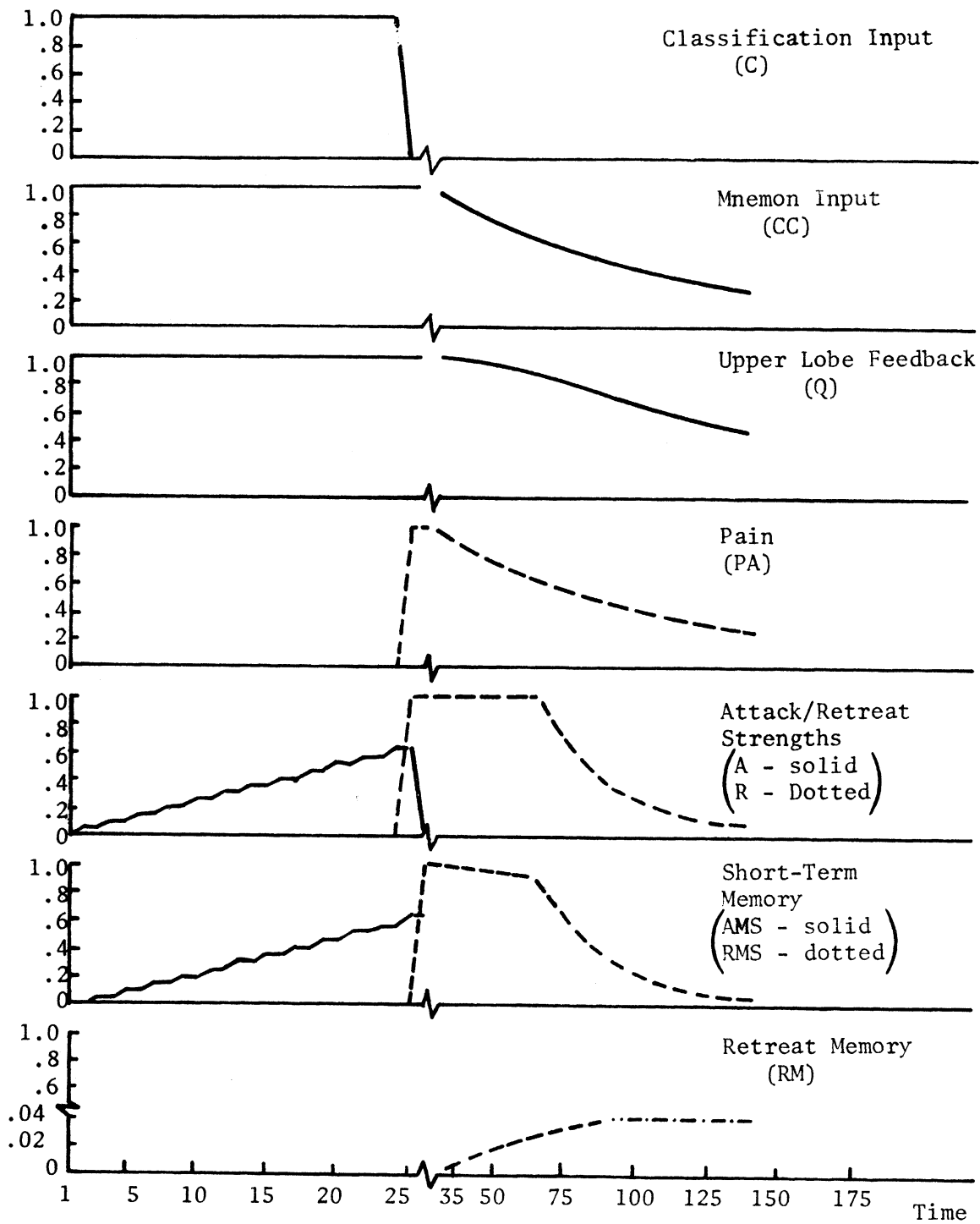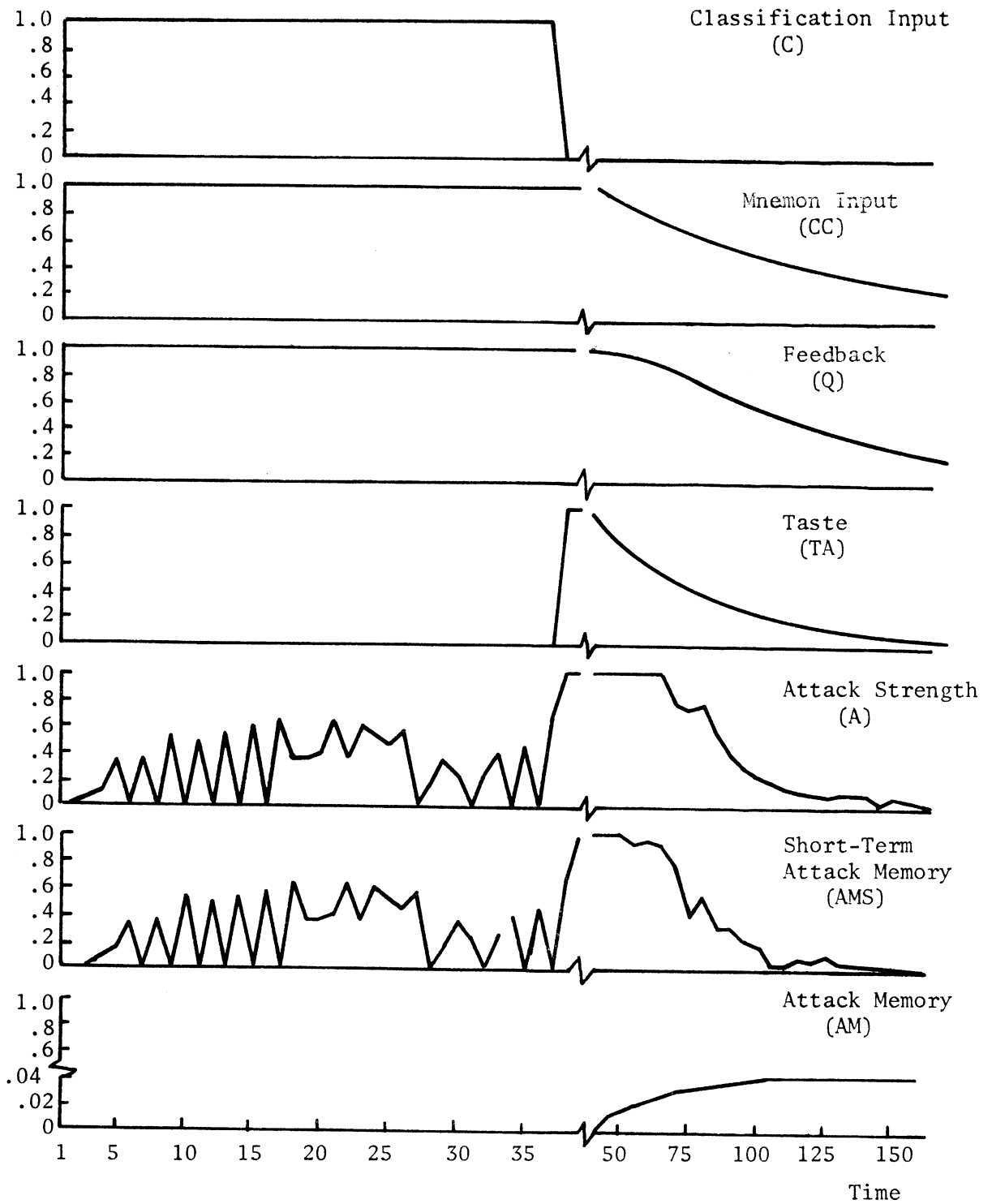Table 5.1
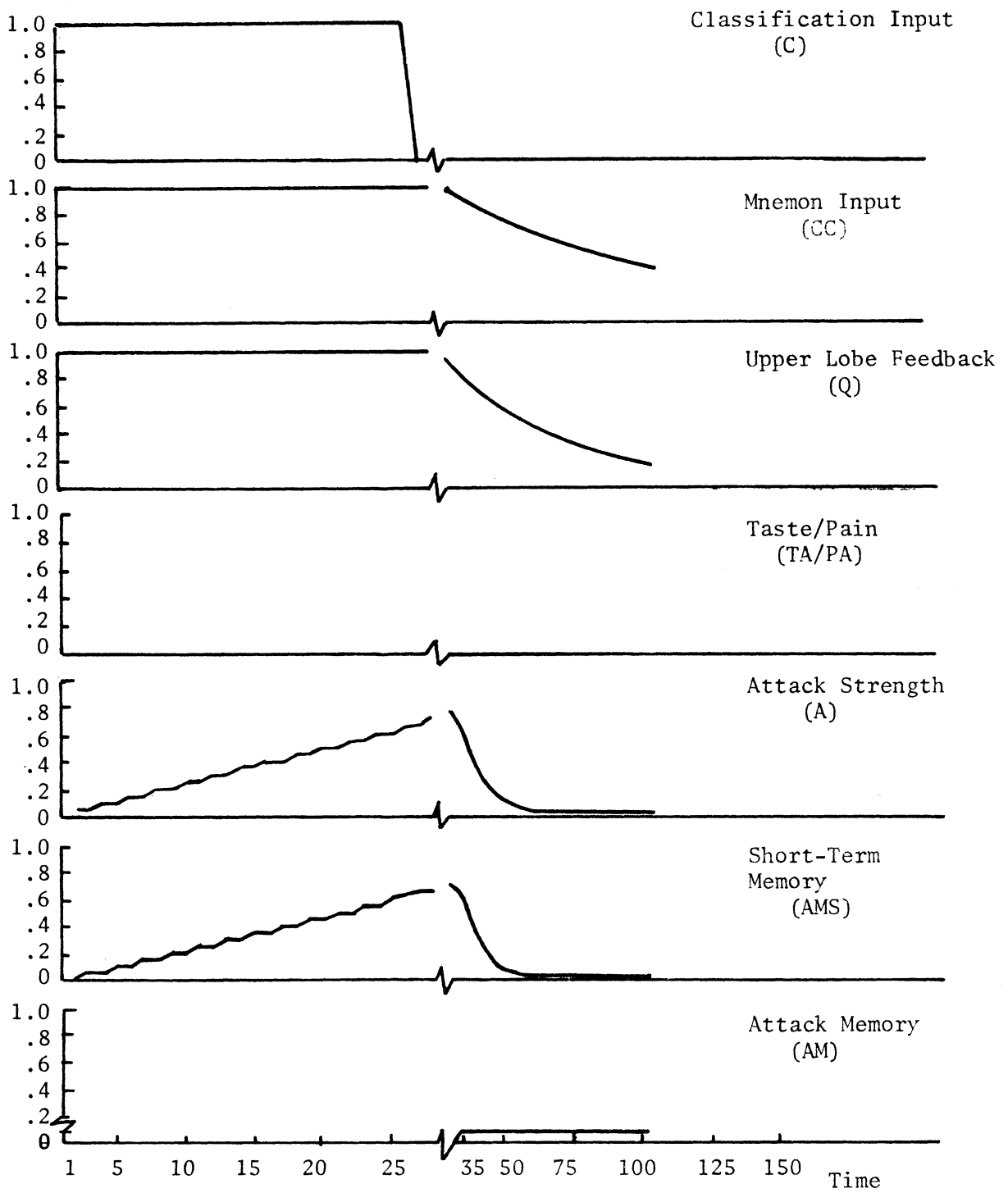
Figure 5.1

Figure 5.2

Figure 5.3

Figure 5.4

Figure 5.5

is reset to zero. From equation 1 of table 4.1 it can be seen that mnemon input (CC) "follows" C, so that CC then begins to decay toward zero with the rate constant shown in table 5.1 after attack. Likewise, the upper lobe feedback factor (Q) also "follows" C, and thus also decays toward zero with a rate constant which is given in table 5.1. Since the stimulus object is neutral in this case, the values of taste and pain remain zero throughout. From equations 2 and 4 of table 4.1 it can be seen that attack strength (A) and short-term attack memory (AMS) interact with one another in the presence of both CC and Q to drive A above the threshold for attack eventually. The rate at which this increase occurs depends on the level of attack memory (AM) and hunger (H). Since a naive mnemon is assumed here (AM = 0 initially), the slope of this rise in A and AMS is determined only by hunger level and the hunger factor (HC). Once CC and Q begin their return to zero after attack, A and AMS decay rather quickly, causing only the very small change of 0.01 in attack memory.

The next figure (5.2) shows exactly the same situation, but with a positive stimulus rather than a neutral one. In this case, taste (TA) goes to one following attack, and its presence forces A and AMS to one also. In addition, the presence of taste also slows the decay of Q. A becomes "locked" at its maximum value for about 30 time steps, and these high values of A and Q keep AMS from decaying. This in turn causes more "read-in" to AM than in the neutral case, so that AM eventually increases to a value of 0.04. (Refer to equations 2, 4, and 6 of table 4.1).

Figure 5.3 shows what happens when the same stimulus is then presented a second time. Attack buildup is much faster (16 seconds vs 26 seconds) because attack memory now combines with hunger to

influence the rate of increase in A and AMS. Otherwise, the form of the curves is much the same as before, and AM increases from 0.04 to 0.07 during the return to quiescence. The rate of attack buildup would continue to increase for subsequent attacks as AM grew larger.

Figure 5.4 shows the results of an attack on a negative object. Attack buildup is the same as in figure 5.2, but pain (PA) rather than taste goes to one following attack and this "switches" the mnemon to the retreat state. R and RMS (shown dotted) are driven to one by PA (equations 3 and 5 of table 4.1), and this forces A and AMS to zero. The presence of PA also slows the decay of Q, and keeps R and RMS at high values. Retreat memory (RM) thus increases from zero to 0.04 during this time when RMS is high.

In actual experimental situations there are, of course, many sources of variation from these idealized results. Extraneous noises and small variations in light level or temperature may have some effect on behavior, and stimuli with overlapping components may add to the animal's confusion. There may also be past associations with some stimulus components, so all of these effects are acknowledged in the model by inclusion of a random factor (RC). Figure 5.5 shows the same situation as figure 5.2, except that such a random factor has been introduced (RC = .7). The buildup of attack is much more erratic and takes much longer in this case, because both A and R receive random inputs according to equations 4 and 5 of table 4.1, and hence tend to cancel one another out. Only the effect of hunger adds a small bias in the direction of attack, and this eventually leads to an attack on time step 37. (Note: R and RMS are non-zero prior to attack but were not shown to prevent cluttering the figure.) Once attack occurs, the curves strongly resemble the no-noise curves of figure 5.2, but with minor perturbations. The end

results are equivalent, in the sense that AM eventually goes to 0.04 in both cases.

It is true in general, as a matter of fact, that once all the transient effects of an encounter with the environment have died out and the mnemon has returned to a quiescent state, the overall result from this encounter will be a change in the values of attack or retreat memory (AM or RM). The amount of change to AM is roughly proportional to the area under the short-term memory curve for AMS, as can be seen from figures 5.1 and 5.2. In figure 5.1 the stimulus is neutral and AMS drops rather quickly following attack, so that the change to AM is small. In figure 5.2, however, the positive result of the attack keeps AMS from decaying immediately, and AM thus changes by a greater amount.

Memory continues to increase with repeated attacks on the same type of stimulus, and figure 5.6 shows the shape of the resulting curves for AM-RM when the stimulus is positive, neutral, or negative. The curves show a characteristic exponential growth because change is proportional to the difference in AMS and AM (RMS and RM) rather than to AMS (RMS) alone, according to equations 6 and 7 of table 4.1.

A portion of the negative curve is shown dotted to indicate that this is the shape it would have were it not for the fact that attacks cease after two encounters. Positive encounters insure continuing attacks on the positive figure, but RM becomes sufficiently large after two negative encounters to cause retreat every time, in the case where there are no random effects. The necessity for introducing some random element thus becomes clear. Information about the environment is only acquired during attacks, but the environmental feature represented by a mnemon may not always have the same value. In some cases it may be positive and in other cases negative. The desire to avoid pain must therefore be
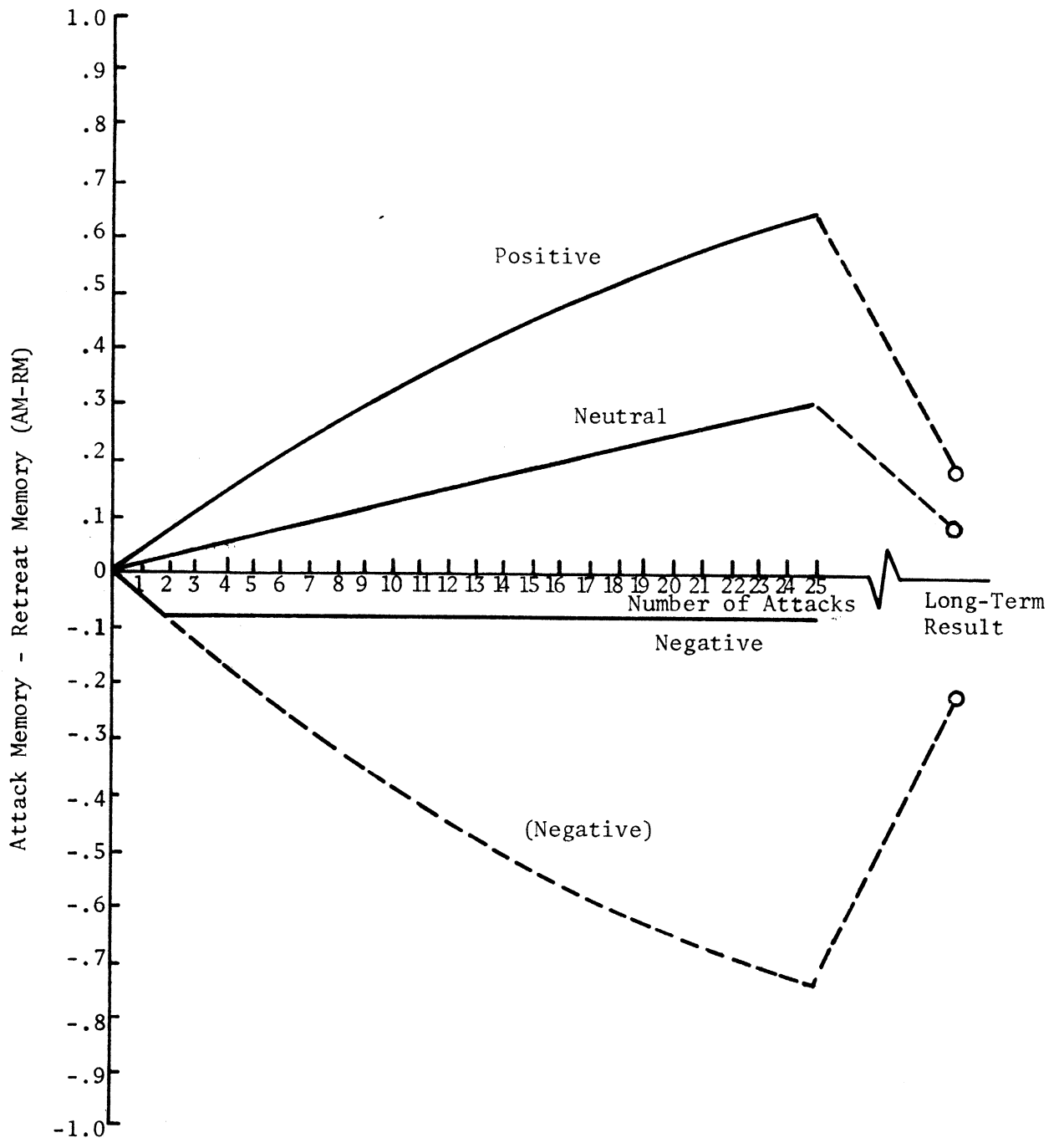
Figure 5.6

balanced against the necessity for obtaining additional information about the environment. The problem can be resolved by using memory as a bias in the random selection between attack and retreat. In this case, the probability of retreat from a negative stimulus continues to increase with each attack that is made on it, and therefore the interval between attacks on a negative stimulus will also increase with each attack that is made on it. The probability of attack never goes completely to zero, however. (Unless retreat memory were pre-set to one, corresponding perhaps to an innate, instinctual response.)

In any event, figure 5.6 shows how memory changes with each attack during a session. Corresponding to equations 4.9 and 4.10, however, there is also a permanent change in long-term memory (AML or RML) that takes place between sessions and is controlled by the parameter DF. With DF = .3, figure 5.6 also shows this final long-term result for the twenty-five attacks. Clearly, "within session" improvement in response will occur as memory increases, and performance will be better towards the end of a session than at the beginning of the next session. On the other hand, there will also be some overall "between-session" improvement. This seems to be a characteristic of real animals.

Another characteristic of real animals is that the time delay before attack tends to decrease with experience. Figure 5.7 shows latency of the model as a function of attack memory (in the absence of hunger effects). At very low values of memory the latency is quite large, but it decreases rapidly as AM becomes larger. This delay is related to the slope of the buildup curves for A and AMS, as noted earlier in the discussion of figure 5.1. As AM increases, the rate at which AMS and A build up becomes much faster, and this fact is reflected in the shape of the latency curve of figure 5.7.

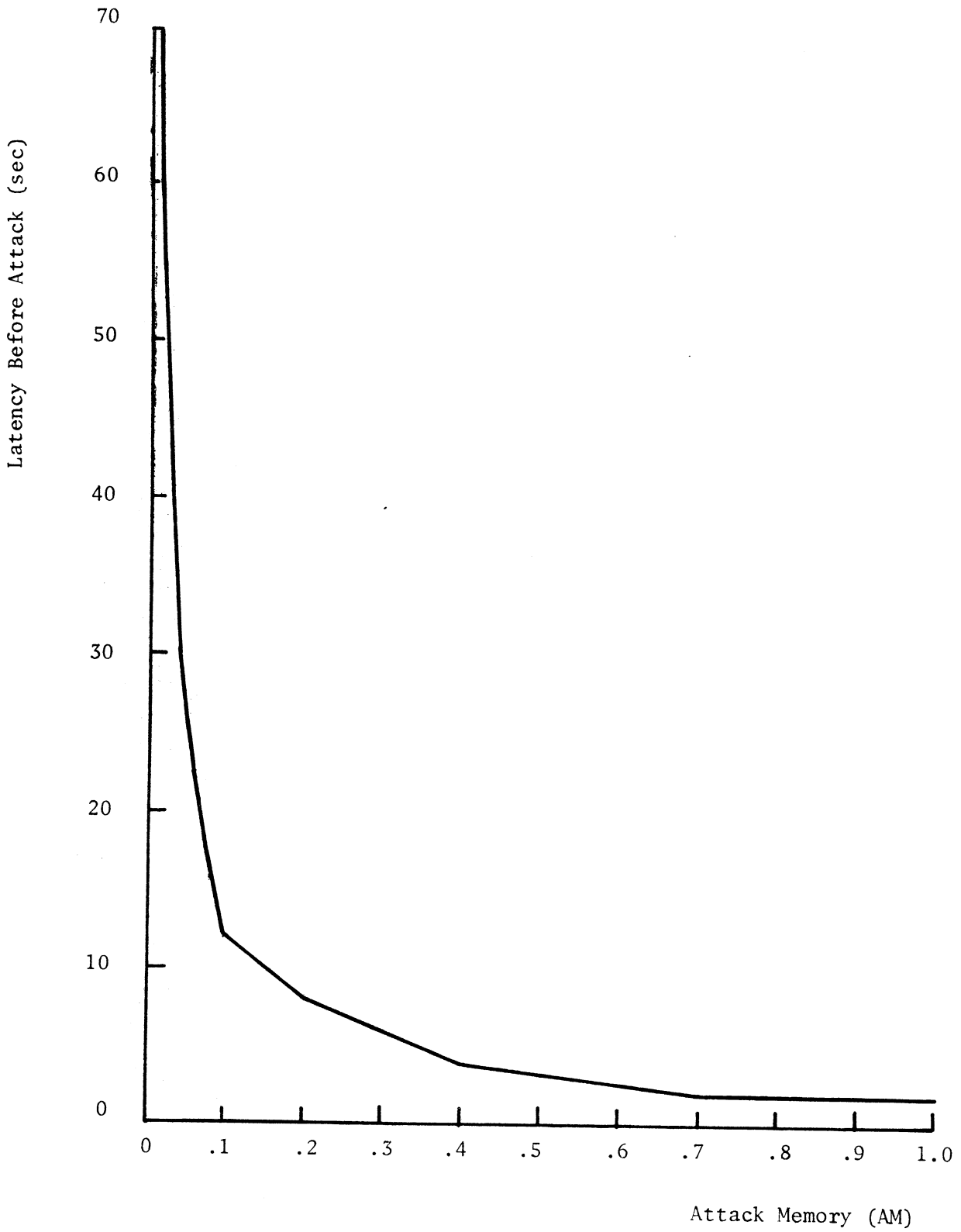Since the effect of hunger in the model is simply to add a bias

Figure 5.7

in the direction of attack, it can be seen from equation 4 of table 4.1

that hunger level becomes somewhat equivalent to attack memory. Figure 5.7

would thus have the same form if attack memory were zero and H·HC

were the abscissa instead of AM. Normalized hunger lies between

zero and one, so that a choice of 0.1 for HC assures that the most

variable portion of the latency curve is encompassed by the range of

H. This is the value listed for HC in table 5.1.

One additional characteristic of the model is shown in figure

5.8, which illustrates the effect of upper lobe feedback on the ability

of the mnemon to learn. In the discussion of section 2.2, it was noted

that the octopus has a set of upper lobe structures which are thought to

be involved in reading to and from memory through the action of rever-

beratory pathways, and that various operations which interfere with

these pathways tend to reduce learning performance. In section 4.1,

these lobes were incorporated in the model as the feedback variable, Q.

Hence operations which interfere with this feedback circuit can be

characterized by limiting the maximum value which Q can take on. When

this is done for several values of QMAX, a set of learning curves

corresponding to the positive curve of figure 5.6 can be obtained, and

these are given as figure 5.8. They show a rather continuous degradation

in learning as QMAX is progressively eroded. (Another way in which the

effect of operations might be modeled is by altering the rate constants

for Q. This was tried once or twice and seemed to have about the

same effect as the simpler procedure of limiting QMAX, so that it was

not investigated further.) Clearly, the value of QMAX is an important

determinant of learning and the effect of QMAX on group performance
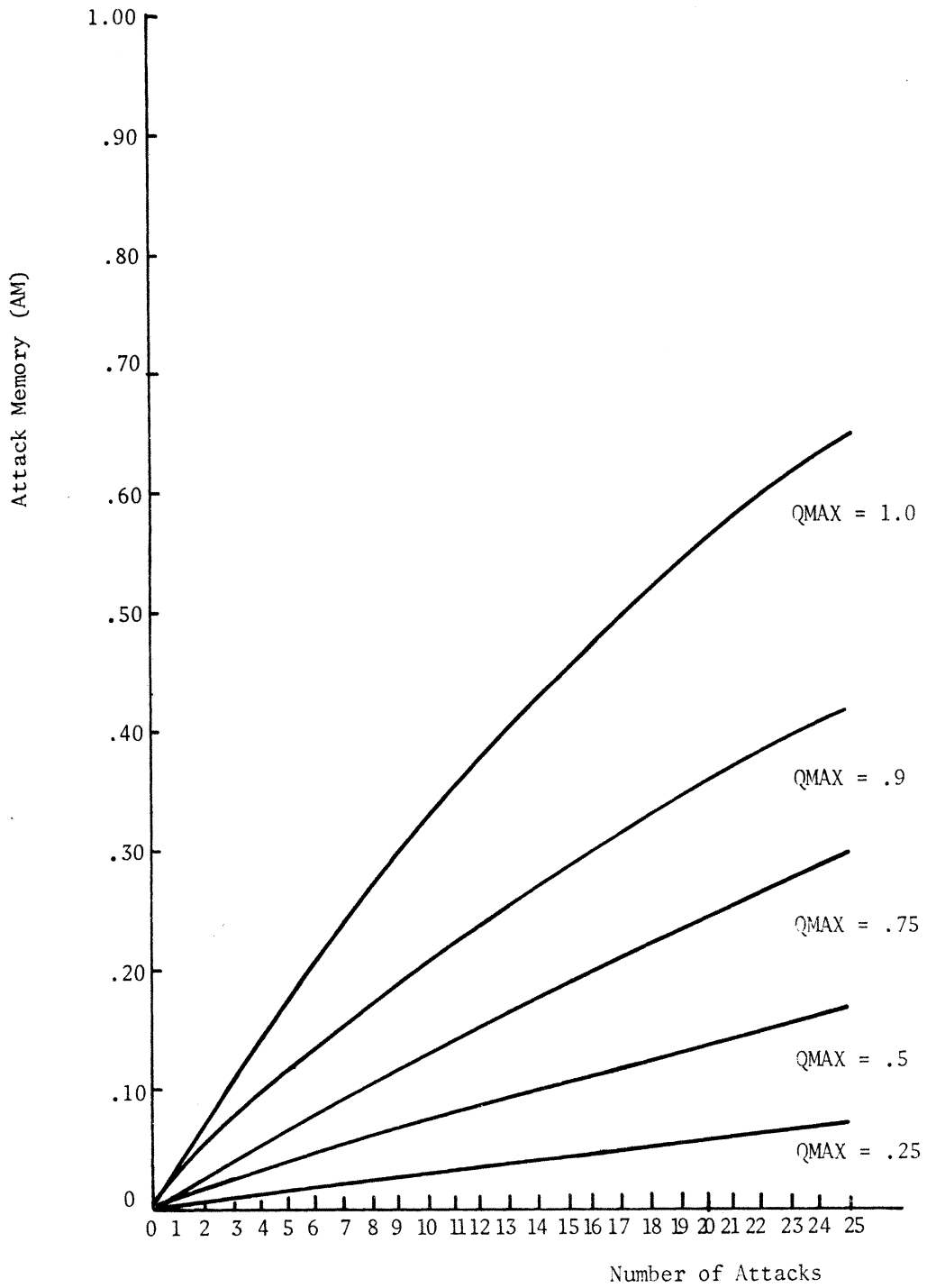
will be discussed later in section 3.

Figure 5.8

2. Comparison to Animal Learning Behavior

Section 1 was concerned with demonstrating the basic operating characteristics of one mnemon in various idealized learning situations, and in illustrating the effect of certain parameters. In order to compare model performance with that of real animals, however, it is necessary to look at statistical trends in group data. Animals tend to exhibit a wide variability in their individual performances, but when averaged together their behavior shows a characteristic form.

Typical "training" sessions consist of sixteen trials with two alternating figures. That is, positive trials (attacks on the positive figure rewarded) alternate with negative trials (attacks on the negative figure punished). During "extinction" sessions the figures are regarded as neutral, and attacks are neither rewarded nor punished. After each session the number of attacks on the "positive" figure and the number of attacks on the "negative" figure are noted and used to compute some performance measure for the session. A plot of this performance measure against sessions can then be regarded as the learning curve for an individual animal, and when the performance measures for a group of animals are averaged together in some fashion, the result is a group learning curve.

As a slight digression, it might be noted that two processes are involved here, viz, defining a performance measure for an individual animal based on its raw "score", and combining these individual measures to produce an overall measure for the group. Many different types of performance measures might be considered and each could have slightly different characteristics. For purposes of discussion here, however, the relatively straightforward difference in percentage attacks on the positive minus percentage attacks on the negative figure will be used. (Note, for example, that this implies both types of errors are equally

"bad". In the more general case, a relative weighting factor might be introduced to represent differences in "cost" between errors of different types.) For the second process (of combining these), the simple average of individual measures will be used to obtain a combined group measure. This simple average may be regarded as the special case of equal weights in the more general scheme of weighted averaging. (In general, all the animals are equal but some could be more equal than others, and weightings might be based on such things as average levels of attack or information content of the scores, for example.)

Table 5.2 shows the behavior of a typical animal[1] and of the model over a comparable set of experimental trials, consisting of two initial extinction sessions in which there was no reward or punishment, eight training sessions, and two final extinctions. Stimuli were presented for 20 seconds, and then removed if no decision to attack or retreat had been reached by that time. The interval between presentations was 5 minutes (300 seconds). Each attack is denoted by the letter "A". Slots are left blank in case of retreat or if no action was taken before the stimulus was removed. Raw scores and performance measures for each session are given to the right. Performance improves during the experiment for both the real animal and the model, but with a great deal of random variation in both cases. In addition, the model seems to learn more slowly and maintain a higher level of attack during the extinction sessions. In particular, animals learn retreat from the negative stimulus much more quickly.

ANIMAL	1 (+)	2 (−)	3 (+)	4 (−)	5 (+)	6 (−)	7 (+)	8 (−)	9 (+)	10 (−)	11 (+)	12 (−)	13 (+)	14 (−)	15 (+)	16 (−)	Total Attacks +	Total Attacks −	Performance Measure
E1		A	A	A	A		A	A									3	3	0
E2	A	A	A	A	A		A	A	A								5	3	25
T1	A	A		A			A		A				A		A		5	2	37.5
T2	A		A		A		A	A	A		A		A		A		8	1	87.5
T3	A	A	A		A		A	A	A		A		A		A		8	2	75
T4			A		A		A				A		A		A		6	0	75
T5			A	A	A		A		A		A		A	A	A		7	2	62.5
T6	A		A		A		A		A		A		A		A		8	0	100
T7			A		A		A		A		A	A	A		A		7	1	75
T8	A		A		A		A		A		A		A		A		8	0	100
E3	A	A	A				A								A		4	1	37.5
E4	A		A	A					A	A	A						4	2	25

Table 5.2a

MODEL	1 (+)	2 (−)	3 (+)	4 (−)	5 (+)	6 (−)	7 (+)	8 (−)	9 (+)	10 (−)	11 (+)	12 (−)	13 (+)	14 (−)	15 (+)	16 (−)	Total Attacks (+)	Total Attacks (−)	Performance Measure
E1		A	A	A	A	A	A		A						A		5	3	25
E2		A	A	A		A	A		A	A	A	A	A	A	A		6	6	0
T1	A	A			A			A	A		A	A			A	A	5	4	12.5
T2	A		A	A		A	A	A	A		A		A		A		7	3	50
T3		A	A		A	A	A	A	A		A		A		A	A	7	4	37.5
T4	A	A	A	A	A	A	A			A	A			A	A		6	5	12.5
T5	A		A		A		A	A			A	A					5	2	37.5
T6	A	A	A				A		A		A		A		A		7	1	75
T7	A		A	A	A		A		A	A	A		A		A		8	2	62.5
T8	A		A		A		A		A		A		A		A		8	0	100
E3	A		A		A		A		A	A	A		A		A		8	1	87.5
E4	A		A		A		A		A		A		A		A		8	0	100

TRIALS

Table 5.2b

Parameters

    N = 2

    H = .5

    HC = .1

    RC = .75

    DF = .25

    QMAX = 1

Rate Constants

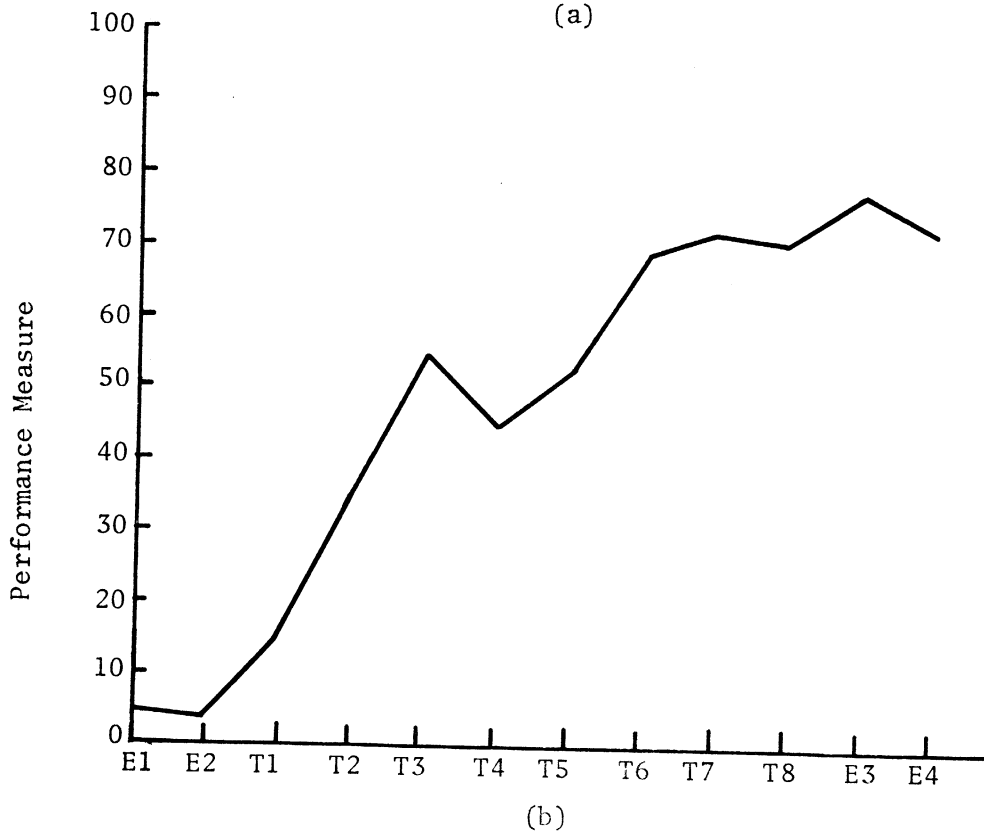|	Rise	Fall
CC	1	60
AMS	1	1
RMS	1	1
A	1	1
R	1	1
AM	900	--
RM	900	--
Q	1	30
TA	1	30
PA	1	60

Table 5.3

(a)

(b)
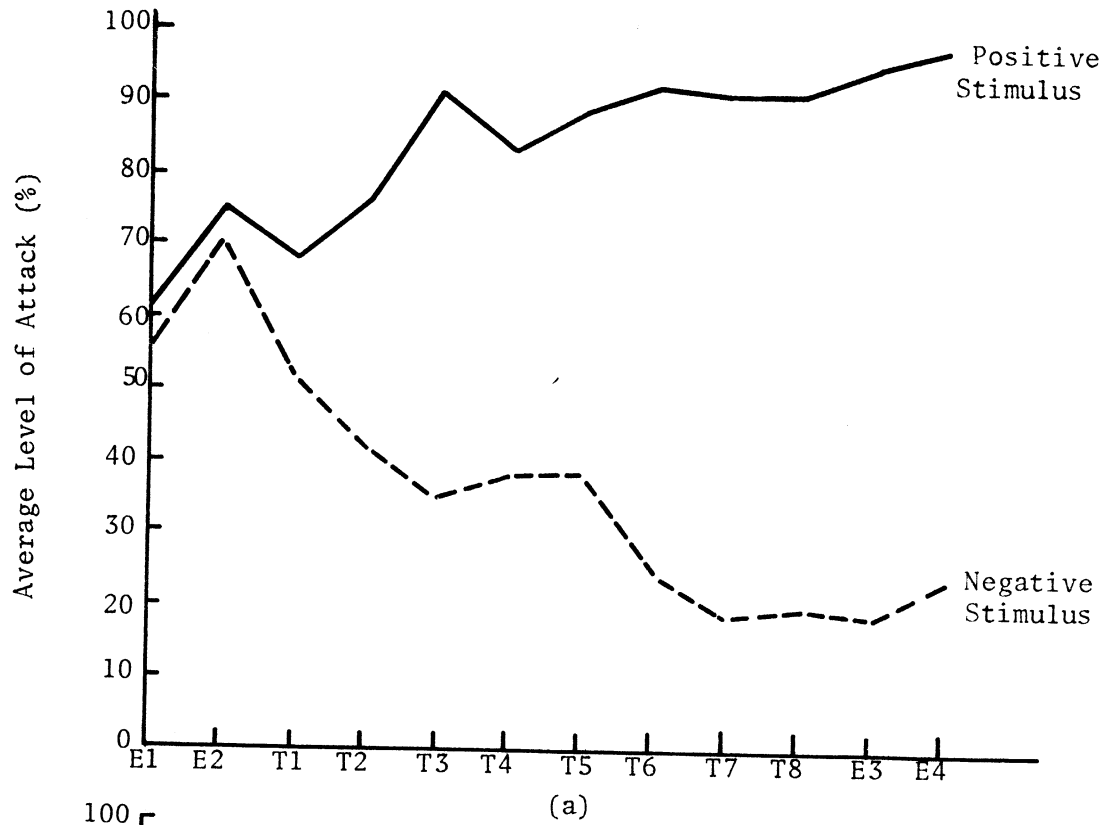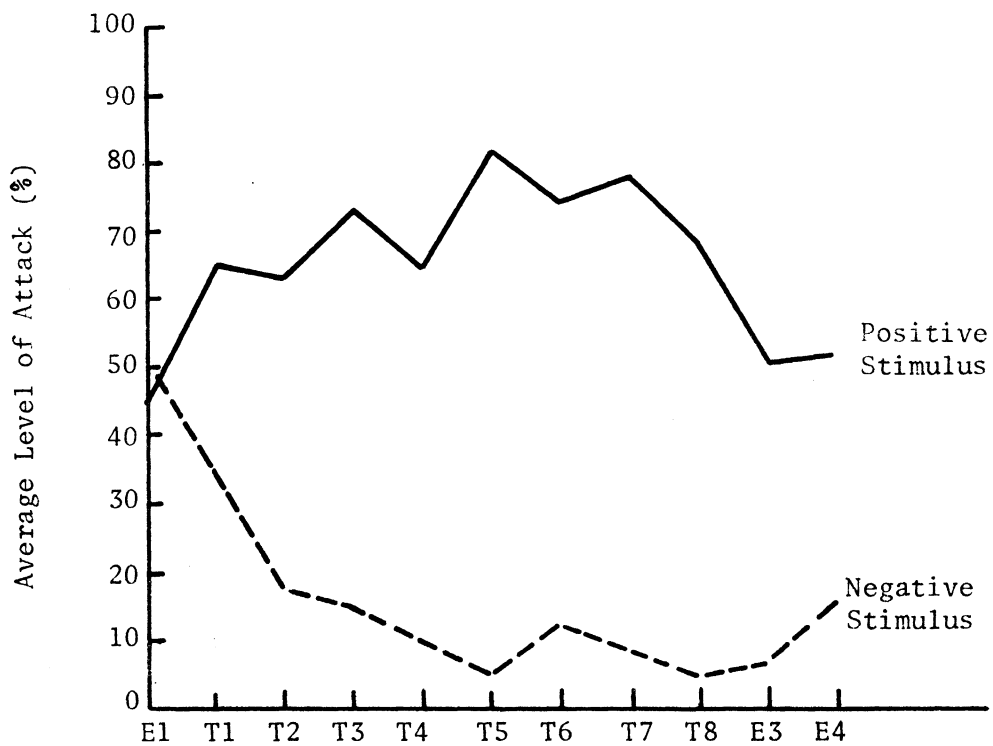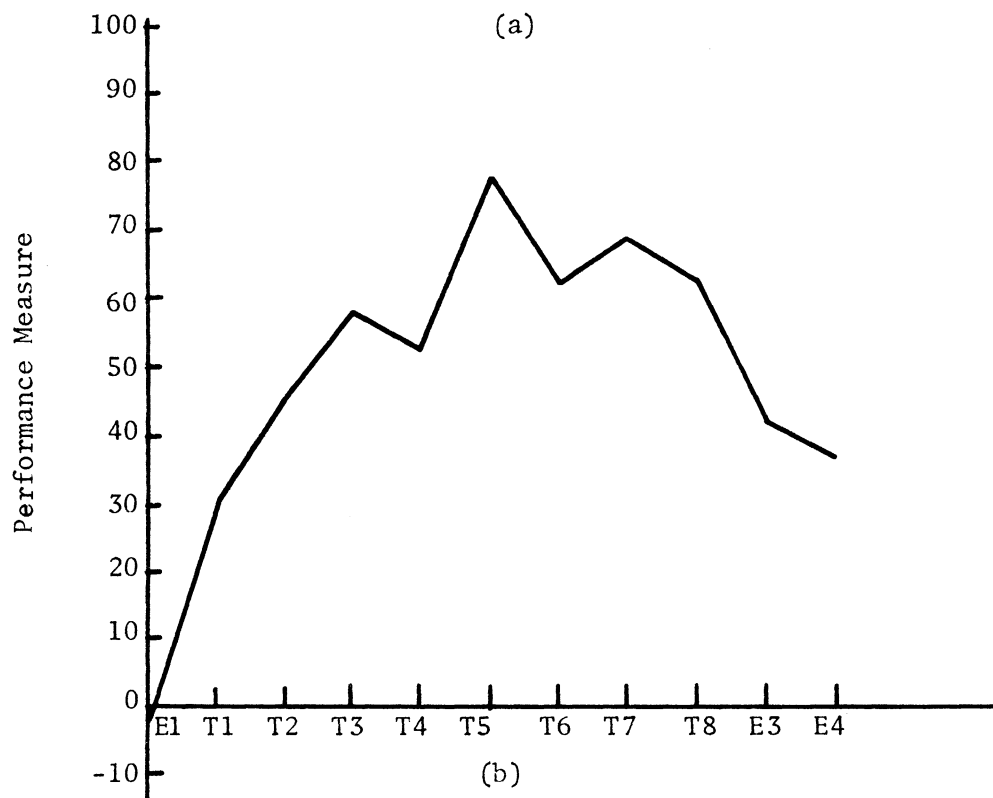
Figure 5.9

Figure 5.10

The parameter settings for this experiment and the following figures are given in table 5.3. The magnitude of the stimulus values of taste and pain for the positive and negative figures were set equal, but the effect of pain is slightly greater than that of taste due to the somewhat slower decay rate for pain (see figure 5.6). A random factor (RC = .75) is also incorporated, and the curves of the next figure (5.9) represent an average of six repetitions of the experiment.

Figure 5.9a shows the performance of the model over the twelve sessions in terms of its percentage attacks on the positive and negative figures. Figure 5.9b shows the corresponding performance measure (difference in these two curves). Notice that the average attack level is about 60% for both the positive and negative figures prior to training, but eventually rises to about 90% for the positive figure and drops to about 20% for the negative figure as a result of the eight training sessions. The performance measure thus rises from around zero to about 70%. With continued training the performance measure asymptotically approaches, but never fully realizes 100%.

Comparable results for a set of nine octopuses[1] are shown in figure 5.10. Conditions of the experiment were about the same, except that there was only one initial extinction session. The training figures used were white and black vertical rectangles. Attacks on the positive figure rise from an original level near 50% to about 75%, while

attacks on the negative figure fall to about 10%. The performance measure thus increases from near zero to about 65%. (The relatively poorer performance in the final extinction sessions may be due to the fact that animals had just received dummy operations. Performance resumed at its previous level during additional training that was then conducted.)

3. Effect of Operations

A central focus of much current work on the octopus is the attempt to elicit functions of the upper lobe structures of the brain, and of the vertical lobe in particular. The fine structure of this lobe has recently been explored by Gray[2] using the electron microscope. Previous work of Young and others in describing its structure and function was discussed in Chapter 2. There it was noted that the paradigm for most behavioral work was the isolation or removal of one or more lobes and the observation of changes in learning that resulted.

One of the primary objectives in designing the model has been to test the hypothesis that deterioration in learning performance following damage to the vertical lobe system results not from the loss of memory units themselves, but rather from interference with the circuits for reading and writing memory. Thus the model incorporates the upper lobes as the feedback variable, Q, which appears as a necessary ingredient in the equations for short-term memory buildup (equations 2 and 3 of table 4.1). The discussion of figure 5.8 in the first section of this chapter pointed out that learning performance is progressively degraded as Q is reduced.

This section takes a look at the effect of QMAX on group performance and compares these results with the behavior of animals from which the vertical lobe has been removed. In these experiments the model is first trained and tested through a set of training and extinction sessions, as before, and then operation is simulated by reducing the value of QMAX, while leaving memory in whatever state it has already achieved through training. A second set of training and extinction sessions is then conducted, and the "post-operative" performance noted. When the data from several repetitions of this experiment have been averaged

together, the result is analogous to the group learning curve for a comparable set of experimental animals.

Figure 5.11 continues figure 5.10. It shows further results in an experiment[1] where nine animals were first trained, then given a "dummy" operation (in which the brain was exposed but nothing removed), and then retrained. Figure 5.11a gives average attack level for the positive and negative stimuli and figure 5.11b shows the corresponding performance measure as before. Figure 5.11c, which plots the ratio of mean performance to standard deviation of the performance distribution for the nine animals, gives an indication of the level of significance to this learning. Notice that the standard deviation is always less than the mean after the first training session. Performance tends to approach some final value of around 70% for these animals, and the dummy operation appears to have no lasting effect on performance.

Figure 5.12 is an analogous set of curves for a group of eleven animals with vertical lobe removal.[1] The difference is striking. The effect of the operation seems to be immediate and permanent. Performance drops drastically following operation, and never improves very much with retraining. Positive attacks remain above negative attacks (except in session E4), but by a small amount. The ratio of mean performance to dispersion in figure 5.12c is always less than one after operation, even though the performance measure itself does remain positive.

The results obtained when the corresponding experiments were run with the model are shown in figures 5.13 through 5.16. The first of these shows the case where QMAX remains at one, corresponding to a "dummy" operation. It is apparent that mean performance continues to improve throughout and remains much greater than the standard deviation of the four repetitions of the experiment, as it should. The next
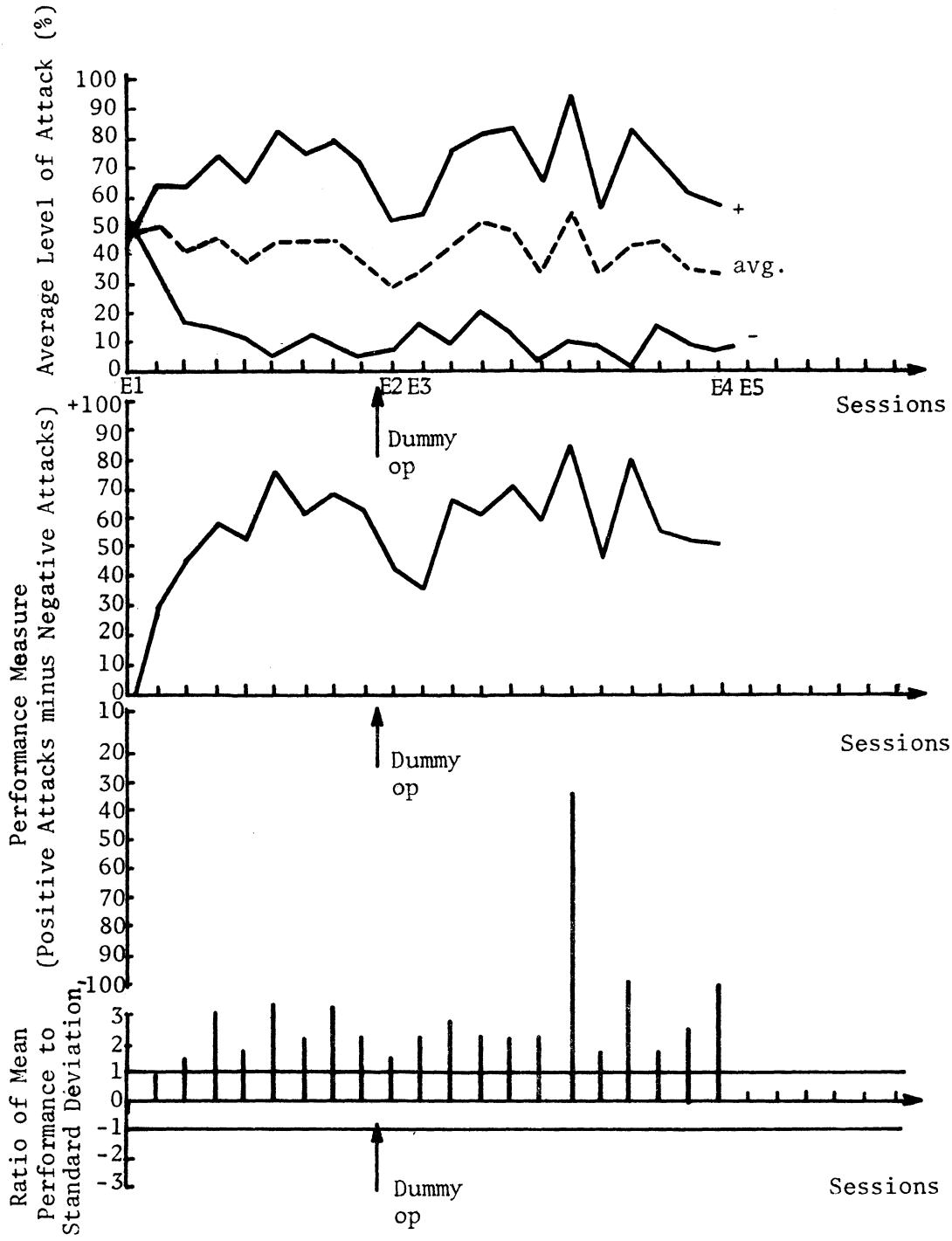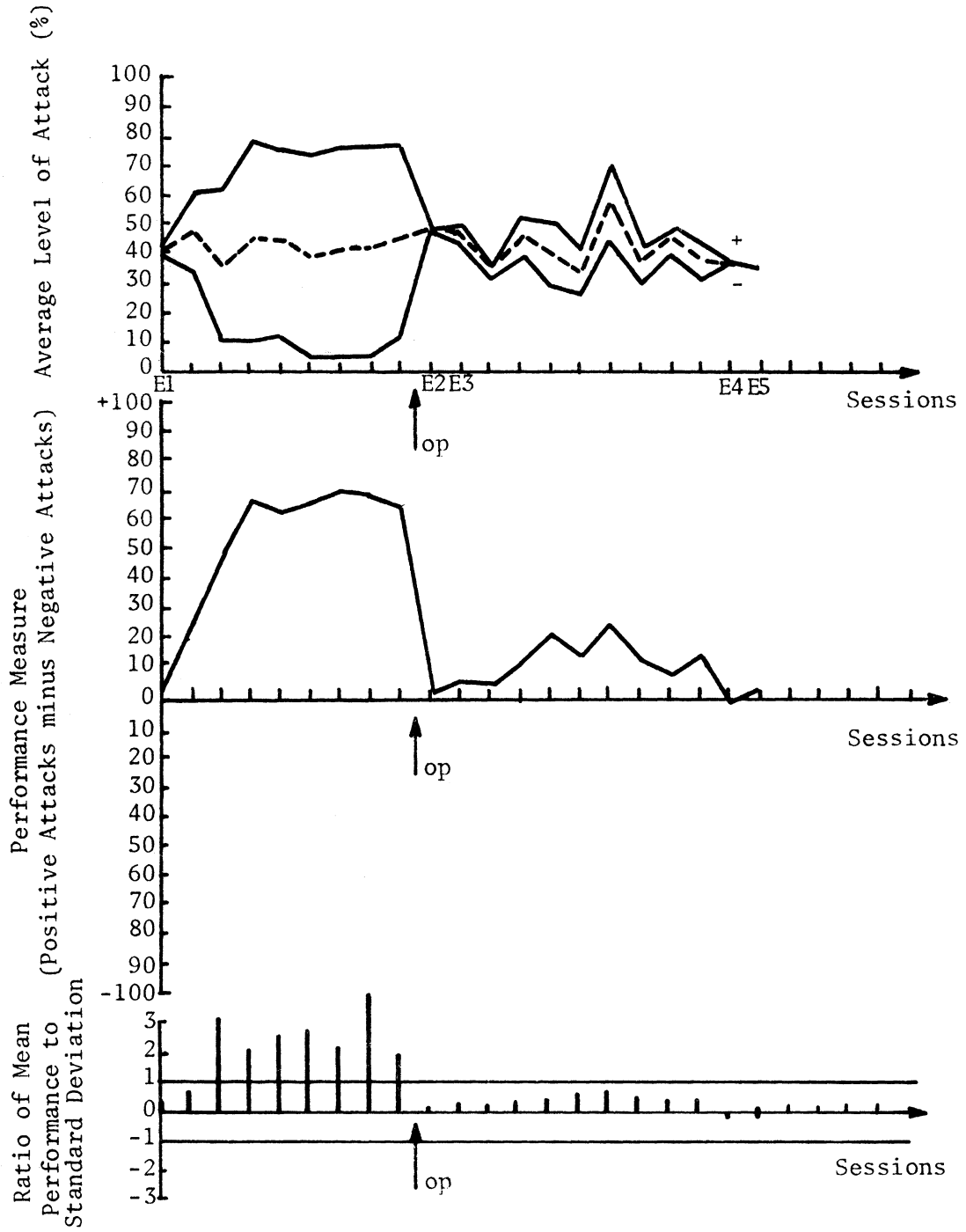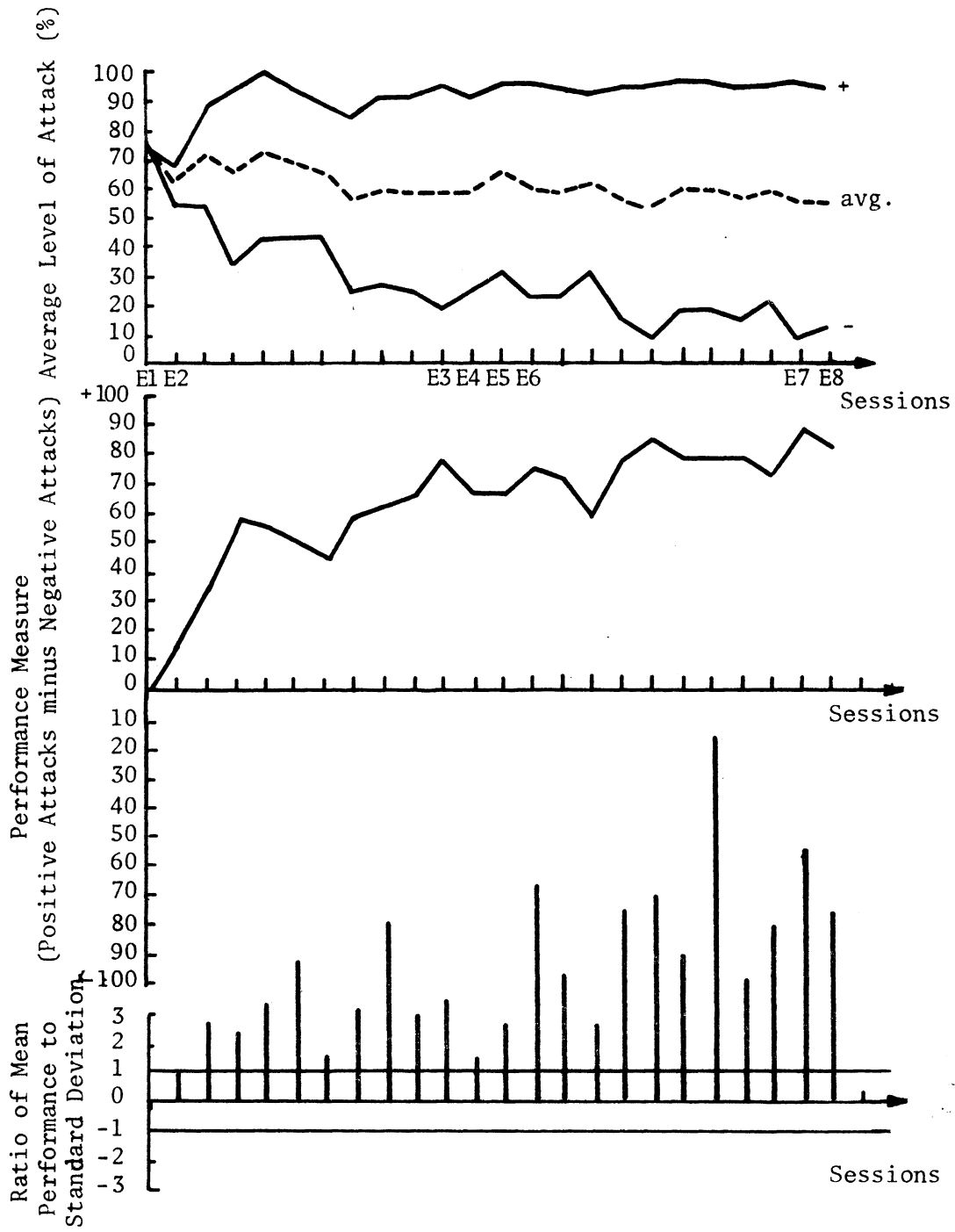
Figure 5.11
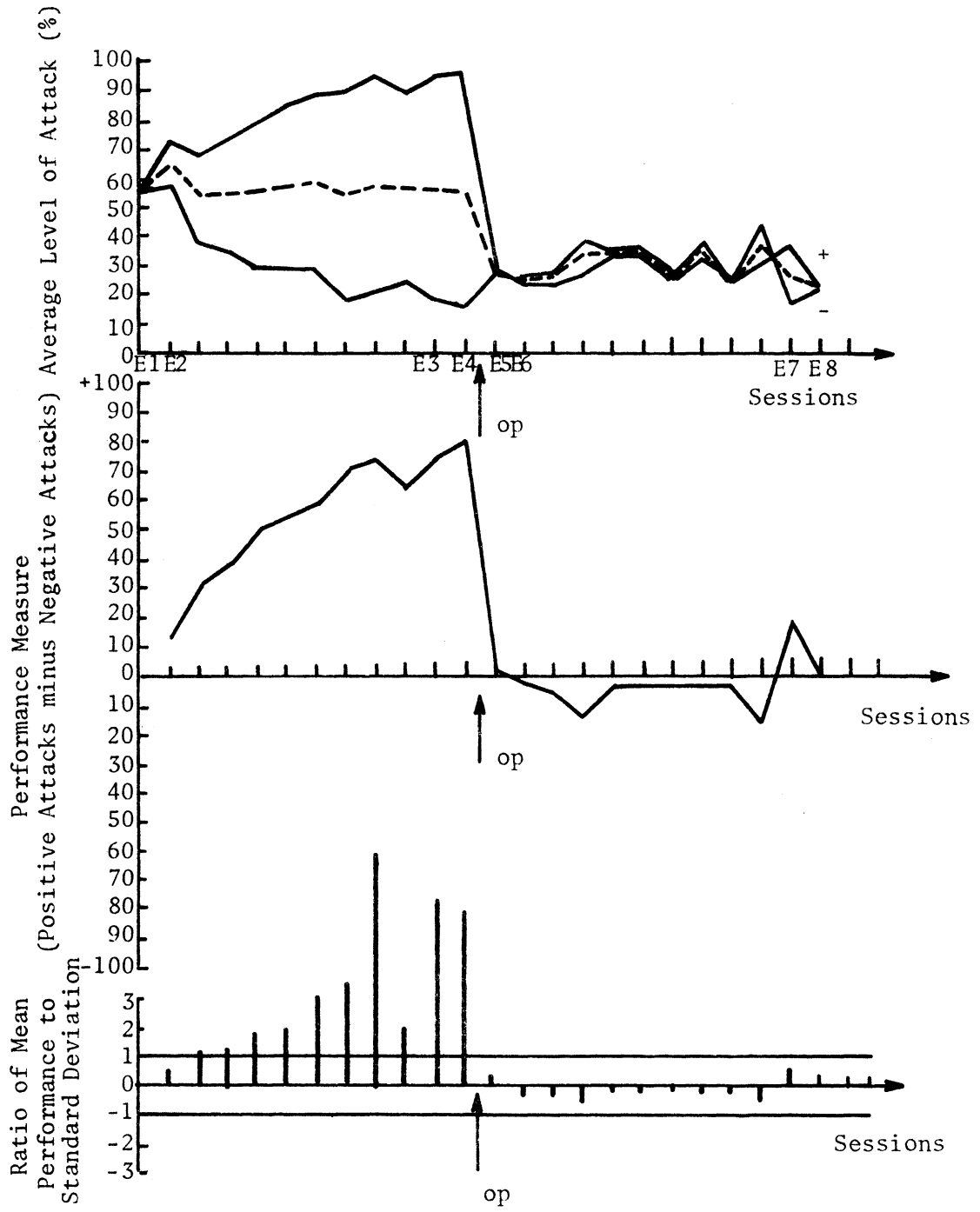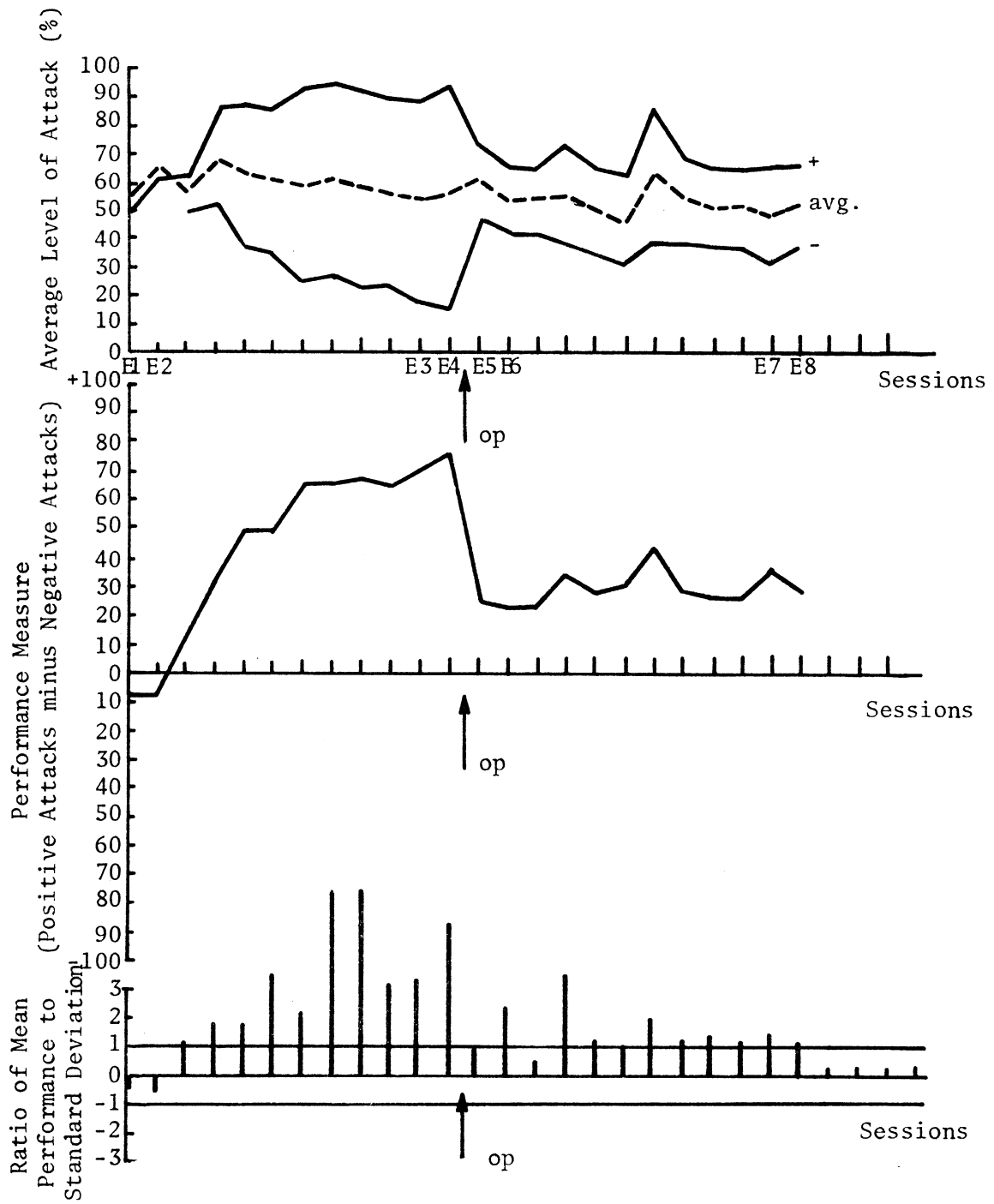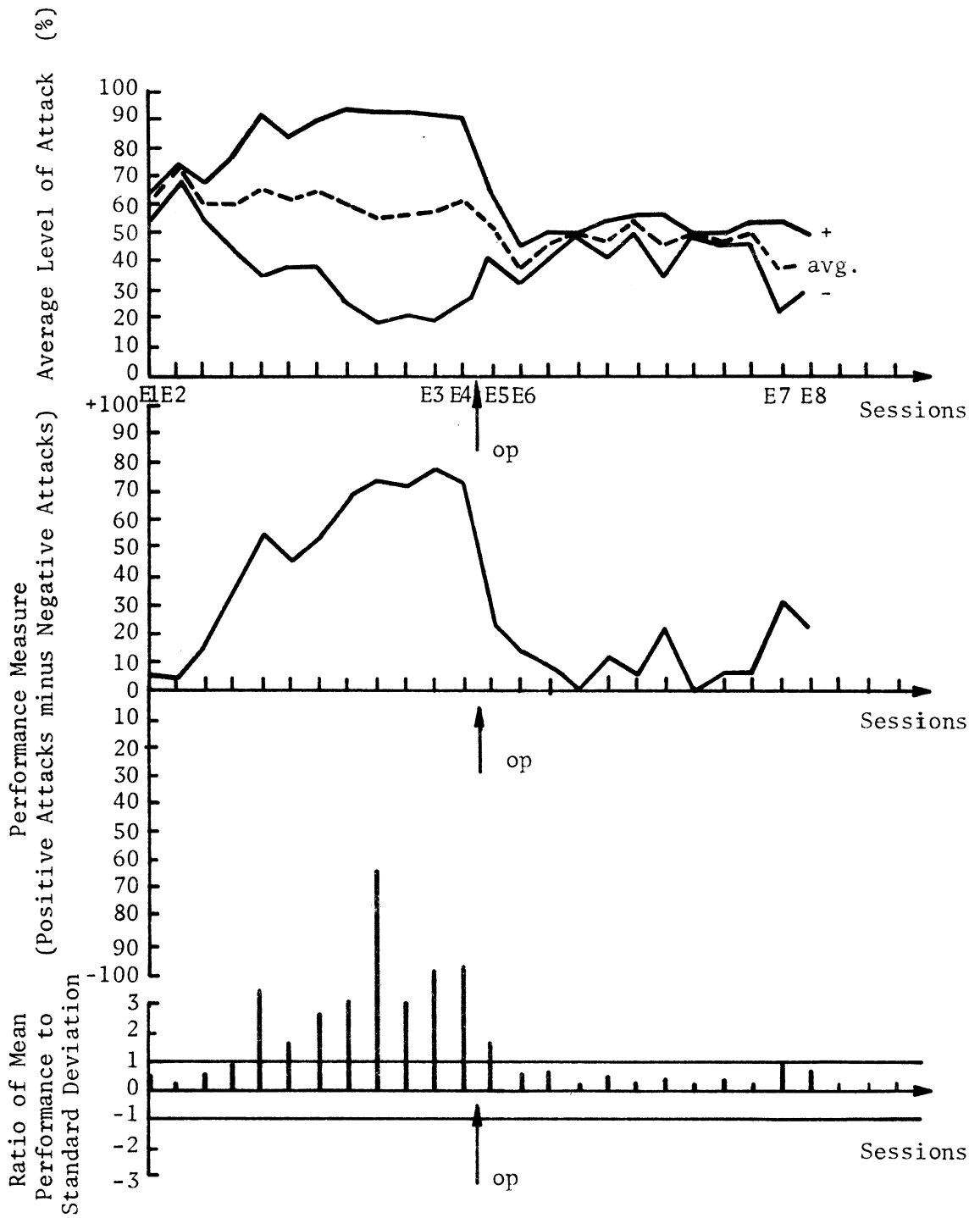
Figure 5.12

Figure 5.13

Figure 5.14

Figure 5.15

Figure 5.16

figure shows what happens when operations are simulated by setting QMAX to zero after the first set of training sessions. Performance immediately drops to a random level near zero, and the dispersion is much greater than the mean for the eight simulated animals, in spite of the fact that memory levels remain the same. The memory value is still there, but it can no longer be "read-out" in the absence of Q.

An intermediate condition is shown in figure 5.15. In this figure, an operation involving incomplete removal is simulated by setting QMAX to some small value ofter operation (QMAX = .3). The result is an intermediate case where mean performance falls, but not entirely to zero. It remains greater than dispersion for the six repetitions, in fact (except for T9), and seems to improve slightly with retraining.

When an experiment is conducted with part of the simulated animals having complete removal (QMAX = 0) and some having partial removal (QMAX = .3), the combined results are shown in figure 5.16 for six simulated animals. In this figure the mean performance remains positive but less than the standard deviation (except for E7). These results are qualitatively reminiscent of those in figure 5.12 for eleven real animals with vertical lobes removed. It thus seems reasonable that the degradation in performance as a result of operation could be due to interference with the read-write mechanism of memory, as hypothesized.

4. Summary

Section 1 surveys basic operating characteristics of the model in the absence of random influences, and follows the time course of variable values during a single encounter with the environment. Three typical model responses are examined, corresponding to positive, negative, and neutral stimuli. It is seen that the ultimate effect of an encounter

with the environment is a small memory change in AM or RM after return to quiescence. Random influences are important for insuring continued sampling of the environment as experience is accumulated. Latency is shown to decrease rapidly with experience. Figure 5.8 shows the effect of QMAX on memory changes, and it can be seen there that interference with this pathway through simulated operations will drastically reduce the ability to learn.

The next section compares statistical data from the model with group performance data of real animals. Learning takes about the same form in both cases, except that real animals more quickly learn to avoid the negative stimulus, and attack a bit less often generally. A short discussion of performance measures and methods for combining individual scores concludes that the ones adopted are simple but probably adequate.

Operations were simulated in section 3 by reducing the value of QMAX while leaving memory contents unaltered. The effect of simulated operations was seen to be much like the effect of real operations. It was concluded, therefore, that the deterioration in performance observed when the vertical lobe is damaged or removed could well be due to interference with the mechanism for reading and writing memory, rather than from loss of memory itself.

CHAPTER SIX

ADDITIONAL MODEL CHARACTERISTICS

1. Attack and Retreat Probabilities

The previous chapter dealt with model performance in certain learning

situations and with comparisons to animal behavior. This chapter continues

that development by looking at additional characteristics of the model

and further experiments having to do with un-learning, time interference,

and stimulus overlap.

In the absence of random influences and hunger, a single encounter with

a positive stimulus would be sufficient to tip the balance toward subsequent

attacks on such a stimulus. Likewise, a single encounter with a negative

stimulus would also be sufficient to ensure subsequent retreats by the model.

A plot of probability of attack vs memory value thus rises from zero to 100%

as memory goes from negative to positive values. When a random factor

(RC) is introduced, however, this curve of attack probability vs memory

value begins to depart from such a sharp step, and figure 6.1 shows

its empirically determined shape with RC = .7. Each circle represents

approximately 400 trials and gives the percentage attacks at that value

of memory. Crosses give 100 minus percentage retreats. To a first

approximation the curve shows a linear rise over the range from -.30

to +.30. The difference between the attack and retreat curves near

zero represents the percentage of trials in which the stimulus was removed

after twenty seconds because no action had been taken in that time, and

can be understood in terms of the latency plot of figure 5.7, which shows
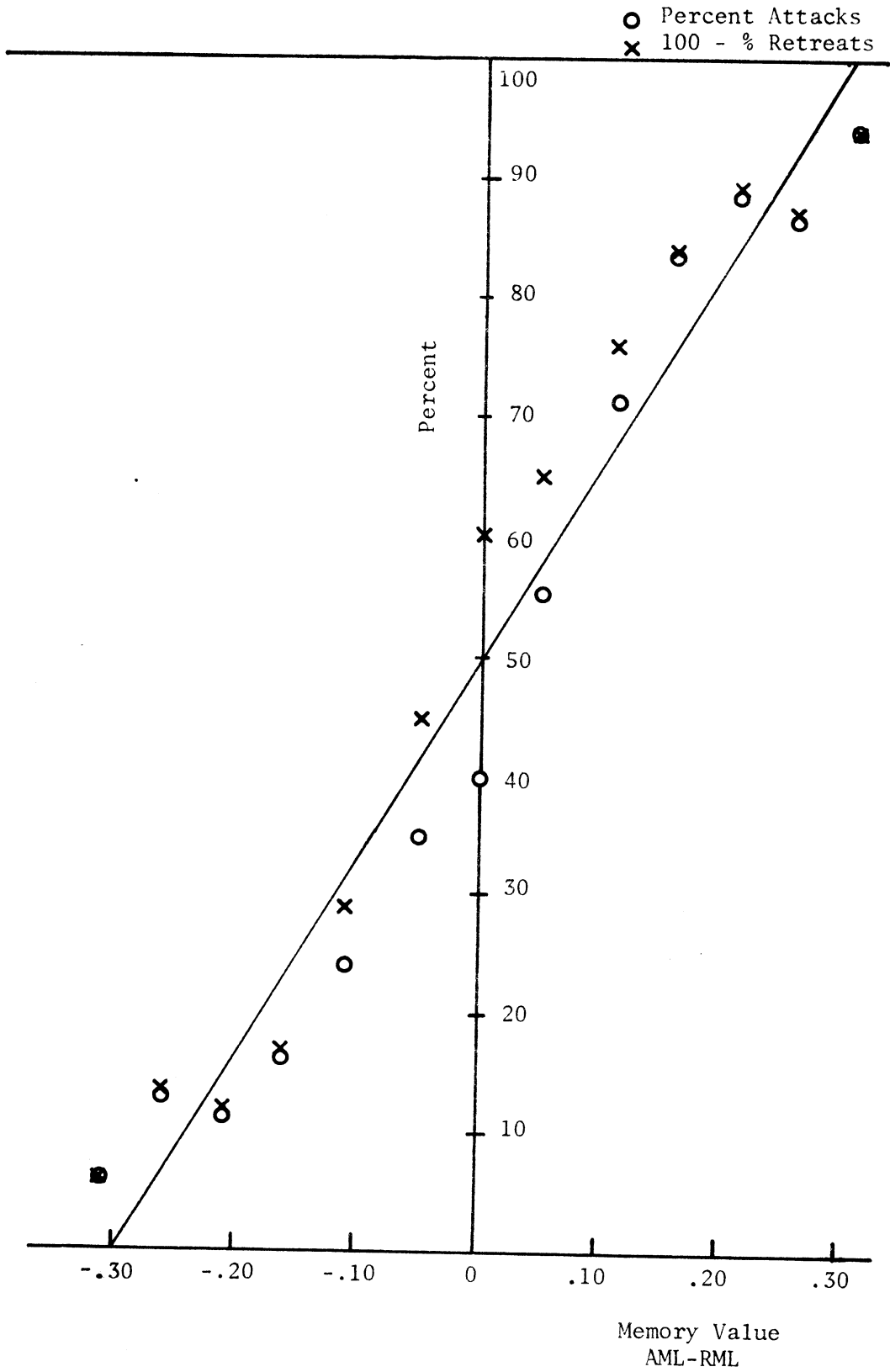
that latency drops off rapidly with learning.

Figure 6.1

2. Un-learning

When a long-term memory value of about .30 (AML - RML = .3) has been attained, the model may be said to be fully trained in the positive direction, based on figure 6.1. If training is then reversed, so that it receives PA input for attacks on the formerly positive stimulus, then it should begin to "un-learn" its former training. This is somewhat analogous to experiments in which animals receive shocks for attacks on crabs placed in their tanks, and eventually learn not to attack them. Figure 6.2 shows the results of this experiment with the model. Attack levels near 95% occur in the two initial extinction sessions, but drop throughout a series of training sessions to a level near 50% in the two final extinctions. A similar curve obtained when the initial memory value was set to .5 is also shown for comparison (dotted curve).

3. Repeated Reversals

Figure 6.3 shows the performance of the model when training is repeatedly reversed in this fashion for one simulated animal. Real animals show an ability to reverse repeatedly, although there is a residual bias in favor of the direction of initial learning, and an eventual disintegration of performance. In this experiment, (which consists of an extinction session, eight training sessions, another extinction and then reversal) it is clear that the model was never really able to overcome its initial training. At the end of the first two reversal series attacks on the two stimuli are about equal, but attacks on the initially positive stimulus always exceed attacks on the initially negative stimulus elsewhere.

Figure 6.4, which gives the value of long-term memory after each
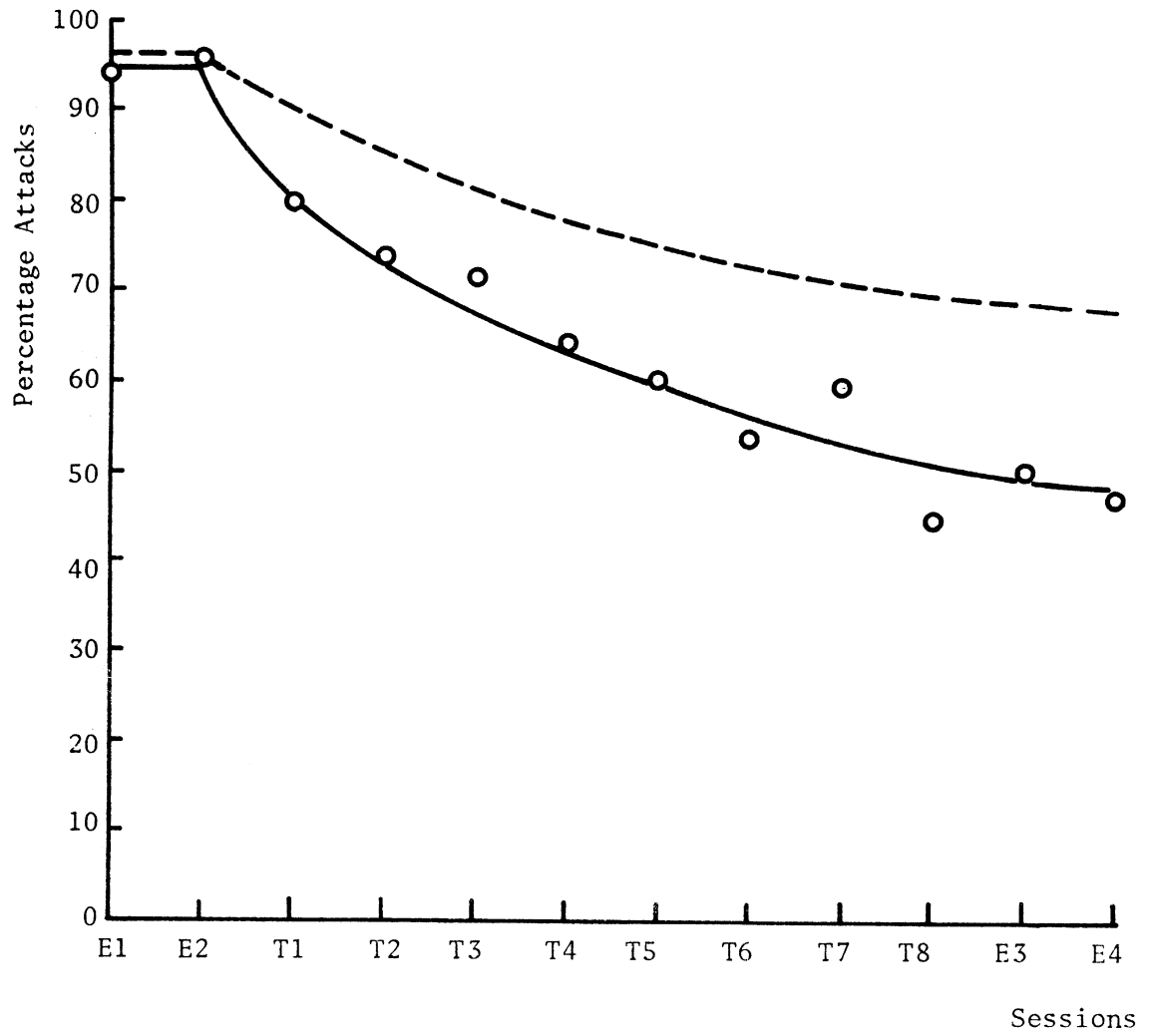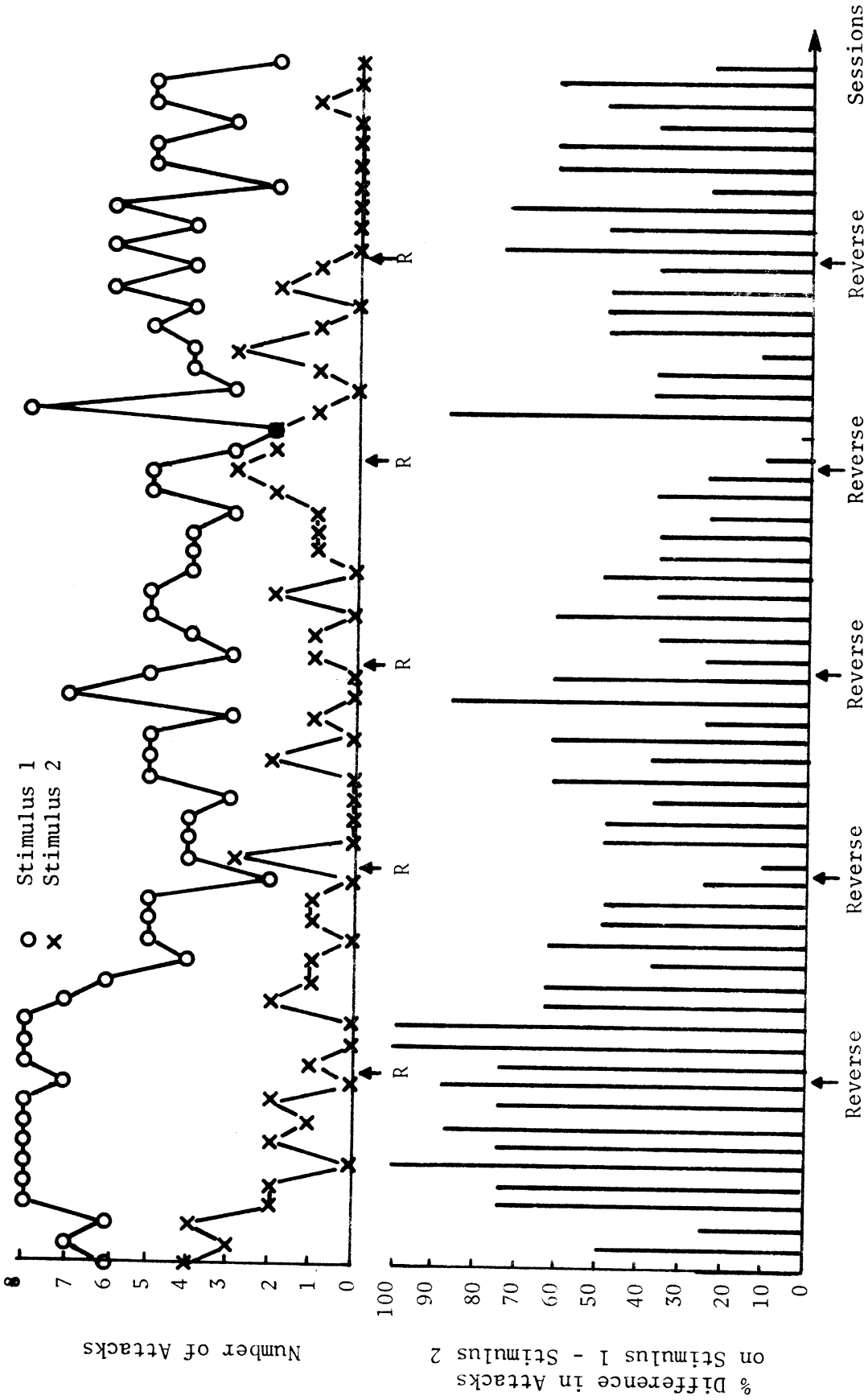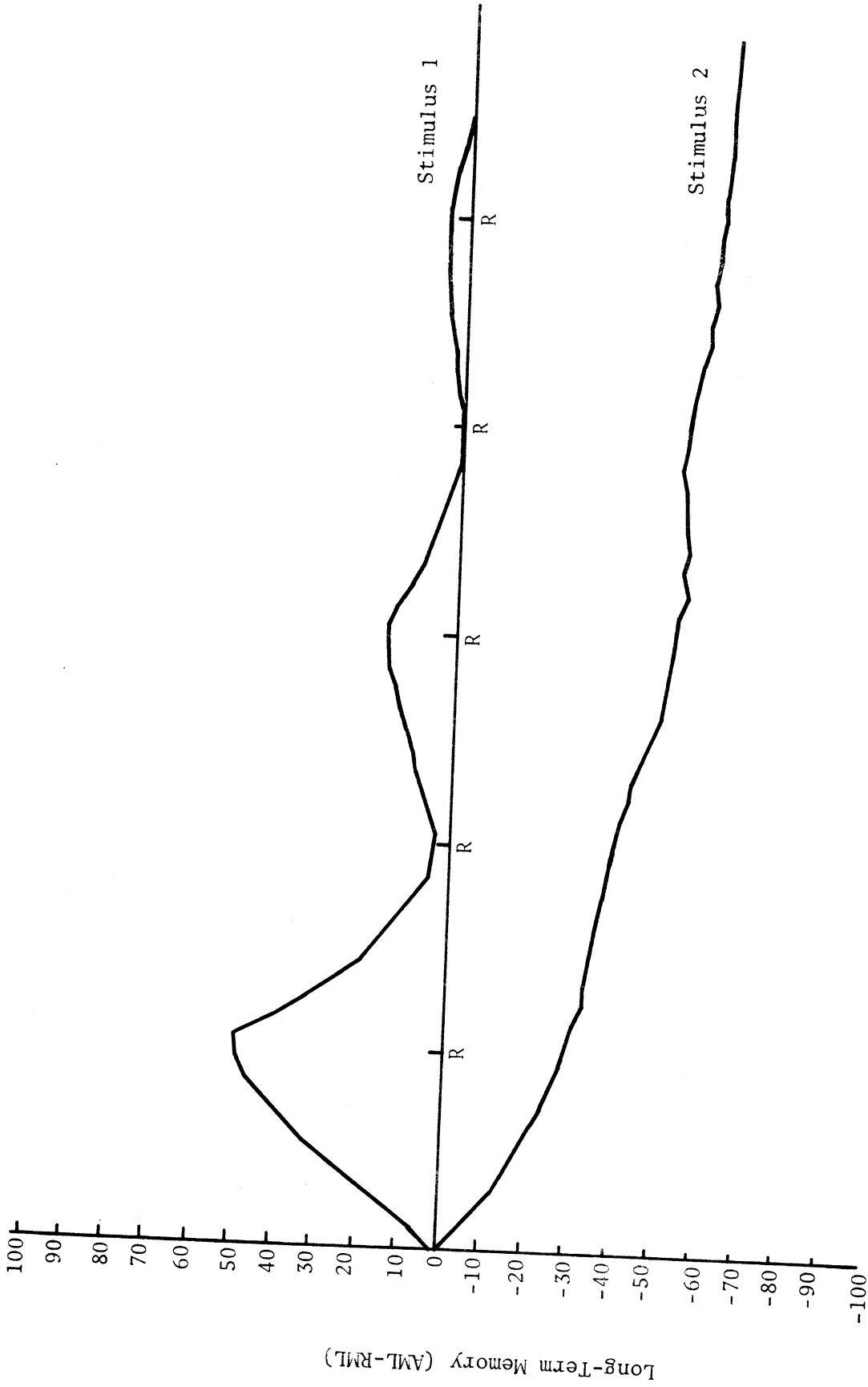
Figure 6.2

Figure 6.3

Figure 6.4

session, shows the reason for this failure. The memory associated
with stimulus 1 (initially positive) increases and decreases throughout
each reversal series by about the same amount. The memory associated
with stimulus 2, on the other hand, continues to decrease, even through
those sessions when stimulus 2 is trained to be positive.

The explanation for this effect is that once the model has been
trained negative the attack level becomes very low, and this small
number of attacks during positive reversal periods is then not sufficient
to overcome the effect of its initial training. Compounding this difficulty
is the problem of too much read-in to negative memory from simply being
in the retreat state (refer to the discussion of figure 5.1). The
principal difficulty with the model, however, is that results of the
recent past do not have sufficient effect.

Notice that the highest level of memory attained for stimulus 1
in each positive portion of the reversal cycle is progressively less
than that attained in the previous cycle. This comes about because
attack memory (AML) increases during the positive portion of the cycle
and retreat memory (RML) increases during the negative portion of the
cycle. The difference thus shows a periodic fluctuation. As each
approaches its maximum normalized value of one, however, this difference
becomes smaller on each cycle. The result is that if the memory associated
with stimulus 2 were behaving properly, as discussed above, then the overall
performance would follow the cycle of reversals. The maximum performance
on each cycle would deteriorate just as that of real animals does, however.

4. Conditional Learning

One of the earliest experiments conducted by Boycott and Young
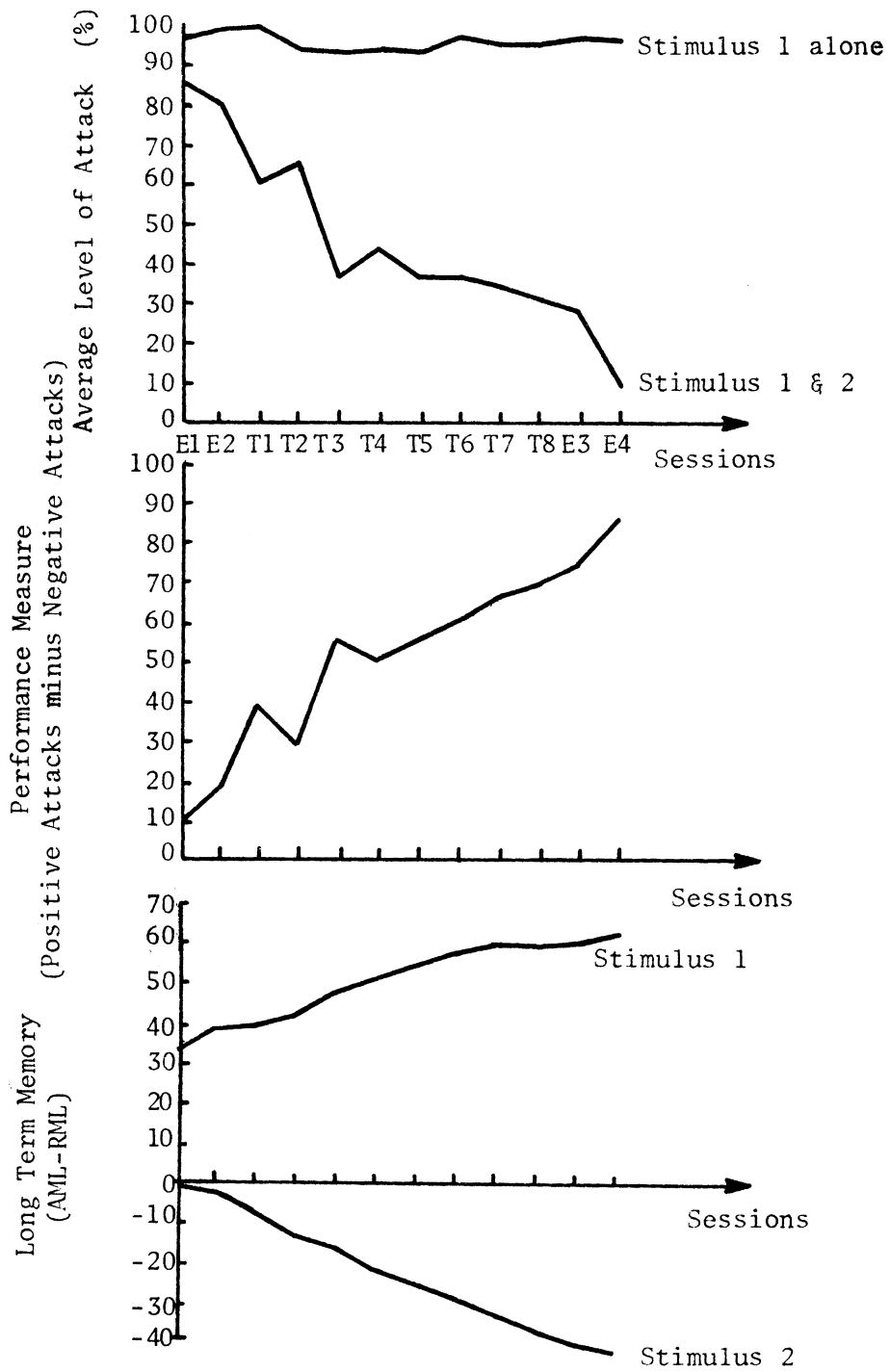was to teach animals not to attack crabs presented along with some

Figure 6.5

special signal, such as a white square. The animals soon learned that they could catch crabs given alone, but that they would receive a mild electric shock for attacking crabs when the white square was present in their tank at the same time.

This suggests an analogous experiment with the model, and figure 6.5 shows the results when this experiment is performed. (Values shown in figure 6.5 are averages over twelve simulated animals with equal numbers at hunger levels of 0, .5, and 1.) Attacks on stimulus 1 alone continue at a high level throughout the experiment, but attacks on the combination fall to about 20% during training. The memory value of stimulus 1 does increase somewhat from its initial value of .3, but when combined with the increasingly negative value of stimulus 2, yields a low memory value for the combination.

5. Overlapping Stimuli

A generalization of the crab/white square type experiment is the case of multi-component, overlapping stimuli. Everything which has been done so far has assumed that for modeling purposes the stimuli could be considered to consist of orthogonal elementary components, in the Hubel-Wiesel sense. In this case, the average output from the set of mnemons comprising each stimulus could be represented as the output of a single "average" mnemon, and this has been done throughout the text thus far. In the non-orthogonal case, however, some mnemons from each set may also be present to some extent in the other set, so that there is an "overlap" between stimuli.

Figure 6.6 shows the results of a standard experiment averaged over two simulated animals when such stimulus overlap is present.
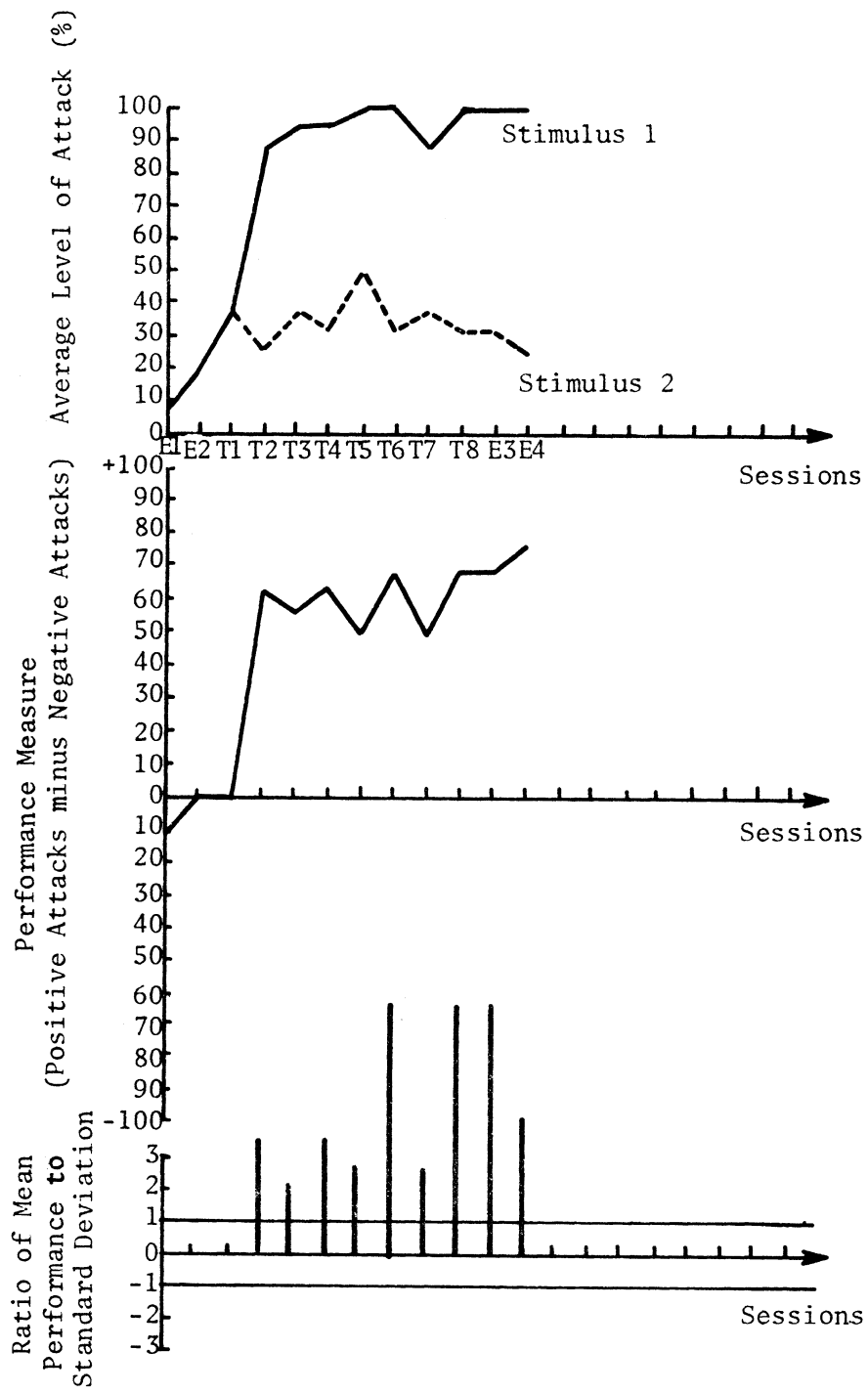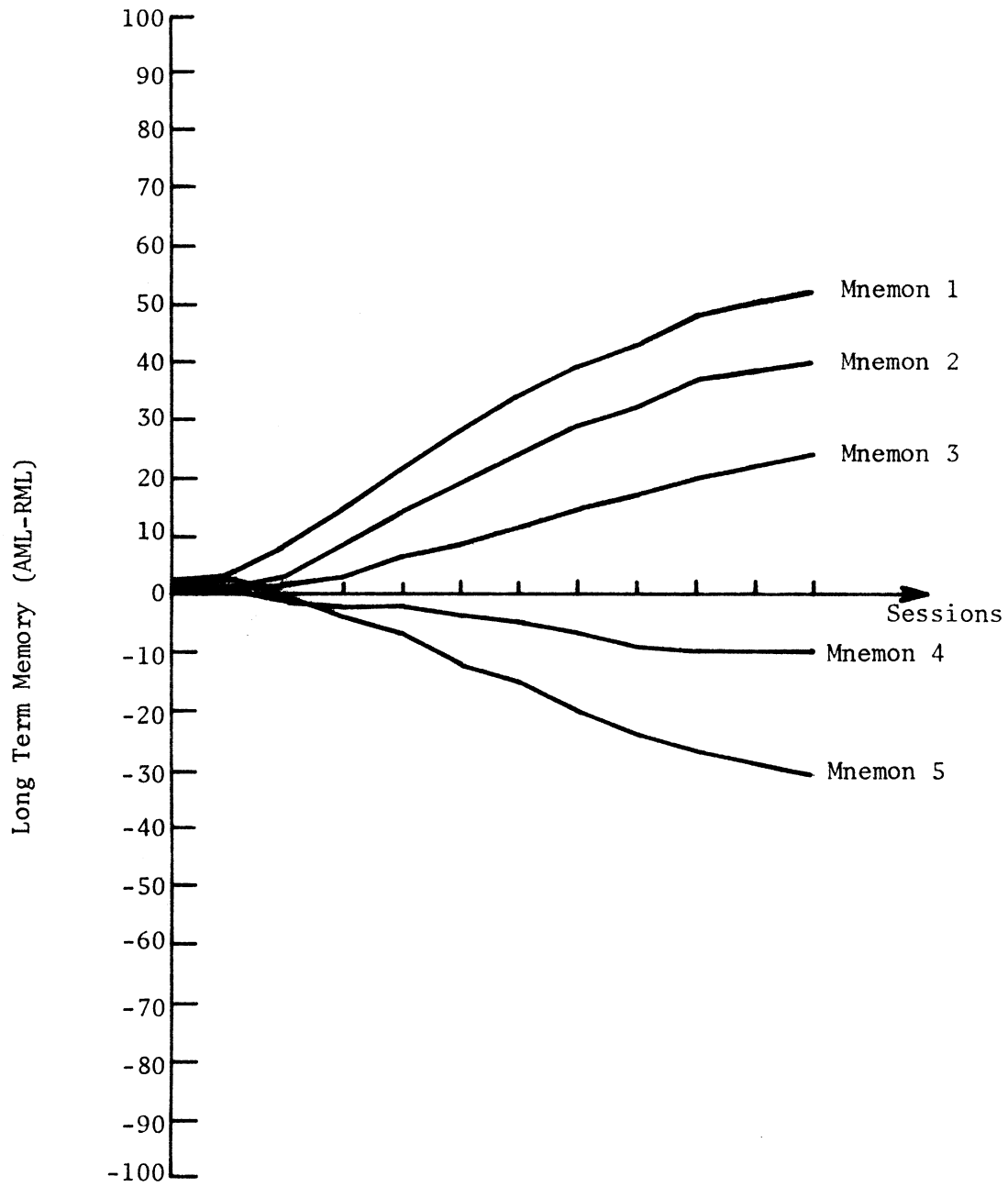
Figure 6.6

Figure 6.7

Each stimulus is assumed to activate five mnemons with input constants of $_1C_i$ = 1,.75,.5,.25,0 and $_2C_i$ = 0,.25,.5,.75,1 respectively. All mnemons overlap except the first and the fifth, therefore, and will be influenced to some extent by both taste and pain. Figure 6.7 shows the memory values acquired by each of the mnemons during training. The overlapping mnemons acquire intermediate values related to the extent of their overlap, as expected, but with some bias in the positive direction. (This results from the fact that there are more attacks on the positive stimulus than on the negative one.)

Figure 6.7 gives some insight into what might be expected from additional experiments with complex stimuli. A different linear combination of mnemons 1 through 5 after training (corresponding to some new stimulus object) could produce attack, retreat, or indecision, depending upon the relative values of the weighting factors. For example, suppose some new stimulus consisted of an untrained mnemon 6 in mild association with the somewhat negative mnemon 4. The model's reaction to this new stimulus could be expected to be mildly negative.

6. Time Interference

In all of the experiments which have been described thus far, the interval between stimulus presentations during each session has been set at five minutes (TMAX = 300). This has given sufficient time for slow changes to memory through the action of the upper lobe read/write mechanism (Q) to be completed, for the time constants chosen, and in most cases for the model to reach a quiescent state before each new

stimulus presentation. As this interval between presentations is reduced,
however, interference effects begin to appear, because the writing into
memory of results of the previous encounter has not been completed
by the time the next presentation is made.

Figures 6.8 through 6.12 show a series of standard experiments in
which the period between trials was set at 6, 4, 2, 1.5, and 1 minute
respectively. (Note: Different numbers of simulated animals were
averaged together for each figure, so there is some variation in
smoothness between the curves. Figure 6.8 represents an average of
two, figure 6.9 an average of six, figures 6.10 and 6.12 an average
of five each, and figure 6.11 an average of four.) Notice the progressive
deterioration in performance as TMAX is reduced. The results in figures 6.8
and 6.9 for TMAX = 360 and 240 closely resemble the TMAX = 300 results
seen in various figures previously. With inter-trial times less than
this, however, the performance deteriorates rapidly.

The explanation for this behavior can be seen from the set of memory
curves, which show that the memory value (AML - RML) associated with
the positive stimulus after each session tends to decrease as TMAX is
reduced. For TMAX less than 90, this positive memory becomes negative,
in fact. The net result is that attacks on the positive stimulus fall.
The effect of time interference in the model, therefore, is that
read-in to positive memory is reduced. This comes about in the following
way. When the inter-trial time is small, short-term conditions following
an attack on the positive figure make an attack on the negative figure
quite likely. This results in some negative read-in to the mnemon
associated with the positive stimulus, however, because this positive
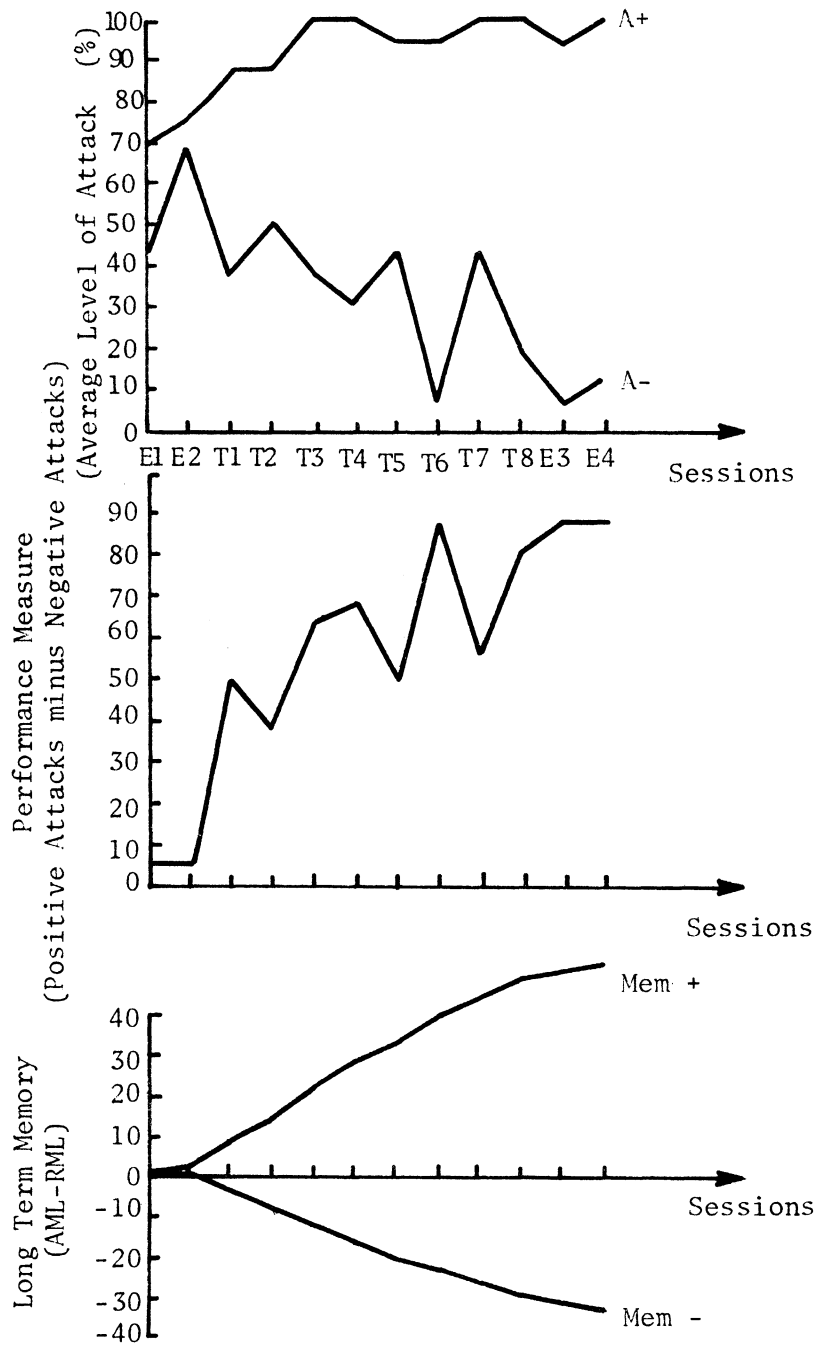mnemon is still partially active. The less the inter-trial time, the

TMAX = 360
No. = 2

Figure 6.8

TMAX = 240
No. = 6
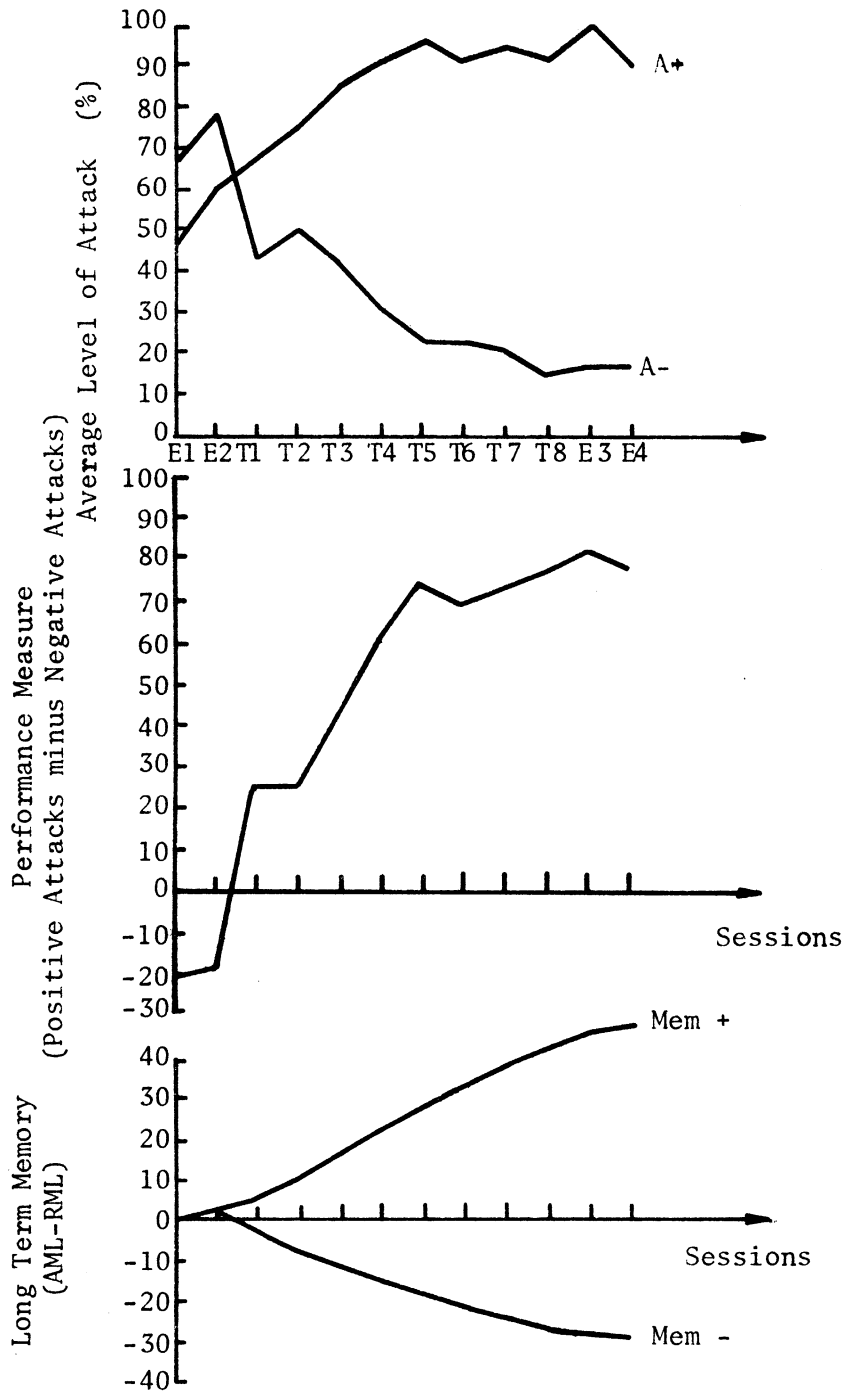
Figure 6.9

TMAX = 120
No. = 5

Figure 6.10

TMAX = 90
No. = 4

Figure 6.11

TMAX = 60
No. = 5
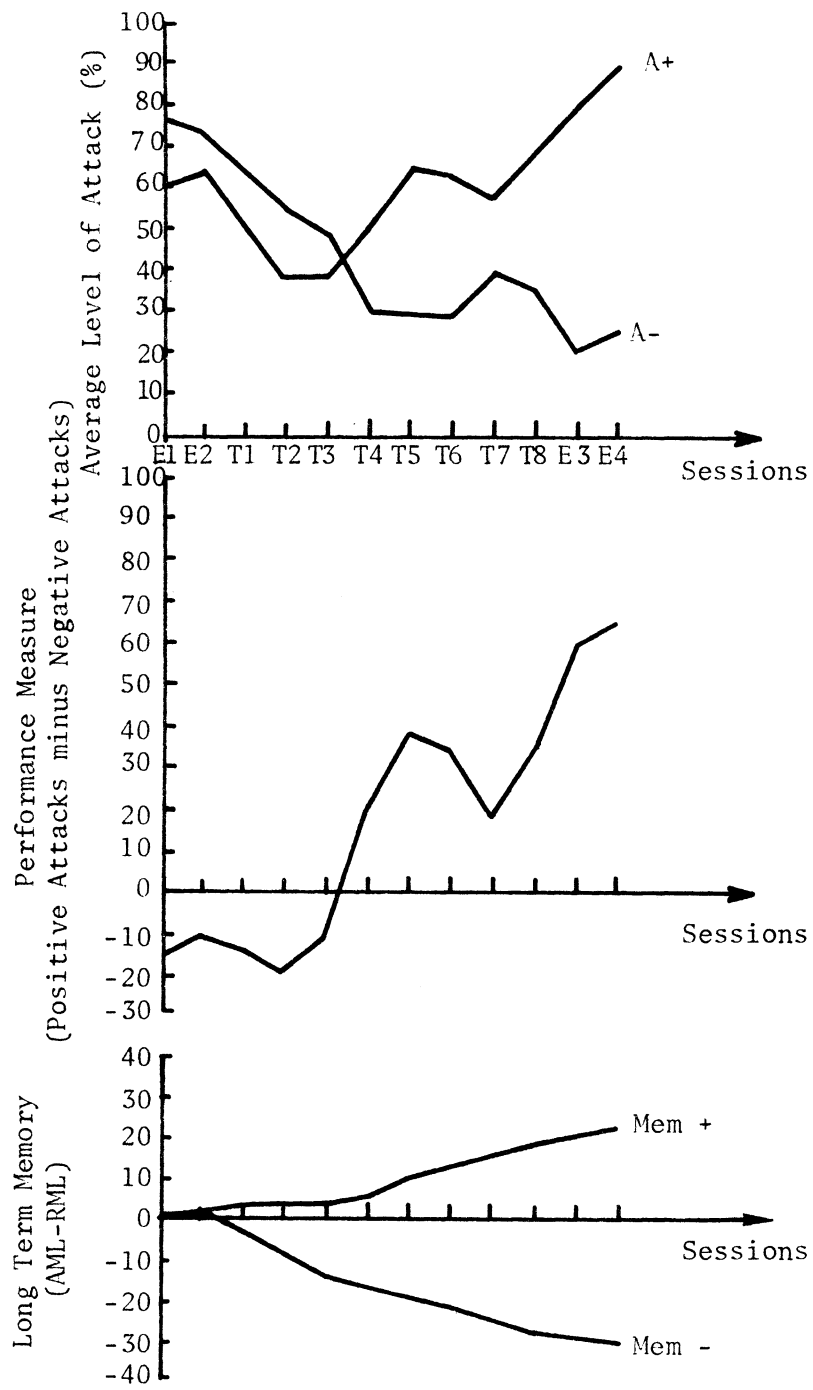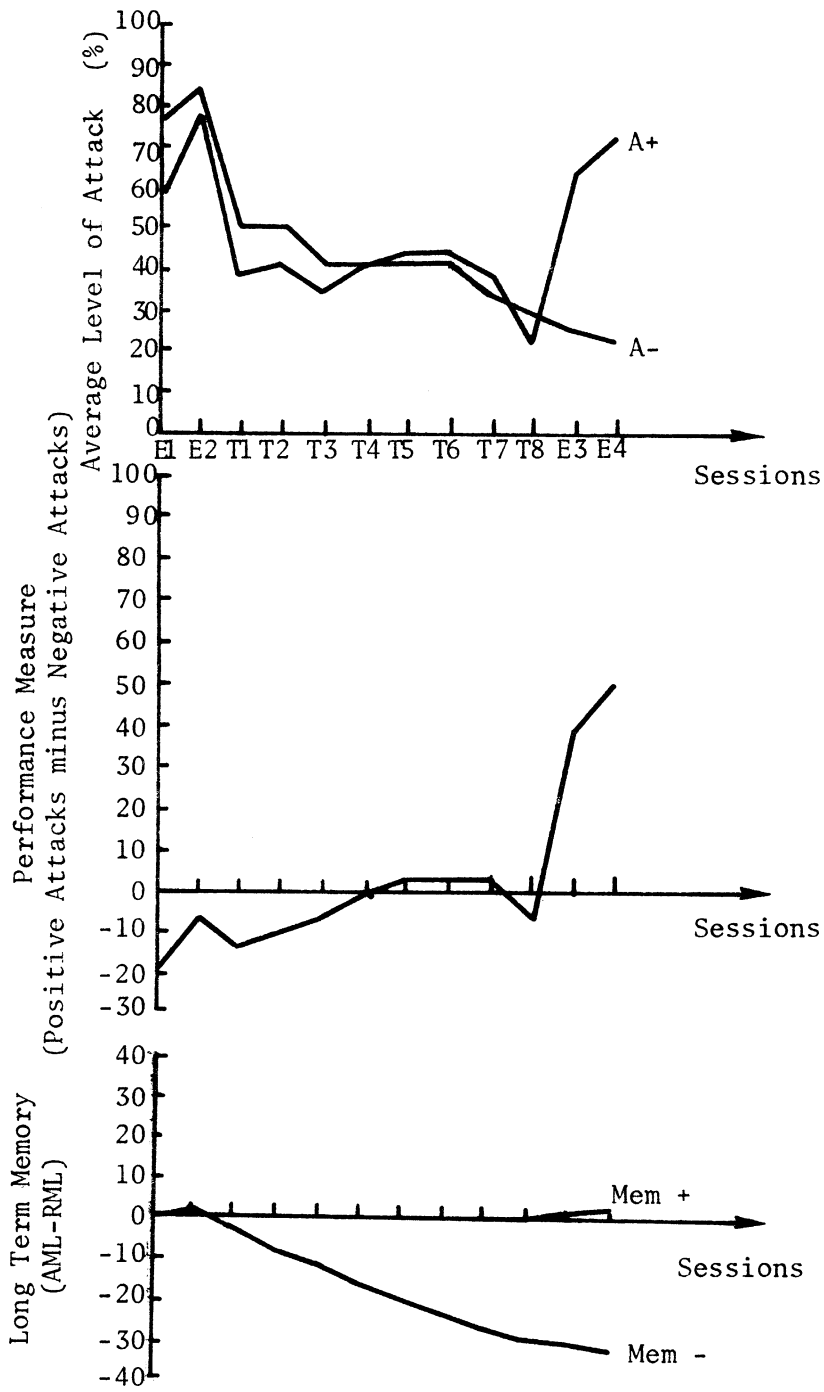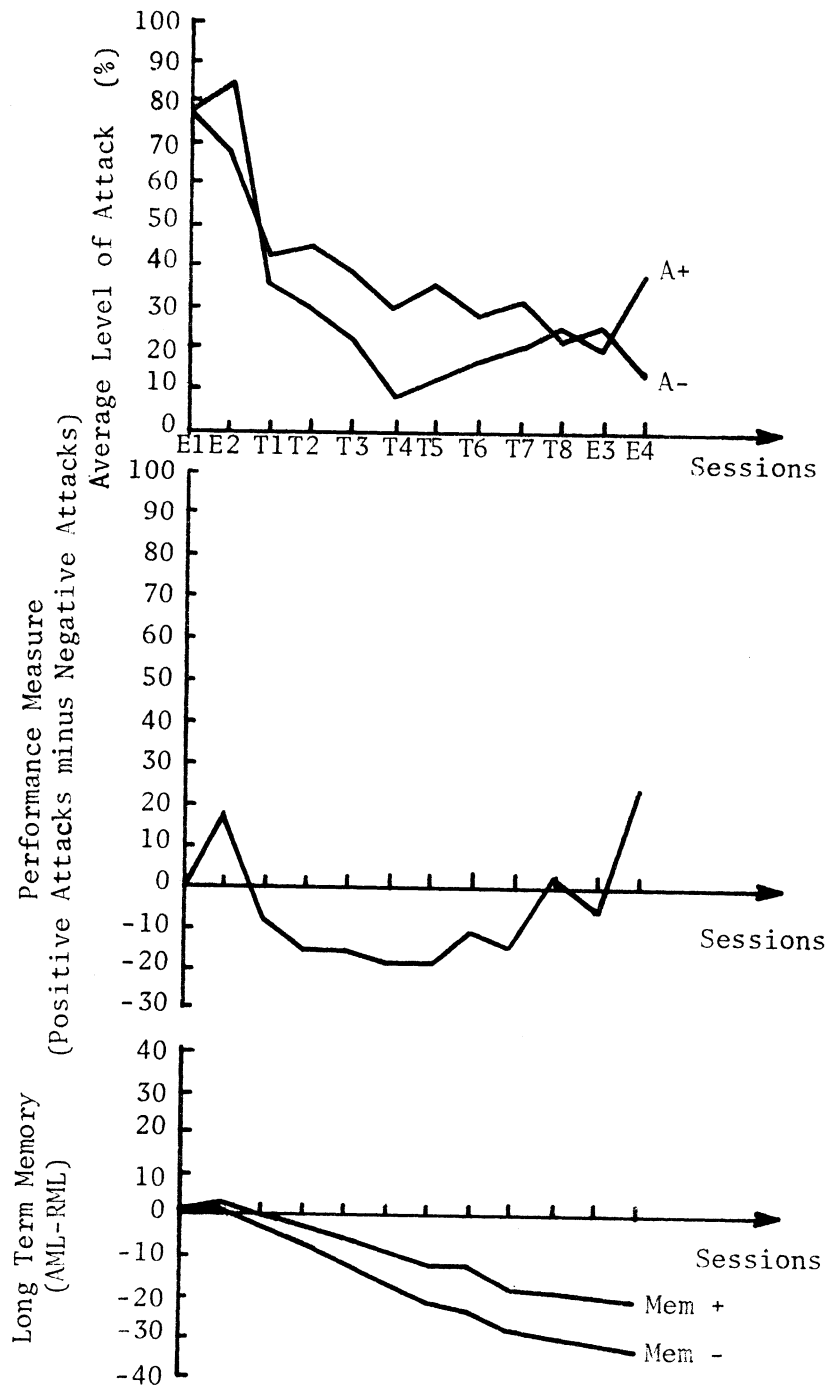
Figure 6.12

greater this effect will be. The exact opposite occurs in the inverse situation, however. An attack on the negative stimulus will make a following attack on the positive stimulus *less* likely, so there will be little interference to learning in the negative mnemon. The net result, which can be seen in figures 6.8 through 6.12, is a tendency for both mnemons to become negative as the intertrial time is shortened.

7. Summary

This chapter continues the development begun in chapter 5 of describing a series of experiments conducted with the model. The first section describes an empirical determination of attack probability as a function of memory value. Sections 2 and 3 deal with reversal-learning situations and shortcomings in the model which these reveal. Sections 4 and 5 describe experiments with multi-component, non-orthogonal stimuli, which give some indication of pattern recognition capabilities in the model. The final section deals with interference effects resulting from inter-trial times which are small with respect to memory time constants.

CHAPTER SEVEN

DIRECTIONS FOR FURTHER MODEL STUDIES

1. Discussion of Results

This chapter summarizes the results obtained from the series of experiments described in chapters 5 and 6, and suggests directions for further work.

Because the model contains a random factor, it is often necessary to repeat an experiment several times to produce curves which are smooth enough to indicate model behavior satisfactorily. Additionally, there is a wide disparity between the basic time scale over which changes occur in the model and periods encompassed by actual animal experiments. Each iteration of model update is taken to represent a time frame of approximately one second, but the experiments to be simulated may last several hours per day for a half-dozen days or more. Once the initial phase of detailed study into operation of the model was completed, therefore, it became important to keep the model as simple as possible and to seek ways in which unnecessary computation could be reduced, such as by using separate update procedures at quiescence.

Even within such constraints the model appears to exhibit many desirable characteristics. The curve of figure 5.9 shows typical performance in a standard learning experiment, and figure 5.13 shows average performance over four repetitions of the experiment. These results are generally similar to learning curves for real animals, except that the attack level is too high and the model is slower in learning to retreat. When operations on the vertical lobe structures are simulated by reducing the value of QMAX, performance collapses for the model in proportion to the extent of this reduction, as shown in

figures 5.14 and 5.15. Real animals suffer a corresponding deficiency, as seen from figure 5.12. Latency also decreases with experience as in real animals, and figure 5.7 shows the form of this dependence for the model.

Reversal studies, on the other hand, point up some deficiencies in the model and its parameter settings. Reversal generally occurs too slowly in comparison to real animals, and mnemons initially trained negative never reverse at all, in fact. The reasons for this are discussed in chapter 6. One basic problem is that the parameters are set for insufficient "within session" learning to occur, so that long-term memory exerts too great an influence in comparison to recent experience. The shape of the memory curve for stimulus 1 in figure 6.4 indicates that otherwise the behavior of the model would have been similar to what is actually observed.

Another apparent problem is that neutral stimuli have a somewhat positive effect. The relatively simple answer to this problem is to define "neutral" stimuli as those which cause no net memory change. These will need to be mildly negative in order to cancel the small positive change from the attack buildup itself, but the problem is only one of terminology.

The interference experiments of figures 6.8 through 6.12 predict a deterioration in performance as the time between trials is reduced in a training experiment. This comes about because previously active mnemons are still active during the next trial, and thus false information can get entered into them. The model results show these effects appearing at inter-trial times of about three minutes or less, but the exact time scale depends on model parameter settings and is less important

than the question of whether such effects exist at all or not. Any system which updates memory by gating taste or pain into storage on the basis of whether an element is currently active or has been active in the recent past should necessarily show such interference. It would be impossible to account for the short-term memory effects discussed in section 2 of chapter 1 if this were not the case, since the distinction between delayed reward or punishment and interference is only one of viewpoint.

The model does show one serious flaw in this respect, however. It should not be possible to "backward condition" a mnemon, but this does occur. That is, in any real system only elements which were active before the signals of results arrived are conditioned. Those which become active after the signals of results arrive are not altered. In the model however, memory in newly active mnemons can be changed by taste or pain signals which are still present from events that occurred previously.

The experiments with stimulus overlap (figures 6.5 through 6.7) are perhaps the most potentially interesting, although also the most costly experiments to run. They show that mnemons which are components of more than one stimulus take on intermediate memory values, as might be expected. More complex associations of mnemons into stimuli should yield interesting results, especially since changes to memory only occur if an attack is actually launched, and this provides a bias in the positive direction over the long run. Some results obtained with a simplified form of the model in such pattern recognition situations are discussed in section 3.

2. Changes to the Model

One difficulty in designing a model which will produce results comparable to experimental observations is that different experiments sometimes suggest different directions for changes to the parameter settings. This makes it necessary to use one version of the model equations and one set of parameter values through a complete series of simulated experiments in order to evaluate their adequacy. That is what has been done in the work reported here. A version of the model equations and a set of parameter values was arrived at by a series of initial tests, and these were then used for the complete set of experiments reported in chapters 5 and 6. The results of these experiments were discussed in section 1, and this section will discuss the changes which should be made if they were to be repeated.

The most important change should be to increase the value of recent experience over longer term influences. My recommendation for the best way to accomplish this would be to add A_i within the brackets of equation 4 in table 4.1b, and to add R_i within the brackets of equation 5. The time constant half-lifes for equations 2 and 3 should then be set to about 1 minute for rise and 5 minutes for fall, and the random factor (RC) should be increased slightly. The half-life constants for equations 6 and 7 should then be increased to about half an hour as compensation. The effect of these changes would essentially be to loosen the coupling between AMS_i and A_i, and between RMS_i and R_i, so that the influence of memory on attack buildup could be separated from the influence of attack level on memory change somewhat.

The results to be expected if this were done is that within session learning should show considerable improvement while overall performance

remained about the same. This would correct problems that the model shows with reversal learning, but would not materially affect the other results. It would also bring the model more into line with real animals in another respect. During actual training sessions animals must be given a small piece of fish initially in order to start them attacking. With these changes to the model, a similar procedure would be necessary to bring AMS up to the level of AM before it could have any effect in the equation for A. That is, the model would also have to have an initial taste input before memory would begin to influence attack buildup.

Another change which could be tried in the parameter settings is to introduce time constants for A and R. (Values of 1 for time constants in tables 5.1 and 5.3 essentially mean that those entries do not enter the model as parameters.) This change should have the effect of lowering overall attack level and increasing latency somewhat.

Rather than pursuing speculation further about improvements that could be made to the current model, however, attention will now be turned to other lines of study suggested by this work.

3. Other Model Work

Two directions for further research along these lines are of special interest. One is the study of more general networks of learning elements, and the other involves pattern recognition capabilities in mnemon-like sets of basic elements. Some preliminary investigations have been conducted into both of these areas, and this work will be discussed next.

For the pattern recognition studies, consider an input grid which is connected to a set of memory elements by classifying units. These classifying units perform a feature extraction function by collecting

their input as a weighted sum of grid elements. The grid elements to
which they connect and the values of the corresponding weights determine
the type of feature extraction. One classifying unit might connect
with the grid units in a single horizontal row, for example. This could
be considered analogous to a neuron with approximately horizontal dendritic
branches in the plexiform layer of the optic lobe, and the weighting
factors might represent the relative density of dendrite arborization.
Similarly, classifying units could connect to vertical sets of grid
elements, or in general, to arbitrary sets of elements.

Corresponding to each classifying unit is a mnemon-like memory
unit which contributes to an attack/retreat decision, and which is
altered by the results of an attack in proportion to this contribution.
An arbitrary stimulus pattern might excite numerous memory elements
in varying degrees, and thus there should be interference, even between
patterns which are orthogonal with respect to the grid elements.

These preliminary studies considered a 10 x 10 grid with 40 class-
ifying and memory units. Half the units connected symmetrically about
the vertical centerline and half connected symmetrically about the
horizontal centerline with weights of .16, .13, .10, .06, .03 and the
reverse of these. Weights were the same for all units. Twenty units
were thus devoted to covering each horizontal row and column, and the
other twenty covered the symmetric variations of half-rows and half-columns.
Two stimulus patterns were considered, a vertical bar two columns wide
and a horizontal bar two rows wide. These were presented alternately
for eight trials each, with one as the positive figure and the other as
the negative figure. On each presentation a probability of attack
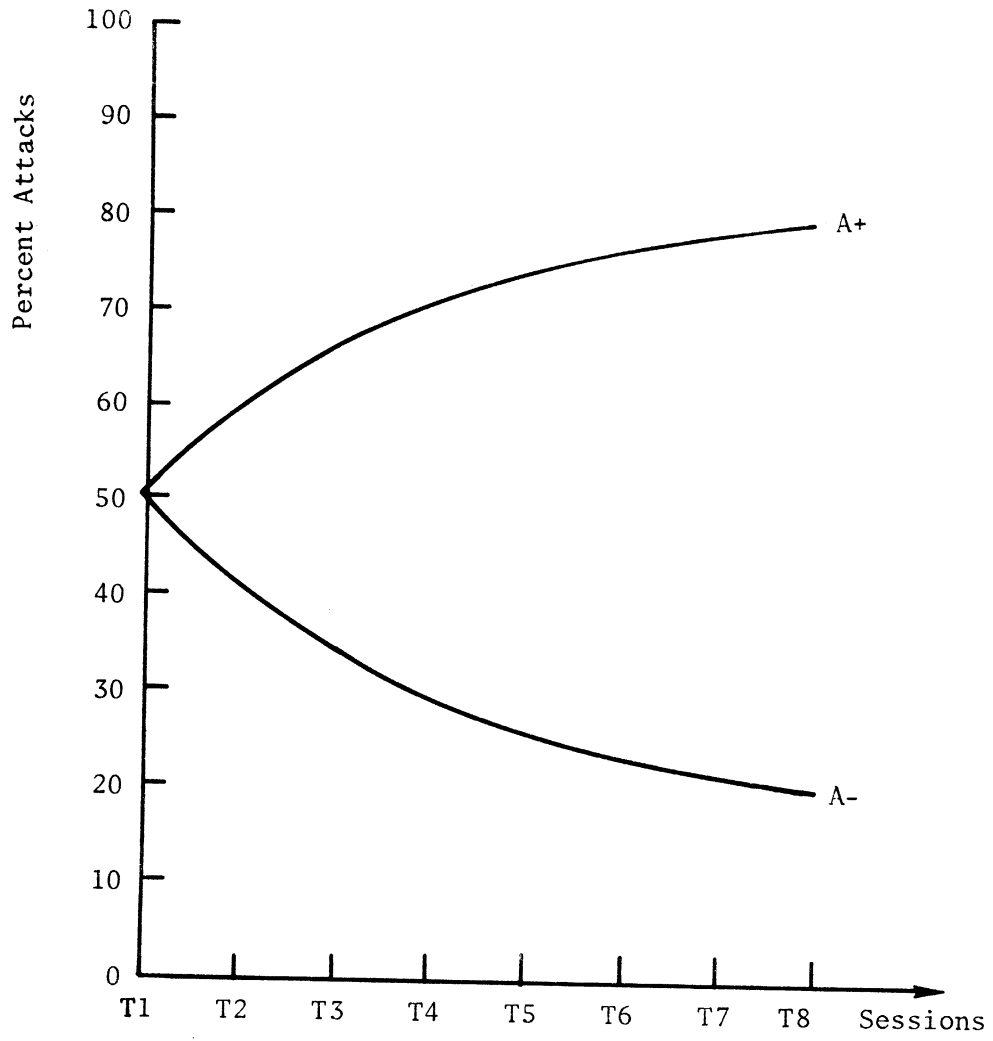was computed for comparison to a random number to determine attack or

Figure 7.1

retreat.[1] Figure 7.1 shows the resulting curve of attack probability. The rate of learning is controlled by a memory smoothing factor, which was set to the intermediate value of .3 in these trials.

Another direction for research which was undertaken in a preliminary study concerns the behavior of arbitrary networks of normalized learning elements which change in accordance with an equation such as 3.1. The approach taken was motivated by the fact that neural material can be conditioned and an association pathway built up. It exhibits a "plasticity" in this sense. In the basic paradigm used here, an input is applied to one element of a randomly connected network on the initial time step, and another input is applied to a different element some n time steps later. After several repetitions of this procedure, the question is whether or not an increase can then be seen in the output of the second element after n time steps when the second external signal is not applied. Under certain conditions this can be seen to occur.[2] Table 7.1 shows some sample values of this second node before and after ten conditioning trials at two values of the delay, n.

4. <u>Summary</u>

The model performs reasonably well in the standard experimental situation. It reacts to simulated operations just about as real animals do, thus demonstrating that performance failure can be due to interference with the memory read/write mechanism even though memory itself remains intact. Reversal studies point up difficulties with the model, however. The time interference and stimulus overlap experiments predict deterioration in performance with short intertrial times and non-orthogonal stimuli.

Model performance could be improved, especially in the reversal case,

	n = 2				n = 3	
T	Before	After		T	Before	After
1	8	2		1	3	3
2	4	9		2	6	32
→ 3	3	22		3	4	3
4	2	9		→ 4	3	27
5	0	14		5	2	2
6	0	6		6	0	16
7	0	8		7	0	1
8	0	3		8	0	9
9	0	4		9	0	0

Table 7.1

by the changes outlined in section 2. These would loosen the coupling between short-term memory and attack or retreat outputs, and should thus increase the influence of recent experiences over longer memory.

Two directions for further study of networks with learning elements are described in section 3, and illustrative examples are given.

CHAPTER EIGHT

SUMMARY AND CONCLUSIONS

1. Review

Biological systems exhibit an astonishing degree of complexity. Nervous systems and brains, in particular, confront researchers with an extremely difficult subject for study. Their basic element, the neuron, is both fragile and complicated in its operation, so that even gross functional relationships are often difficult to determine. These nervous systems permit organisms a wider range of adaptive capabilities, and thus one approach to their study is through hypotheses about what they may be doing, based on the needs of the animal in its environment. The octopus, like most other animals, must concern itself principally with discriminating objects which are likely to yield taste and food value from those likely to lead to pain and harmful effects. One primary function of the higher levels of its nervous system, therefore, must be to learn to make this distinction.

After considerable study of the octopus and the anatomy of its nervous system, Young has proposed an idea for how this discrimination might be made in the visual system. He suggests that classifying cells with dendritic fields in the plexiform layer of the optic lobe are inputs to memory circuits in these lobes, which he terms "mnemons". These mnemon circuits can be switched through the action of taste or pain signals into either an "attack" pathway or a "retreat" pathway, where the outputs are then elaborated into patterns of action at lower motor levels. Each active mnemon is assumed to retain a memory of the direction in which it was switched, so that later recurrences of the same stimuli

will yield surer responses. He further postulates that the action of
circuits located in the vertical lobe structures of the brain is necessary
for this "writing into" and "reading from" memory to occur. Similar
functions are inferred for inferior lobe structures and the tactile
learning system.

Chapters 3 and 4 describe a series of models I developed on the
basis of this mnemon concept, and chapters 5 and 6 give a summary of the
results obtained from a set of experiments conducted with the model
described in chapter 4. The model itself is a compromise between a
desire to incorporate as many details of the postulated circuitry as
possible while still being able to conduct large scale experiments
within reasonable run times on the computer. It thus includes three
layers of memory in the mnemon, for example, but reduces the entire
action of the upper lobe structures to a single feedback variable, Q.
Even this minimal version often required long run times on the 360-67
to produce the curves of chapters 5 and 6.

Figure 4.2 shows my conceptualization of a mnemon and table 4.1
gives the basic set of model update equations. The general form of the
equations is spelled out in 4.1 through 4.5, and figure 4.3 outlines
program operation. Details concerning the behavior of the model and some
insight into its operational characteristics are provided in section 1
of chapter 5. The rest of chapter 5 and chapter 6 are devoted to a
description of various experiments. Chapter 7 discusses the results of
these experiments and suggests changes which might be made in the model
to improve its performance. Some possible directions for further
theoretical work are then given.

2. Conclusions

A simple model of learning and memory based on Young's mnemon concept can produce behavior patterns comparable to experimental behavior in many learning situations. The model shows both short-term effects which fade over time, and also more lasting long-term changes to behavior. The model can be conditioned to discriminate between two stimuli and its performance will approach, but not attain, 100% in this discrimination. Its performance within each training session improves and reaches a higher level than can be measured at the beginning of the next session sometime later.

Operations which interfere with the upper lobe structures can be simulated in the model with the result that performance deteriorates in proportion to the extent of this interference, even though memory levels remain unaltered. This degradation in performance results from interference with the memory "read/write" mechanism rather than from the loss of memory itself. When interference exists but is less than complete, then performance shows the proper direction of learning but at a level which is not statistically significant.

Training can be reversed and the model will learn the new direction of training, but it will still show a strong bias in the direction of its original training. If training is repeatedly reversed, then the model fails to respond properly because its parameters are not set to allow sufficient influence from recent experience. With this deficiency corrected the model would show an ability to reverse repeatedly, but with an eventual approach toward a random level of performance.

The model shows considerable interference when trials are placed close together. Performance falls because attacks on the positive

figure are reduced. The model thus predicts a fall in overall attack

level as the time between trials is reduced in a training experiment.

It also predicts an improvement in overall performance at short intertrial

times from other training disciplines than alternating trials, but says

that these effects will decrease as the time between trials becomes

longer. An experiment could be conducted with real animals in which

the time between trials is shortened to determine if this has an effect

upon performance. If so, the model predicts improvement by changing

the training discipline to some form of non-alternating trials.

The model says nothing about the very interesting question of

innate stimulus perferences and how these are affected by operations

on the upper lobe structures. All memory values are initially set to

zero in the model and each stimulus component is given equal weight

in the attack/retreat calculations.

Latency of response falls off rapidly with training in the model,

and also varies with hunger level. The effect of taste as a short-term

reward is separated from its long-term influence in reducing hunger.

Following simulated operations which reduce memory influences, recent

taste or pain input has a greater effect in determining model response

than in the un-operated condition.

In experiments with overlapping stimuli the model shows a form of

stimulus generalization. Net memory change is dependent on the degree of

association which each mnemon has had with positive and negative stimuli,

so that performance with similar stimuli is similar.

Transfer across the mid-line was not discussed, but would be accounted

for by the model in the same way that it was accounted for by the

Maldonado model discussed in chapter 3.

The model itself shows many shortcomings, but as the simplest

version I could conceive which would follow the basic anatomical constraints

and utilize realistic half-life values for the memory constants, it

displays a surprising number of characteristics observed in real animals. The work has been motivated by a belief that there is fertile ground for cooperation between computer scientists and biologists. Biological organisms are enormously complex, but models and computer simulations are one more tool to aid in their study, and we should see models of increasingly powerful descriptive capability emerging as our knowledge of local properties in biological systems continues to increase.

FOOTNOTES*

Chapter 1

1. Horridge (1968), p. 377

2. Young (1964), p. 25

3. Young (1964), pp. 91-92

4. Young (1964), p. 69

5. Horridge (1968), p. 362

6. Wooldridge (1963), p. 185

7. Young (1964), p. 82

8. Young (1964), p. 177

Chapter 2

1. Young (1966), p. 24

2. Young (1964), p. 202

3. Young (1964), pp. 205-206

4. Young (1964), p. 207

5. Young (1964), pp. 212-213

6. Young (1964), pp. 215-217

7. Additional references to this work will be found in the bibliography

Chapter 3

1. Maldonado (1963)

Chapter 5

1. All octopus data quoted in this chapter are used with the kind permission of Professor J.Z. Young and are taken from his experiments at the *Stazione Zoologica* in Naples during July 1971 (A. 187, H21-H25).

*References are to items in the bibliography.

FOOTNOTES

(Cont'd)

2. Gray (1970)

Chapter 7

1. Attack was actually forced in these trials by always setting the random number to zero.

2. When other cycles do not develop in the network first.

REFERENCES

Arkadev, A.G. and Braverman, E.M. *Computers and Pattern Recognition*, Thompson, 1967.

Boycott, B.B. and Young, J.Z. "A memory system in *Octopus Vulgaris* Lamarck," *Proc. Royal Society B*, 143:449-480, 1955.

Gordon, Geoffrey, *System Simulation*, Prentice-Hall, 1969.

Gray, E.G. "The fine structure of the vertical lobe of the Octopus brain," *Phil. Trans. of the Royal Society of London B*, *258*, 379-395, 1970.

Horridge, G.A. *Interneurons*, W.H. Freeman, 1968.

Hubel, D.H. and Wiesel, T.N. "Receptive fields, Binocular interaction, and functional architecture in the cat's visual cortex," *Journal of Physiology*, 160:106-123, 1962.

Lettvin, J.T.; Maturana, H.R.; McCulloch, W.S. and Pitts, W.H. "What the frog's eye tells the frog's brain," *Proceedings IRE*, 47:1940-1951, 1951.

Maldonado, H. "The visual learning system in *Octopus Vulgaris*," *J. Theoretical Biology*, 5:470-488, 1963b.

Minsky, Marvin, "Steps toward artificial intelligence," *Proceedings IRE*, 49:8-30, 1961.

Mize, J.H. and Cox, J.G. *Essentials of Simulation*, Prentice-Hall, 1968.

Samuel, A.L. "Some studies in machine learning using the game of checkers," *IBM Journal of Research and Development*, 3:211-229, 1959.

Selfridge, O.G. "Pandemonium: A paradigm of learning," *Mechanization of Thought Processes*, Her Majesty's Stationery Office, 513-526, 1959.

Sutherland, N.S. "The visual system of *Octopus* (3) Theories of shape discrimination in Octopus," *Nature*, 186:840-844, London 1960.

Wells, M.J. "A touch-learning center in *Octopus*," *J. Exp. Biology*, 36:590-612, 1959a.

___, "Proprioception and visual discrimination of orientation in *Octopus*," *J. Exp. Biology*, 37:489-499, 1960a.

___, "Centres for tactile and visual learning in the brain of *Octopus*," *J. Exp. Biology*, 38:811-826, 1961b.

Wells, M.J. and Wells, J. "Repeated presentation experiments and the function of the vertical lobe in *Octopus*," *J. Exp. Biology*, 34:469-477, 1957a.

___, "The function of the brain of *Octopus* in tactile discrimination," *J. Exp. Biology*, 34:131-142, 1957b.

Wells, M.J. and Young, J.Z. "Learning at different rates of training in the *Octopus*," *Animal Behavior*, 17:406-415, 1969.

___, "The effect of splitting part of the brain or removal of the median inferior frontal lobe on touch learning in *Octopus*," *J. Exp. Biol.* 50:515-526, 1969.

___, "Split-brain preparations and touch learning in the *Octopus*," *J. Exp. Biol.*, 43:565-579, 1965.

Wooldridge, D.E. *The Machinery of the Brain*, McGraw Hill, 1963.

Young, J.Z. "The failures of discrimination learning following removal of the vertical lobes in *Octopus*," *Proc. Royal Society B*, 153:18-46, 1960c.

___, "Learning and discrimination in Octopus," *Biological Revue*, 36:32-96, 1961.

___, "Reversal of learning in Octopus and the effect of removal of the vertical lobe," *Quarterly J. Exp. Psychology*, 14:193-205, 1962f.

___, "Memory mechanisms of the brain," *J. of Mental Science*, 453:120-132, 1962.

___, "Some essentials of neural memory systems; paired centres that regulate and address the signals of the results of action," *Nature*, 198:626-630, 1963.

___, *A Model of the Brain*, Oxford Press, 1964.

___, "The organisation of a memory system," *Proc. Royal Society B*, 163:285-320, 1965.

___, *The Memory System of the Brain*, California Press, 1966.

___, "Influence of previous preferences on the memory of *Octopus vulgaris* after removal of the vertical lobe," *J. Exp. Biol.*, 43:595-603, 1965.

___, "Short and long memories in *Octopus* and the influence of the vertical lobe system," *J. Exp. Biol.*, 52:385-393, 1970.