

THE UNIVERSITY OF MICHIGAN
COLLEGE OF LITERATURE, SCIENCE, AND THE ARTS
Department of Communication Sciences

Interim Engineering Report

RESEARCH IN THEORY OF ADAPTIVE SYSTEMS

M. R. Finley, Jr.
S. T. Hedetniemi
C. V. Page

ORA Project 06114

under contract with:

AF AVIONICS LABORATORY
SYSTEMS ENGINEERING GROUP
AIR FORCE SYSTEMS COMMAND
CONTRACT NO. AF 33(615)-1162
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

administered through:

OFFICE OF RESEARCH ADMINISTRATION ANN ARBOR

January 1965

Engn

UMR

1413

FORWORD

This report was prepared by The University of Michigan under USAF Contract No. AF33(615)-1162. The contract was initiated under Project No. 4160, "Engineering Bionics," Task No. 416004, "Bionic Sub-System Techniques." The work was administered under the direction of the Bionics Branch, Electronic Technology Division, Air Force Avionics Laboratory, Dr. Donald E. Lewis, program monitor.

This program of research was carried out at The University of Michigan by the Logic of Computers Group, Professor Arthur W. Burks, director. The principal investigator for this research program was John H. Holland, Associate Professor, Communication Sciences Department. Assisting Dr. Holland in this program were Marion Finley, Jr., Stephen Hedetniemi, and Carl Page, Research Assistants in Communication Sciences, and Dr. Harvey Garner, Professor of Electrical Engineering at The University of Michigan.

ABSTRACT

The research on adaptive systems described in this report may be subdivided into two areas according to the approach used: (1) automata theory and Turing machines, and (2) neural network simulation.

Automata Theory and Turing Machines

S. T. Hedetniemi, "Studies in Cellular Automata." John von Neumann's development of a space of cellular automata, which he used to demonstrate the construction and reproduction of automata did not allow certain primitives for the 29-state automaton used as the cells in his spaces. In particular, his primitives would not allow the simultaneous crossing of two channels of information. C. Y. Lee gave a solution to this cross-over problem which was more efficient than that of von Neumann. In this paper, the author gives a further improvement which allows simultaneous cross-over, but indicates that the construction of the latter would be difficult.

C. V. Page, "Formulating a Game-Theoretic Problem in Probabilistic Sequential Machine Theory." The author gives an example of a game which is relevant to the theory of adaptive systems and rephrases it as a problem in probabilistic sequential machines; then discusses whether or not a nesting property may be obtained for these machines which would yield information about the expected payoffs after some substring of plays.

Neural-Network Simulation

M. R. Finley, Jr., "Experimental Study of Neural Networks by Means of a Digital Computer Simulation." The author gives the development of a class of abstract neural network models, based on certain neurophysiological evidences and the extrapolations of D. O. Hebb in his development of a theory of learning. He discusses Hebb's notion of cell-assembly, then describes a series of initial experiments to test the model for basic desired properties, and the results obtained. From these, considerable information was gained concerning the nature of the basic network functions, such as threshold, fatigue, etc. A derivation is given for the form of the threshold curve, and empirically derived arguments are given for the fatigue and synapse-value curves.

Publication of this technical documentary report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

TABLE OF CONTENTS

	Page
STUDIES IN CELLULAR AUTOMATA	S. T. Hedetniemi 1
FORMULATING A GAME-THEORETICAL PROBLEM IN PROBABILISTIC SEQUENTIAL MACHINE THEORY	C. V. Page 7
EXPERIMENTAL STUDY OF NEURAL NETWORKS BY MEANS OF A DIGITAL COMPUTER SIMULATION	M. R. Finley, Jr. 17

STUDIES IN CELLULAR AUTOMATA

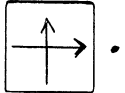
S. T. Hedetniemi

Manuscript released by authors January 1965 for publication as an RTD Technical Report.

The following is an indication of some of the work being done in an attempt to understand and achieve optimal designs within von Neumann's framework of 29-state cellular automata.

Perhaps one of the most perplexing problems von Neumann had to solve in constructing a self-reproducing cellular automaton was that of transmitting information from one point in his cellular two-space to another without its being in any way distorted. Although it is easy to construct a path (of cellular automata) capable of transmitting information from one origin to one destination point, it is not nearly so easy to construct paths from several origins to several destinations, for invariably in two-space at least two paths must cross.

Unfortunately or not, von Neumann's basic 29-state automata were not capable of accepting inputs from either of two directions (say from the east and south) and outputting the same in the corresponding opposite directions (i.e., to the west and north, respectively). Von Neumann did not design a crossover primitive, one which might be symbolized, as did C. Y. Lee* as



The problem was effectively solved first by encoding differently the information from the various origins, and then sending all the information along a common channel (path) which passes by each of the destinations, at which are located corresponding decoding devices which recognize (ideally) only the encoded information from their respective origins.

Using this procedure, von Neumann was able to construct a network for transmitting information which had the desired crossover properties, but in

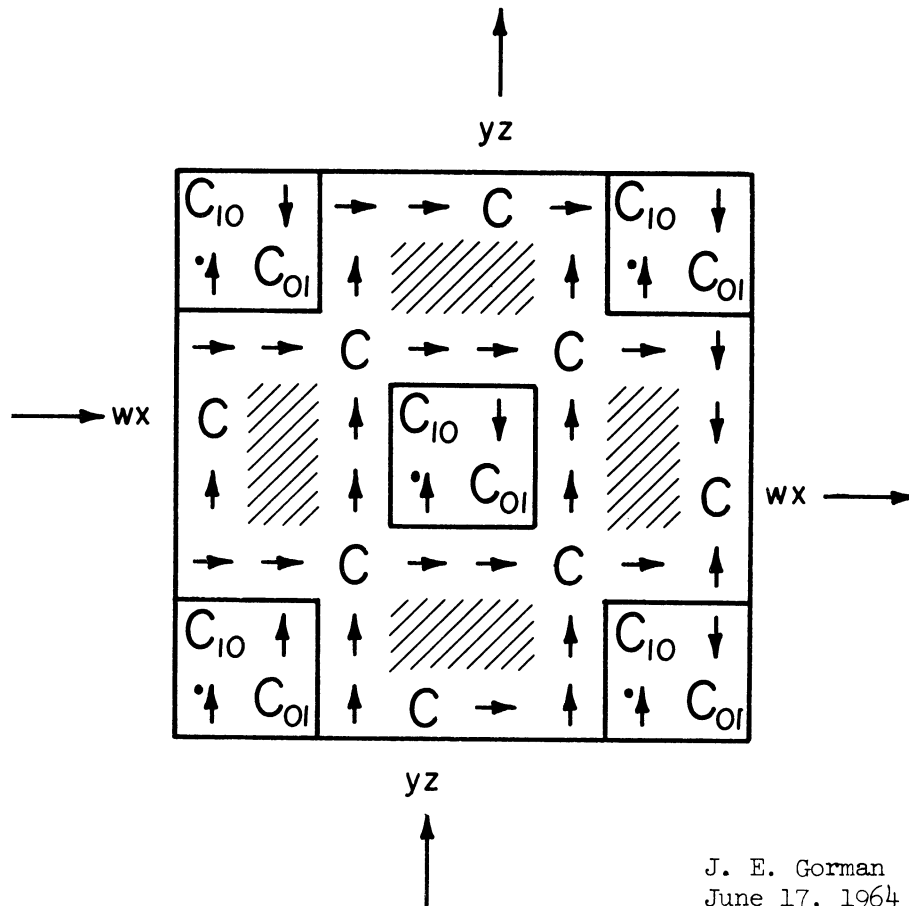
*Lee, C. Y., "Synthesis of a cellular computer using the 29-state model of von Neumann," Engineering Summer Conferences, Automata Theory, The University of Michigan, Summer 1964.

work is time dependent.

A particularly ingenious 8x8 crossover network (see below) has been designed which allows for simultaneous information crossover, however, it has one major drawback. Since this network is always in an active state, and since construction in the von Neumann model is designed primarily for passive networks, it appears as though the construction of this network will be particularly difficult.

On the other hand, since all of the previously mentioned crossover networks are passive, their construction presents no problems.

THIS DEVICE ALLOWS SIGNALS IN WIRES TO CROSS SIMULTANEOUSLY



J. E. Gorman
June 17, 1964

8x8 Synchronous Crossover
von Neumann Cellular Automata

FORMULATING A GAME-THEORETIC PROBLEM
IN PROBABILISTIC SEQUENTIAL MACHINE THEORY

C. V. Page

The Game-Theoretic Problem

Consider the following game which is important as an example in the theory of adaptive systems. Two players alternately choose either a 0 or a 1 for a total of $2N$ choices. After the sequence of choices is made, a payoff is supplied depending only on the binary number of $2N$ bits defined by the sequence. If the payoff is positive, it goes to one player, otherwise to the other player. Of interest to us is the case when one player (designated by FM) uses a fixed mixed strategy dependent on the previous plays of the game, i.e., after any string $x = i_1 \dots i_k i_j \in \{0, 1\}$ player FM chooses a 1 with probability p_x^1 and 0 with probability $p_x^0 = 1 - p_x^1$. Does there exist for the other player (designated by D) a fixed sequence of plays which

- (1) maximizes the expected payoff for player D against play FM?
- (2) has the property that after some substring of plays, each succeeding play increases the expected payoff for D against the fixed mixed strategy of FM?

Condition (1) seems quite likely to occur for arbitrary probabilities of FM. On the other hand, condition (2) which we will call the nesting property depends on the relationship between the mixed strategy probabilities and the game values attached to strings of length $2N$. Probabilistic sequential machines provide a framework in which necessary and/or sufficient conditions for the nesting property may be obtained.

An Equivalent Probabilistic Sequential Machine

Following a definition of probabilistic sequential machines and related terms, the above game-theoretic problem is expressed as a problem in probabilistic sequential machines. The insight gained from the study of such machines should prove fruitful in the study of this problem.

Definition 1.1. A probabilistic sequential machine A is a system of inputs, internal states and outputs specified by

$$A = \langle n, I, S, \Sigma, A(0), \dots, A(K-1), F, \mathcal{O} \rangle$$

where

n : a finite natural number, the number of states.

I : an n -dimensional stochastic row vector, the initial state vector.

S : the set of state vectors

$$S = \{S_1 = (1, 0, \dots, 0), \dots, S_n = (0, \dots, 0, 1)\}.$$

Σ : input alphabet set $\Sigma = \{0, 1, \dots, K-1\}$.

$A(i)$: $i = 0, 1, \dots, K-1$: $n \times n$ switching matrix for input symbol i .

$A(i)_{pm}$ is the probability of a transition from state p to state m when symbol i occurs.

F : output vector, a n -dimensional column vector whose entries are real numbers.

\mathcal{O} : output function $\mathcal{O}(S_i) = S_i \cdot F = F_i$ S_i in S where \cdot is just matrix multiplication. (In instances where no confusion occurs the symbol "." is left out.)

The correspondence between the game and a probabilistic sequential machine.

<u>Game Interpretation</u>	<u>Probabilistic Sequential Machine</u>
(i) A player specifies 0 or 1.	(i) Input of 0 or 1 to the machine.
(ii) A player will specify a choice using a mixed strategy dependent on previous inputs.	(ii) An input of ? to the machine.
(iii) Number of moves for each player N .	(iii) Number of states $n = 2^{N+1} - 1$.

<u>Game Interpretation</u>	<u>Probabilistic Sequential Machine</u>
(iv) Play of game of length $2K$ with mixed strategy player moving second.	(iv) Input string x of the form $i_1 ? i_2 ? \dots i_k ?$ where i_j in $\{0, 1\}$ $j = 1, \dots, K$.
(v) Partial sequence of plays observed i.e., 01011 .	(v) State of machine i.e., S_{01011} .
(vi) A priori changes in game tree caused by specification of $0, 1$ or $?$.	(vi) Switching matrices $A(0), A(1)$ and $A(?)$ which tabulate change of state of the game tree with input.
(vii) Start of the game.	(vii) Initial state of the machine $(1, 0, \dots, 0)$.
(viii) Value of payoff for game with plays 01011101 .	(viii) Output of machine which is $F_{01011101}$ from state $S_{01011101}$.

Remarks

Let $x = i_1 i_2 \dots i_r$ $i_j \in \{0, 1, ?\}$ $j = 1, \dots, r$ be an arbitrary string. Then the switching matrix for x can be found from the switching matrices of its symbols by matrix multiplication, i.e.,

$$A(x) = A(i_1) A(i_2) \dots A(i_r) .$$

The expected value of output from A for a string x is a bilinear form in I and F with form matrix $A(x)$ i.e., $E_A(x) = I A(x) F$.

In order to reflect the game problem we define the average expectation of a game sequence $x = i_1 j_1 \dots i_n j_n$ where the i_k 's are given by player D and the j_k 's by player FM using "?" to symbolize the move of player FM .

$$\frac{x}{E(x)} = \frac{1}{2^n} \sum_y E_A (i_1 ? i_2 ? \dots i_n ?)$$

where $y = (i_1 \dots i_n) \in \{0, 1\}^n$

For a game beginning with the fixed sequence by player D of $z = i_1 \dots i_k$ $k \leq n$ (and then random choices for the rest) we have

$$\frac{y}{E(zy)} = \frac{1}{2^{n-k}} \sum_y E_A (i_1 ? i_2 ? \dots i_k ? \dots i_n ?)$$

where $y = (i_{k+2} \dots, i_n) \in \{0, 1\}^{n-k}$

Example: A machine with the nesting property.

We show a 31 state machine which has the nesting property for games of length 4 in which the mixed strategy player FM moves second. There seems to be no theoretical difference in whether D or FM moves first. The general form of A(?) and F are shown in Figure 1. A(0) and A(1) are not shown, but one can consider them as special cases of A(?). A(1) is just A(?) with all p_s^1 equal 1 while A(0) is just A(?) with p_s^0 equal 1. Hence the rows of A(?) are convex combinations of the rows of A(0) and A(1). Figure 2 presents the special case which illustrates the nesting property with machine A_N .

Note that all entries of F which are not terminal states of the game are zero. To complete the representation it is assumed that any move after a game is concluded restores the game to its initial state.

For machine A_N and game strings of length 4, if player D plays at random he can obtain the average game value

$$\frac{x}{E(x)} = 4.05$$

but if D begins with 0 and chooses the second move at random

$$\frac{y_1}{E(0?y_1?)} = 2.675 \quad y_1 \text{ in } \Sigma = \{0, 1\}$$

while starting with a 1 gives

$$\frac{y_1}{E(1?y_1?)} = 5.425 \quad y_1 \text{ in } \Sigma$$

choosing 1 for both moves, D obtains

$$\frac{\quad}{E(1?1?)} = 8.100$$

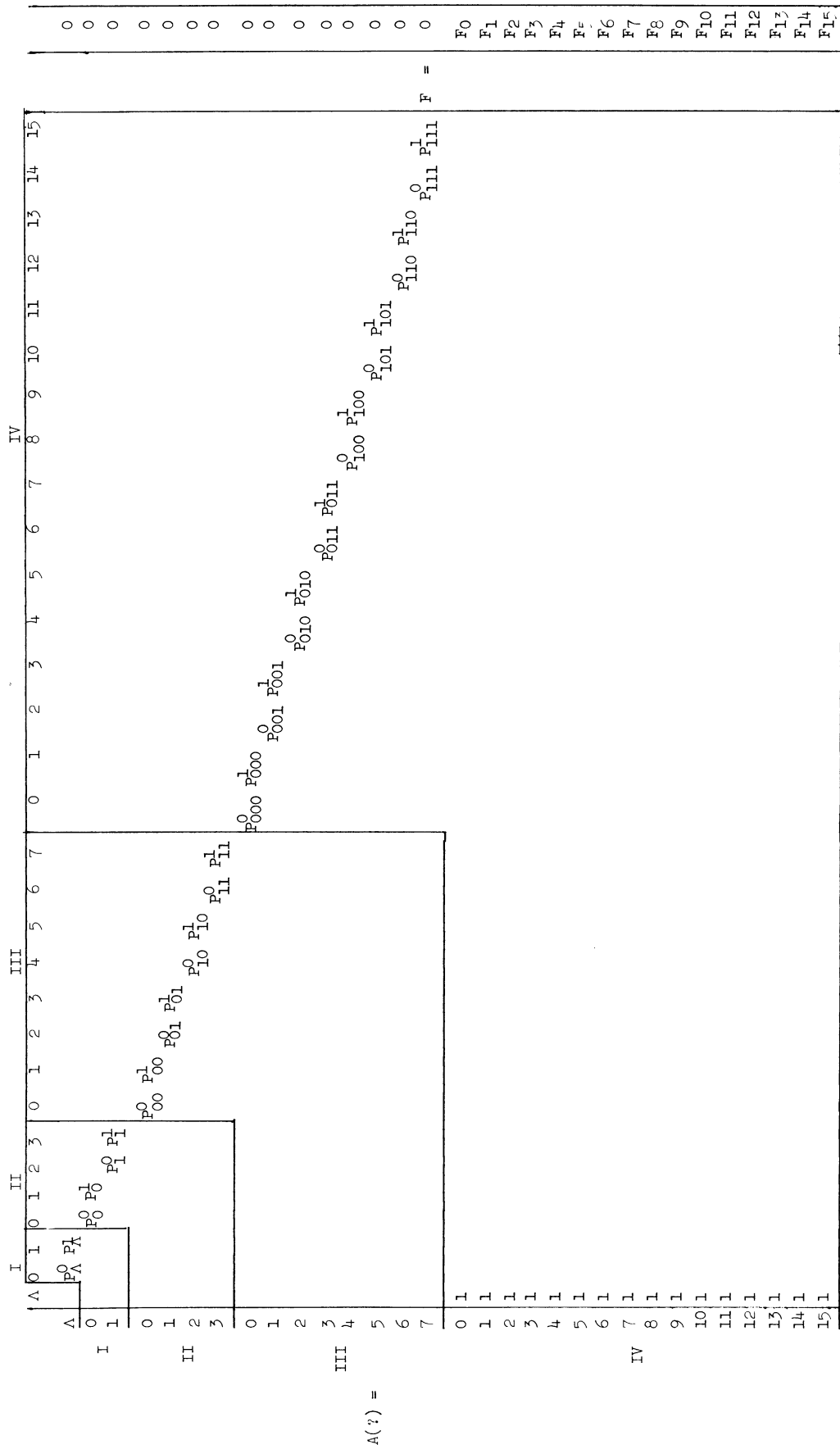


Figure 1. Mixed matrix $A(?)$ and output vector F . Final states are mapped into the initial state. For $A(0)$ set all $P^0_S = 1$ and $P^1_S = 0$. For $A(1)$ set $P^1_S = 1$, $P^0_S = 0$. Note that 0 in phase II is 00 while 0 in phase IV is 0000. All entries left blank are zero.

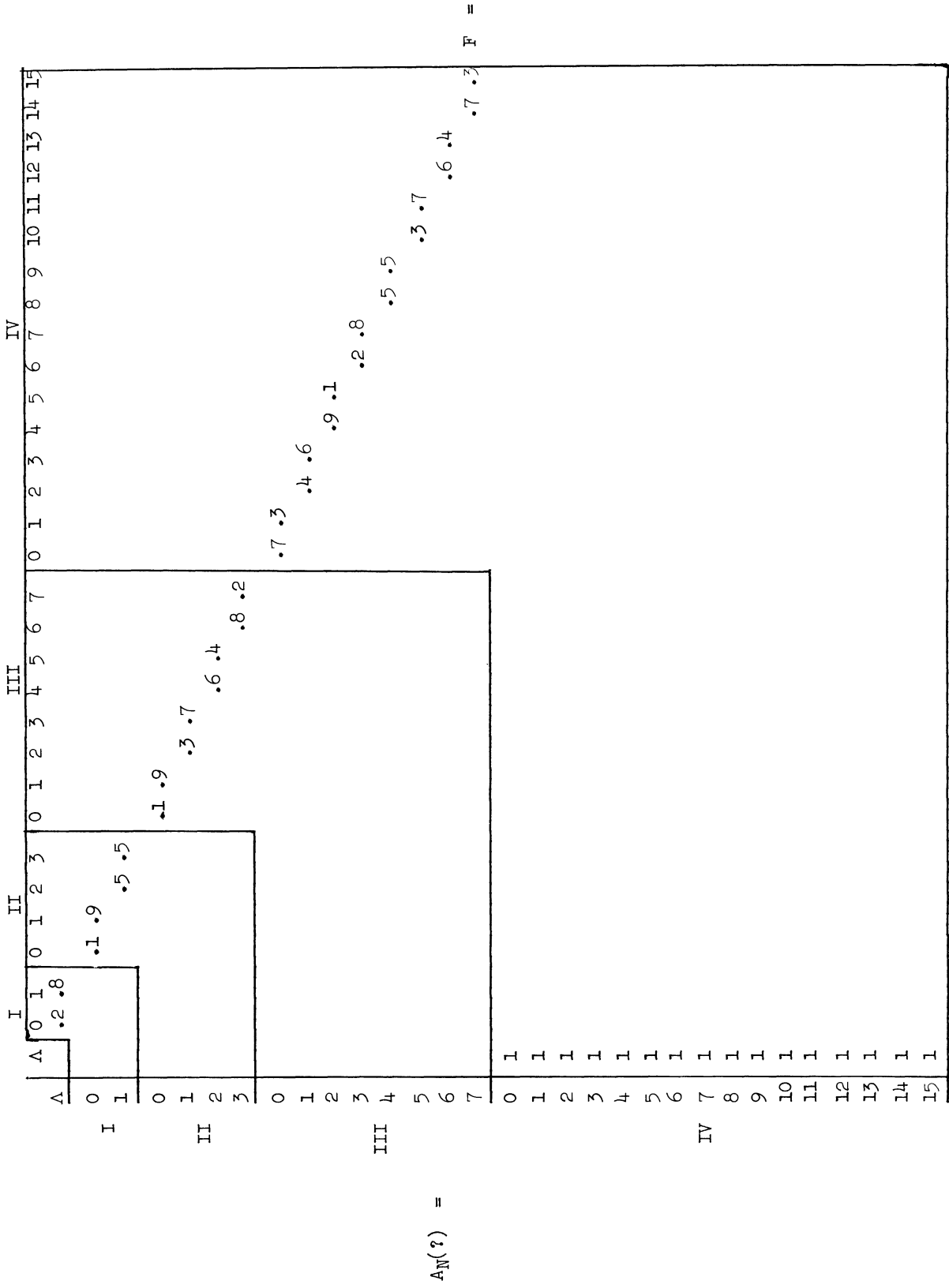


Figure 2. $A_N(?)$ and F of machine A_N which has the nesting property.

Hence the string 11 for D has the nesting property in A_N

$$E(1?1?) > \frac{y_1}{E(1?y_1?)} > \frac{y_1 y_2}{E(y_1?y_2?)} \quad y_1, y_2 \text{ in } \Sigma .$$

General Results

The special form of the matrices $A(?)$, $A(0)$, and $A(1)$ make it possible to write simple general expressions for $\frac{x}{E(x)}$ and $\frac{y}{E(zy)}$.

$$\frac{x}{E(x)} = \frac{1}{2^n} \sum_x P_{i_1}^{j_1} P_{i_1 j_1 i_2}^{j_2} \cdots P_{i_1 j_1 i_2 j_2}^{j_n} \cdots i_n F_{i_1 j_1} \cdots i_n j_n$$

where $x = (i_1, j_1, i_2, j_2, \dots, i_n, j_n) \in (\Sigma)^{2n}$

If $z = i_1 ? \dots i_k ?$

$$\frac{y}{E(zy)} = \frac{1}{2^{n-k}} \sum_y P_{i_1}^{j_1} P_{i_1 j_1 i_2}^{j_2} \cdots P_{i_1 j_1}^{j_n} \cdots i_n F_{i_1 j_1} \cdots i_n j_n$$

where $y = (j_1, j_2, \dots, j_k, i_{k+1}, j_{k+1}, \dots, i_n, j_n) \in (\Sigma)^{2n-k}$.

A start z provides a better return for D than random iff:

$$\begin{aligned} & \sum_y P_{i_1}^{j_1} P_{i_1 j_1 i_2}^{j_2} \cdots P_{i_1 j_1 i_2 j_2}^{j_n} \cdots i_n F_{i_1 j_1} \cdots i_n j_n \\ & y = (j_1, \dots, j_n, i_{k+1}, \dots, i_n) \in (\Sigma)^{2n-k} \\ & > \frac{1}{2^k} \sum_x P_{i_1}^{j_1} P_{i_2 j_1 i_2}^{j_2} \cdots P_{i_1 j_1}^{j_n} \cdots i_n F_{i_1 j_1} \\ & x = (i_1, j_1, \dots, i_n, j_n) \in (\Sigma)^{2n} \quad (*) \end{aligned}$$

Further Research

The major goal of this research is not yet realized. However, it is clear that the problem of finding those games of the form described in the first section which have nested sequences which improve the expected output has been reformulated in terms of equation (*). Hence, further investigation will use the methods of convex sets to study inequalities among the

output weights and game values which guarantee the nesting property to occur.

Calculation of $\frac{x}{E(x)}$ and $\frac{x}{E(zy)}$ can be simplified and done much more efficiently than by matrix manipulations. Research is underway to develop a simple algorithm which calculates these quantities by tracing the expectation from the final states back up the tree to the initial state.

Another problem of interest would be to characterize all those machines of the same size which have the same nested sequence.

EXPERIMENTAL STUDY OF NEURAL NETWORKS
BY MEANS OF A DIGITAL COMPUTER SIMULATION

M. R. Finley, Jr.

TABLE OF CONTENTS

	Page
1. INTRODUCTION	21
1.1 Statement of the Problem	21
1.2 Basic Premises and Theory: Relation to Neurophysiological Fact	22
1.3 The Cortical Neuron and Systems of Cortical Neurons	25
1.3.1 Structure	25
1.3.2 Input and Output, Threshold	26
1.3.3 Synapses	28
1.3.4 Fatigue, Spontaneous Firing	29
1.4 Previous Neural Net Studies	31
1.5 Plan of Research	33
2. FORMAL DESCRIPTION OF THE MODELS	35
2.1 Introduction	35
2.2 The Network Equations	35
2.3 The Network Functions R, F, and S	42
2.3.1 Control of Firing Rate	42
2.3.2 The Threshold Function	43
2.3.3 The Fatigue Function	45
2.3.4 The Synapse Value Function	47
2.4 Note on the Simulation Program	50
3. CORRELATION EXPERIMENTS, CYCLE-LESS CASE	54
3.1 Introduction	54
3.2 Correlation	55
3.3 Network Configurations for the First Stage	57
3.3.1 Overview	57
3.3.2 Specification of the Networks for the First Stage	58
3.3.3 Network Functions, Initial Conditions, Environment	60
3.4 Three-Neuron Experiments	60
3.4.1 Experiment 1	61
3.4.2 Analysis and Comment on Experiment 1	62
3.4.3 Experiment 2	64
3.4.4 Analysis and Comments on Experiment 2	65
3.5 Nine-Neuron Experiments	85
3.5.1 Experiment 3, Synchronous Case	86
3.5.2 Analysis and Comment on Experiment 3	87
3.5.3 Experiment 4, Asynchronous Case	88
3.5.4 Analysis and Comment on Experiment 4	90
3.6 Thirty-Three Neuron Experiments	91
3.6.1 General	91
3.6.2 Some Theoretical Considerations	93
3.6.3 Experiment 5, Fatigue Curve Tests	95
3.6.4 Analysis and Comments on Experiment 5	96
3.6.5 Experiment 6, Initial Trend Studies	99
3.6.6 Analysis and Comments on Experiment 6	100

TABLE OF CONTENTS (Concluded)

	Page
3.6.7 Experiment 7. Further Tests on the Fatigue Function	101
3.6.8 Analysis and Comments on Experiment 7	106
3.7 Comments on the Network Functions V , ϕ , and S	106
3.7.1 The Threshold Function, $V(r)$	106
3.7.2 The Fatigue Function, $\phi(l)$	107
3.7.3 The Synapse-Value Function, $S(\lambda)$	109
4. CONCLUSION	111
APPENDIX: DETAILED HISTORY OF SYNAPSE-VALUE CHANGES FOR RUN 4, EXPERIMENT 5	113
REFERENCES	117

1. INTRODUCTION

1.1 STATEMENT OF THE PROBLEM

A class of models of neural nets is given which purports to represent, admittedly in an approximate fashion, a fragment of the (association layer) mammalian cortex. Such a model usually will be visualized in an environment together with appropriate sensory and motor apparatus, thus allowing, for example, detection of objects and movement in the environment. The main problem is to determine whether the models presented have the capacity to learn, in the sense that, as a consequence of feedback from the environment to the model, certain internal changes occur in the model with a resulting (eventual) improvement in behavior.

This class of neural net models has at least one distinctive feature: that is, it is interpreted directly into a computer program. Thus, one has a rigorous expression of (the particular interpretation of) the class of models, from which any specific model is obtained merely by specification of certain parameters. Inasmuch as any program is a formal expression of certain formal operations (analogous to the specification of a list of functions used in the definition of partial recursive functions), it possesses some of the advantages found in the study of formal systems. On the other hand, there also is the advantage that any property of the class of models which is deduced a priori can be, in the interpretation afforded by the computer program, subjected to a well-defined test.

Because of the ease with which operations of the models are interpreted into digital computer operations (more realistically, subroutines), the computer simulation of such models is lifted out of the realm of a mere programming application. That is, in a sense, the program itself is a model.

Study it—i.e., its behavior—and you are studying the model.

1.2 BASIC PREMISES AND THEORY: RELATION TO NEUROPHYSIOLOGICAL FACT

The original source for the specification of this class of neural net models and of the neural as well as behavioral processes involved in learning stems back to the theory which was developed by D. O. Hebb [3] and later modified somewhat by P. M. Milner [4]. The theory, which integrates knowledge of neural events, taking place in time intervals of up to a hundred milliseconds or so, with behavioral events, taking place in time intervals of seconds on up, has as its basis the proposed mechanism of the cell-assembly: informally characterized, this is a system of cortical (association layer) neurons which are capable of acting as a closed autonomous functional unit for a brief period of time. These neurons are anatomically diffuse, but functionally connected. The functional unity of the cell-assembly results from the initial existence of the proper inter-connections among the neurons of the system together with a particular (i.e., selective) sequence of cortical events that forces these neurons to act briefly as a unit via a growth of synaptic strength at the connections such that after a period of time the assembly may be activated by appropriate excitatory stimuli.

The cell-assembly is a hypothetical structure; its physiological existence has not been demonstrated—on the other hand, the concept does not conflict with current neurophysiological knowledge. Moreover, the formation of a cell-assembly rests upon three main premises: (a) the initial existence of the proper inter-connections among the neurons of the system, (b) an initial selective sequence of cortical events that forces the neurons of the system to act briefly as a unit, and (c) the law of the change in synaptic strength between neurons. This latter premise is taken by Hebb as his basic

neurophysiological premise. Stated more fully, it reads:

When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency as one of the cells firing B, is increased.

While there is evidence that is very suggestive, the validity of this hypothesis has not yet been demonstrated neurophysiologically: again it does not conflict with known properties of neurons. It was demonstrated, conclusively shortly after the appearance of Hebb's book (for example, Eccles [2]) that some neurons send out inhibitory connections as well as excitatory connections. Milner [4] argues effectively for the inclusion of inhibitory connections, subject to the same law of effect (c) and his suggestion is adapted here.

It should be noted here, that many properties of cortical neurons are inferred from the known properties of peripheral neurons. There seems to be no reason, at this time at least, for not doing this as it may be some time before techniques are evolved that will allow the fine, detailed study on the cortical neuron that has been carried out on neurons in the spinal ganglia, etc. This is obviously one area where new knowledge will be of the greatest interest in the study of models such as the one developed here.

There is one other premise which, although not explicit in the above formulation of the cell-assembly, is in some respects the most important of all. That is that the system of neurons under consideration be large enough and the inter-connections among these neurons be dense enough such that the probabilities of existence of "the proper inter-connections" in premise (a) above be of a magnitude such that cell-assemblies may actually come into existence. Here the evidence from neuro-anatomy is encouraging: the human cortex has of the order of 10^{10} neurons; (peripheral) neurons have been observed with approximately 1500 synaptic endings on them (i.e., 1500 input lines). Moreover, a given cortical neuron (association layer) seems to send

out connections to all points in the region surrounding it up to a distance of one millimeter away.

Hebb's theory is in some respects a stimulus-response theory, where "response" does not mean immediate (muscular) response. This is reflected most strongly in premise (b), where the "initial selective sequence of cortical events" refers to the "priming" of the initial skeletal pathway assumed in (a) by massive "training" stimuli together with the stimulus which alone is to activate the assembly later on. The massive "training" stimuli may result from a sensation, e.g., hunger, from some environmental feedback, the action of other, already established assemblies, etc.

Thus, referring back to the statement of the problem given above, the main problem reduces to that of testing the role of the cell-assembly in learning—i.e., Hebb's theory—via the digital computer simulation of the models involved. One of the objects of this study is that of giving—in terms of the model—a precise characterization of the formation and the role of a cell-assembly in a learning process. While thus far an intuitive notion of learning has been referred to, it is hoped, that in the context of a well-defined experiment, some account can be given for a non-trivial learning process.

A final observation on the character of cell-assemblies and phase sequences of cell-assemblies is in order: That is, that they allow one to discuss learning and associated problems at a "molar" level (as Hebb puts it)—i.e., in terms of aggregates of neurons, their statistical properties, etc.—just as, for example, in statistical mechanics one works with aggregates of point masses, with little if any attention being paid to the individual bodies of the system.

1.3 THE CORTICAL NEURON AND SYSTEMS OF CORTICAL NEURONS

The advent of the micro-electrode and associated probing techniques in the last fifteen years has allowed physiologists to determine electrical properties of neurons from direct inter-cellular readings and, as a consequence, a wealth of knowledge has been gained about the electrical behavior of neurons, axonal propagation of pulses, etc. Most of this knowledge has been gleaned from studies on non-cortical neurons, e.g., neurons in the spinal ganglia, etc. A good, though slightly outdated, account of this is given in Eccles [2]. It is assumed that the properties of non-cortical neurons carry over to those in the cortex. Histologically, the cortical neuron is a neuron; while direct electrical studies on the cortex are hard to interpret, they tend to support this assumption.

It is manifestly impossible to simulate the real neuron in all its complexity. In fact, even if it were possible to do so, it would probably be unnecessary, as some of the properties of the neuron most likely are unessential to the problem at hand. As in any science, simplifying assumptions have to be made, albeit with great care, trying to retain the most essential properties of the object described. The following description of the neuron is adopted here.

1.3.1 Structure

The gross-structure of the physiological neuron is as follows:

The main part of the organ is the cell body or soma, S, which sends out one fiber called the axon, A, which may later branch out quite profusely. A number of axons from other cells impinge on the soma of the body, sometimes on extensions of the soma—which often are quite profuse—called the dendrites, of the given cell. The point of contact of an incoming (afferent) axon with the soma or dendrites is the synapse and is usually characterized by a nodal

swelling or button-like ending. Moreover, there is a very narrow gap between this ending and the cell body, called the synaptic gap. Neurons have been observed with of the order of 1500 synaptic endings on their soma. A given incoming axon may make contact several times with a given soma. The afferent or incoming axons are, in effect, input lines; the axon sent out from the soma, an output line. Thus the neuron is a multiple input, single output device.

There are neurons of different structure than this, but their use in the nervous system seems to be specialized and not of relevance here (e.g., bipolar neurons in the optic nerve). It should be noted that in the cortex there are neurons with very complex dendritic branching and small—if any—axons as well as neurons with dendritic branching and quite long axons.

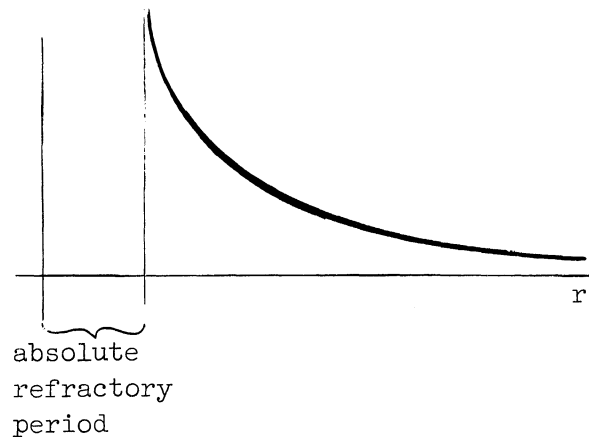
1.3.2 Input and Output, Threshold

The axon of a neuron is capable of transmitting a pulse of electric potential (called the action potential) with no significant decrease in amplitude throughout its length. The pulse originates in the soma of the nerve cell as a consequence of input-pulses on the incoming fibers (synapses) to the cell and spreads down the cell's axon to its various endings. A cell is said to fire when it sends out such a pulse. The neuron (and its axon) is a threshold device in the sense that, as a result of summation of its inputs (at the synapses) and depending on the length of the time interval since the last firing, it either fires completely or not at all, i.e., the amplitude of the outgoing pulse is independent of the magnitude of the input pulses.

The net input to the cell at a given time is determined by the number of impulses present at the synapses at that time and the level of activity (recall hypothesis (c), 1.2 above) at these synapses. Actually summation of this potential activity over a brief interval of time probably takes place.

The inputs thus sum, in a fashion as yet unknown, spatially and temporally. In the model, the inputs (see below) are added. If the summed stimuli exceed the threshold at that time, the neuron fires—if not, it does not fire.

Once the neuron fires, it cannot be made to fire again for a period of time, the absolute refractory period. After that period of time, it maintains a high threshold which gradually decreases to its quiescent or resting value. This time interval, after the absolute refractory period, required for recovery to the quiescent state is called the relative refractory period. Thus, the neuron has the following threshold characteristic:



The time interval since the last firing of the neuron is called the recovery state.

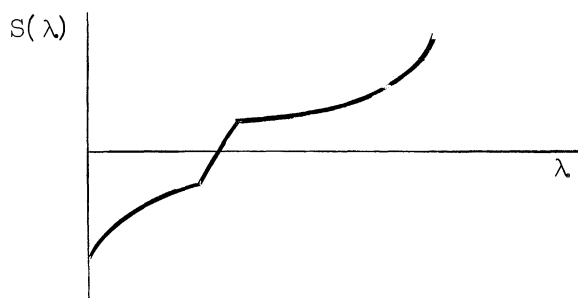
In the model, time is quantized, $t = 0, 1, 2, \dots$. Thus, a neuron fires at time $t + 1$ depending upon

- (1) whether it fired at time t . If it did, then it cannot be made to fire until time $t + k$, where k , a positive integer and a parameter of the system, represents the absolute refractory period.
- (2) whether the sum of the inputs exceeds the threshold at time t . If so, it fires at $t + 1$; otherwise it remains refractory.

(3) a spontaneous firing mechanism which is explained below.

1.3.3 Synapses

The exact nature of transmission across the synaptic gap and summation of the incoming pulses is as yet unknown. Here, it is assumed that each input line has an associated synapse level, λ . This synapse level in turn is used to determine the synapse value for that line, usually by a table of the following sort:



If there are n active input lines, then the total input at time t is $\sum_{i=1}^n S_i(t)$ where $S_i(t)$ is the synapse value corresponding to the i -th line at time t . Notice that in general there will be negative values of the synapse values: these correspond to inhibitory connections.

According to the hypothesis (c), Section 1.2, the synapse levels are subject to a law of effect as follows: suppose there is a synapse from neuron A to neuron B—i.e., neuron A sends, via its axon, one connection to neuron B. Then, if A fires at time t and B fires at $t + 1$, the synapse level from A to B, λ_{AB} , is increased by a uniform amount $\delta\lambda$. If A fires at t and B does not fire at $t + 1$, λ_{AB} is decreased by $\delta\lambda$; otherwise no change in λ_{AB} is made: symbolically,

$$A(t) \ \& \ B(t+1) \ \Rightarrow \ \lambda_{AB} \rightarrow \lambda_{AB} + \delta\lambda$$

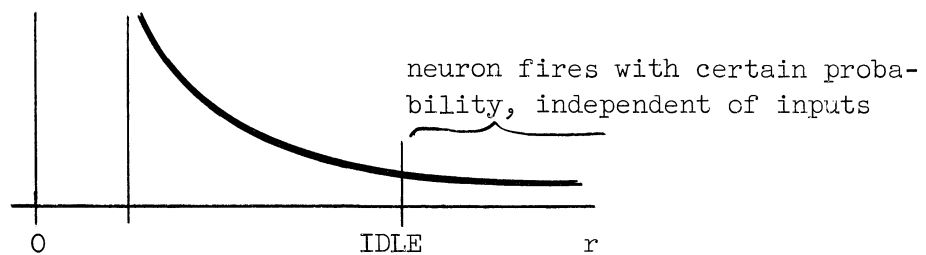
$$A(t) \ \& \ \overline{B(t+1)} \ \Rightarrow \ \lambda_{AB} \rightarrow \lambda_{AB} - \delta\lambda$$

λ ranges in value from 0 to a maximum. In addition to the law stated above,

there is a probabilistic mechanism in the model that serves to "slow down" the λ change. Essentially, if λ is to be changed (i.e., either $A(t)$ & $B(t+1)$ or $A(t)$ & $\overline{B(t+1)}$), then a probability particular to that level is consulted: if it exceeds a certain amount, then the change takes place, otherwise no change occurs. This mechanism can be used to bias the direction of synapse-level change.

1.3.4 Fatigue, Spontaneous Firing

In addition to the threshold function, there is a long term mechanism which delays full-recovery, called fatigue. The evidence for this from neurophysiology, in the case of peripheral neurons, is fairly definite. The fatigue function and its implementation will be discussed at length in a later chapter. The effect of fatigue is one of the subgoals of this study, as is that of spontaneous firing. There is also fairly good evidence that cortical neurons fire spontaneously (see, for example, Sharpless, S. K. and Halpern, L. M. [5]). In the model this is defined as follows: if the recovery state of a neuron exceeds a certain value (called IDLE), then the neuron fires with a certain probability.



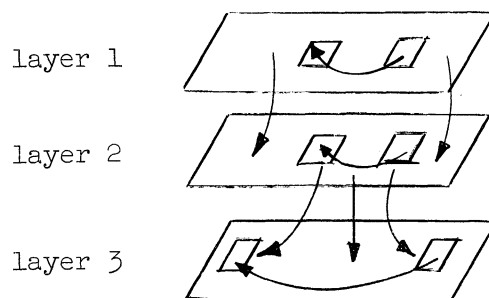
The role of spontaneous firing seems to be essential—it may act as a form of drive if it is a function of the time since the last reward or the like—i.e., a non-specific global disturbance. As the mechanisms of fatigue and spontaneous firing can be defined very exactly in the model, their effects can be studied under tightly controlled conditions.

This completes the discussion of the neuron for this time.

The mammalian cortex consists of several layers of neurons of different structure. The outer layer, for example, consists of neurons with axons which spread out horizontally over large distances; the inner layers consist of neurons with very complex axonal branching in the immediate vicinity of the cell; axons from within the cortex and perhaps those from subcortical structures descend up through all the layers and back down again, probably with complex branching along the way, etc. (see, e.g., Eccles, ibid., pp. 229-331).

Moreover, there are regions of the cortex, into which sensory input is projected (e.g., the visual cortex) and other regions from which motor control is effected.

These features can be simulated to some degree in the model. First of all, a neighborhood relationship for a group of neurons may be defined that determines the neurons to which the neurons of the given group are connected and the density of connections sent out by these neurons. This neighborhood relationship thus permits structuring several layers of neurons with different connections for the different layers as well as inter-layer connection. For example, in the figure below, layer 1 may have very dense local connections, similarly for layer 2, while layer 3 may be more diffuse, neurons sending out connections over greater distances; layer 1 may connect to layer 2 in an approximate one-one fashion, while layer 2 may send out diffuse connections to layer 3, etc.



From the discussion so far, it is evident that there are many parameters and functions that can be varied in the given class of models; threshold function, fatigue, spontaneous firing, neighborhood relation, density of interconnections, etc. Moreover, the relationship between the various possible choices may be complicated and subtle. Hence, the great value of the simulation approach: hypotheses, such as those described in 1.5 concerning such relationships can be tested—hypotheses whose validity (in the models) simply may not be rigorously demonstrable a priori.

1.4 PREVIOUS NEURAL NET STUDIES

This study is not the first in its field. Rochester, Holland, et al. [6] experimented first with a "discrete-pulse model," using a simulation program for the IBM 701, then with an "FM model," using a simulation program for the IBM 704. In the first case, they exhibited "diffuse reverberation," a phenomenon somewhat akin to the sustained activity discovered in isolated cat cortex by Burns [1], but could not demonstrate any tendency on the part of the neurons to form cell-assemblies. While the "diffuse reverberation," in the authors' eyes, might serve as a mechanism for short term memory, they felt that additional structure must be imposed upon the net to allow formation of cell-assemblies. They conferred with Milner and followed his suggestion [4] of introducing negative synapse values into their model. At the same time, taking advantage of the larger and faster IBM 704 computer, they reprogrammed their model in such a fashion that the detailed firing history of the neurons was lost, to be replaced by a frequency of firing for each neuron. This frequency varied with the time, hence the term "FM model." They simulated a net of 512 neurons with six inputs each. In their experiments with this model they observed the formation of cell-assembly-like structures, i.e., sets of neurons such that within each set the connections between the neurons had large,

excitatory synapse values while between the various sets themselves the interconnections had large, inhibitory synapse values. They also observed phenomena somewhat like the fractionation and recruitment of neurons, as required by Hebb's theory. On the other hand, the cell-assembly-like structures they observed could not arouse one another, as Hebb's theory again requires, that is, their model was too environment-dependent.

In later studies with this model, Holland and Rochester demonstrated binary learning (Holland—personal communication). However, for a variety of reasons, the project was abandoned and not resumed by any of its originators. It was continued, however, at the Logic of Computers Group at The University of Michigan, under the supervision and inspiration of John H. Holland, by J. W. Crichton [7][8].

Crichton and Holland [8] proposed a new method of simulating neural nets which took advantage of the increased storage of the IBM 704 computer and which would allow simulation of up to 2000 neurons with about 150 inputs per neuron. This gives rise to the so-called "variable-atom" model, in which all neurons with the same characteristics (i.e., firing history, threshold, fatigue, etc.) are lumped together into an "atom." Computation of the number of active inputs to a neuron is performed by reference to appropriate Poisson tables.

This model was never simulated on the IBM 704. The availability of an IBM 709 computer, a machine which represented a considerable advancement over the IBM 704 in that much improved input-output equipment and procedures were available and new powerful operation codes were added, caused a major change in plans and the model was to be reprogrammed for the IBM 709, taking advantage of its new features. Crichton was joined by Finley at this point.

Crichton and Finley modified the model, putting it in almost the form in which it is used in this study and programmed it for simulation on the

IBM 709 [9]. Early experiments with this model revealed the distressing fact that the model was not capable of sustained activity such as Burns observed [1]. Stimulated "slabs" would not maintain activity indefinitely, but in fact died down rather rapidly. Marked epileptic behavior resulted—that is, intense activity alternated with low activity, leading rather quickly to "death," i.e., no activity at all. No modification of the net parameters seemed to produce a cure for this behavior and we were forced to re-appraise the whole model. This led to the discovery that the statistical techniques used in the model contained a fatal flaw, basically that it would not allow a small number of neurons to produce a sufficient stimulus to fire a single neuron. Several modifications of the original technique were tried with little success. This forced us back to basic principles and led to the implementation of a new technique aimed at introducing greater statistical disuniformity into the model. It is on the basis of this modified model [10] that the study to be described here is based.

Crichton has developed in the appendix to his doctoral dissertation [7] an interesting and fruitful analysis which is especially useful in considering large systems of neurons and their interactions. This will be referred to later on in this study.

1.5 PLAN OF RESEARCH

This study represents the first of several stages of the long-range study and is concerned only with simple, cycle-less nets in which a single neuron, C, is presented with inputs from two sets, A and B, of neurons. The behavior of neuron C depends upon the average firing rates of the neurons of A and B respectively. The essential hypothesis is that the firing of neuron C will correlate with "patterned" versus random inputs, and this will be seen to be the case. At the same time, this simple configuration of neurons all

connected to a single neuron provides an opportunity to study in detail the basic neuron parameters, i.e., threshold, fatigue, etc.

The second stage, not reported here, will be concerned with a generalization of the first obtained by introducing progressively more complicated feedback cycles and replacing the single neuron C by a set of neurons. Again the hypothesis of "patterned" inputs applies and one is led rather naturally to alternation experiments where, for example, group A will be active, suppressing activity in B, thus controlling C, then become less active, allowing B to become active, in turn forcing A into inactivity and gaining control of C, etc.

In both cases, the experiments are graduated, going from the simple to the complex. The theory for each stage is developed separately and its relationship to the general theory indicated. Likewise, the feedback from experiment to theory—an essential component of a work of this sort—is indicated as the occasion arises.

2. FORMAL DESCRIPTION OF THE MODELS

2.1 INTRODUCTION

In Section 1.3 a general description was given of cortical neurons and systems of cortical neurons together with the abstraction of properties describing the neurons of the model. The discussion was informal, going from certain salient known properties of cortical neurons to their abstracted counterparts. In this chapter, the structure and operation of the models of the class being considered are defined formally. The notions of run and experiment are clarified and, using the network equations, the abstract prototype for all experiments is given. Recursive definitions are given for the various network functions, such as threshold, fatigue, etc. Following this, in the next section, an attempt is made to clarify the role of the various functions and to display possible functional forms for them, though no attempt is made at this point to give formal derivations. Finally, a note is given on the network simulation program, followed by a reference list of symbols used in this chapter.

2.2 THE NETWORK EQUATIONS

A neural network, of the class of models considered in this study, consists of a set of N elements called neurons with a set of specified directed connections between these neurons, where "directed" implies, for example, that neuron A may send a connection to neuron B, but not conversely, i.e., there is a connection A-to-B, but not B-to-A. Such a connection is referred to as the output of A, the input to B. A neuron of the model may have many inputs, but it always has only one output; however, this output may branch and go to several neurons, including the source neuron, as inputs or go to the environment. All that is external to the network itself but which influences,

and is influenced by, the network, is called the environment. Thus, in general the environment will supply input to selected neurons of the network and receive output from selected neurons. Included in this concept of environment would be, for example, reflex mechanisms, a simulated biological environment, a human observer, etc.

Time is quantized in these models, $t = 0, 1, 2, 3, \dots$. At any time t , the state of the network, $S(t)$, is determined by the functions (see below) performed by the model; likewise the state of the environment, $E(t)$, is determined. From $S(t)$ plus the input to the network at t from the environment, $I(t)$, is determined the state at time $t + 1$, $S(t + 1)$. Also, $S(t)$ determines the output at t to the environment, $O(t)$, and we have symbolically

$$S(t + 1) = F_N(S(t), I(t)) \quad (t = 0, 1, 2, \dots) ,$$

where F_N is the state-transition function for the network. (In general, F_N is far too complicated to define explicitly, however it is defined implicitly by the network equations given below.) Likewise, $E(t + 1)$, the state of the environment of time $t + 1$, is determined by $E(t)$ and $O(t)$ and again

$$E(t + 1) = F_E(E(t), O(t)) \quad (t = 0, 1, 2, \dots) .$$

Since $I(t) = g(E(t))$, for some function g , then

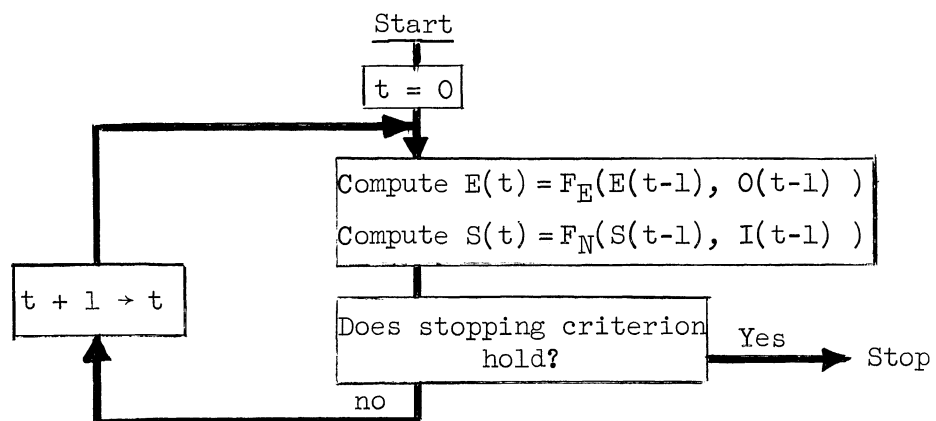
$$\begin{aligned} S(t + 1) &= F_N(S(t), I(t)) \\ &= F_N[S(t), g[F_E(E(t), O(t))]] . \end{aligned}$$

This is a recursive equation for $S(t)$; $S(0)$ and $E(0)$ form the initial conditions for the network and the environment respectively. Given $S(0)$ and $E(0)$, and a starting signal, the network and environment proceed automatically over the time steps $t = 0, 1, 2, \dots$ until a stopping condition,

determined in the environment, is reached. Notice that the cycle, network \rightarrow environment \rightarrow network, forms a closed feedback loop. The procedure of running the system \langle network, environment \rangle , given a $S(0)$ and $E(0)$, from $t = 0$ or $t = t_0 (\geq 0)$ down to a t_f will be called a run. The sequence of outputs $O(0)$ (or $O(t_0)$), ..., $O(t_f)$ form the behavior of the network. However, the term "behavior" will be used in the broader sense of reaction of the network to the environment. The specification of a network-environment pair, the initial conditions, and a set of hypotheses about the behavior of the network constitutes an experiment.

Thus, the abstract prototype of all experiments has the following structure:

(Given: Behavioral Hypotheses, $S(0)$, $E(0)$)



As mentioned, the state-transition function is too complicated to be defined explicitly and must be defined implicitly. This is done as follows: At any time t , a neuron may fire or not fire. If it fires, it puts a 1 at its output, if not—a 0. The set of neurons that fire at time t , together with input from the environment, will determine the set that fire at $t + 1$. The condition for the firing of the i -th neuron at time $t + 1$ is given as a

recursion relative to the real-valued functions R, F, S, and I which in turn are defined relative to recursions on $r_i(t)$, $l_i(t)$ and $\lambda_{ji}(t)$ by the functions V, ϕ , S, and I. Once these functions are given, then the behavior of the net is determined for all t from the initial conditions. This condition is

$$T_i(t) : \left[\delta_i(t + 1) = 1 \right] \equiv [R_i(t) \cdot F_i(t) < \sum_j S_{ji}(t)\delta_j(t) + I_i(t)]$$

(i = 1, 2, ... , N)

where $\delta_i(t) = 1$ means "neuron i fired at t." Thus, T says that neuron i fired at t + 1 if and only if the condition

$$R_i(t) \cdot F_i(t) < \sum_j S_{ji}(t)\delta_j(t) + I_i(t)$$

holds. $R_i(t)$ and $F_i(t)$ are the threshold and fatigue values of neuron i at time t respectively. $S_{ji}(t)$ is the weight or synapse value of the directed connection from neuron j to neuron i at time t. For neurons j which do not send connections to neuron i, S_{ji} may be considered as equal to zero. $I_i(t)$ is input to neuron i at t from the environment; it will be referred to as the pre-stimulus to neuron i. R, F, S, and I are all real numbers; R and $F \geq 0$, S and I either \geq or < 0 . Negative values of S are called inhibitory inputs, positive values are called excitatory. They are defined recursively as follows:

$$R_i(t) = V(r_i(t))$$

where V, the threshold function, is a real-valued function of $r_i(t)$; $r_i(t)$ is the recovery-state of neuron i at t defined as follows:

$$r_i(t) = \begin{cases} 0 & \text{if } \delta_i(t) = 1 \\ r_i(t-1) + 1 & \text{if } \delta_i(t) = 0 \\ r_{\max} & \text{if } \delta_i(t) = 0 \text{ \& } r_i(t-1) = r_{\max} - 1 \end{cases}$$

For $r_i(t) = 0, \dots, r_\alpha$, $V(r_i(t)) = \infty$. r_α is the absolute refractory period; i.e., if $\delta_i(t) = 1$, then neuron i cannot fire again until $t + r_\alpha + 1$. Note that the function V is the same over all neurons of the net.

$$F_i(t) = \phi(l_i(t))$$

where ϕ , the fatigue function, is a real-valued function of $l_i(t)$; $l_i(t)$ is the fatigue-level of neuron i at t defined as follows:

$$l_i(t) = \begin{cases} l_i(t-1) + \Delta_2 & \text{if } \delta_i(t) = 0 \\ l_{\max} & \text{if } \delta_i(t) = 0 \text{ \& } l_i(t-1) = l_{\max} \\ l_i(t-1) - \Delta_1 & \text{if } \delta_i(t) = 1 \\ l_{\min} & \text{if } \delta_i(t) = 1 \text{ \& } l_i(t-1) = l_{\min} \end{cases}$$

where $\Delta_1 > \Delta_2 > 0$. Δ_1 and Δ_2 are extremely important parameters, determined from the nominal system firing rate or frequency, f_b , by the relation

$$f_b = \frac{\Delta_2}{\Delta_1 + \Delta_2}$$

$$S_{ji}(t) = m_{ji}S(\lambda_{ji}(t))$$

where S is the synapse-value function, taking positive, negative, and zero values, m_{ji} is the multiplicity of the connection $j \rightarrow i$, while $\lambda_{ji}(t)$ is the synapse-level of the connection $j \rightarrow i$ at time t . It is defined as follows:

$$\lambda_{ji}(t) = \begin{cases} \lambda_{ji}(t-1) + 1 & \equiv \delta_j(t-1) = 1 \text{ \& } \delta_i(t) = 1 \text{ \& } \rho_i(t) > U(\lambda_{ji}(t-1)) \\ \lambda_{ji}(t-1) - 1 & \equiv \delta_j(t-1) = 1 \text{ \& } \delta_i(t) = 0 \text{ \& } \rho_i(t) > D(\lambda_{ji}(t-1)) \\ \lambda_{ji}(t-1) & \text{otherwise.} \end{cases}$$

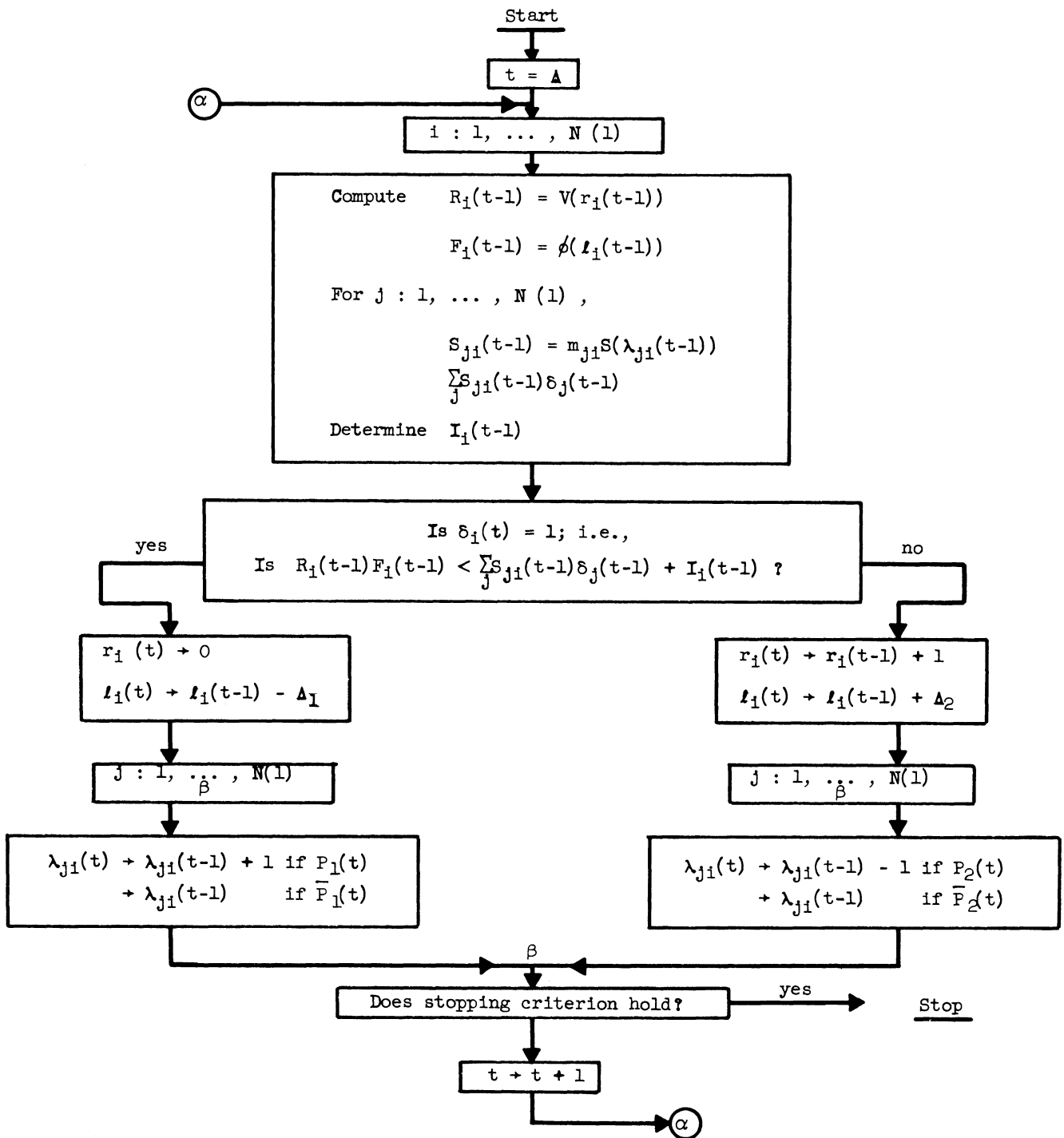
$\rho_i(t)$ is a number drawn randomly and independently for all i and t from the open interval $(0, 1)$. $U(\lambda)$ and $D(\lambda)$ are the probabilities of change up and change down of synapse-levels respectively; notice that U and D in general vary with λ . If $\lambda = \lambda_{\max}$, then $U(\lambda) = 0$; if $\lambda = \lambda_{\min}$, then $D(\lambda) = 0$. The condition $\rho_i(t) > U(\lambda_{ji}(t-1))$ says simply that $\lambda_{ji}(t-1)$ is incremented by 1 with probability $U(\lambda_{ji}(t-1))$ at t . As with Δ_1 and Δ_2 , U and D are extremely important quantities, and relate to the nominal system average f_b as follows:

$$f_b = \frac{U(\lambda)}{U(\lambda) + D(\lambda)} \quad (\text{for all } \lambda) .$$

The law for incrementing or decrementing λ is the implementation in the models of Hebb's law of effect for synapse change.

The multiplicity m_{ji} of the connection $j \rightarrow i$ determines the density of the connection $m_{ji} = 0, 1, 2, \dots$. $m_{ji} = 0$ corresponds to the case of no connection from j to i . Specification of the set of m_{ji} 's for all i, j essentially determines the connection scheme of the model at hand.

Thus, with these recursive definitions in mind, the flow-chart given above representing the abstract prototype of all experiments takes on the following more specific form: (Given: behavioral hypotheses, $r_i(0)$, $l_i(0)$, $\lambda_{ji}(0)$ for all $i, j = 1, \dots, N$.)



In this diagram, the notation "A→B" means that the value of A is replaced by the value of B; "i = 1, ... , N(1)" means that the computation from the occurrence of this statement down to the point β is first done for i = 1, then repeated for i = 2, i = 3, ... , down to i = N. (This is just a "loop" on the index i.) $P_1(t)$ is the condition for incrementation of $\lambda_{ji}(t)$ given earlier, $P_2(t)$ that for decrementation.

2.3 THE NETWORK FUNCTIONS R, F, AND S

In the preceding section, a formal characterization of the functions R, F, and S was given, with no attention being paid to their specific analytic forms. As was mentioned in the Introduction, the study of these forms is a subgoal of this paper, since prior to this there has not been a rigorous demonstration for any one of these functions assuming a given functional form. Since these functions may be specified as one wills, they in fact are parameters of the network in the sense that given a specification of these parameters a specific model of the class under consideration is determined.

2.3.1 Control of Firing Rate

From the network equations $T_i(t)$ one can see that the function of the threshold value $V_i(t)$ of a neuron, as modified by the multiplicative factor $F_i(t)$, is to determine whether or not neuron i of the network fires at t. If the combined input to the neuron is at least as great as the product of $R_i(t)$ and $F_i(t)$, then neuron i fires, otherwise it does not. The function V, which determines R, then controls the firing rate of the neurons of the net. Immediately after neuron i fires, V is infinite and i cannot fire. After a few time steps (r_α —the absolute refractory period), it "recovers" slightly, that is a very large input stimulus can cause it to fire, after a few more, less stimulus is required, down to the point where—if it has not yet fired, a minimal stimulus is required to cause it to fire. This point is called the

resting or quiescent value of V . The function ϕ which determines $r_i(t)$ modulates the control of V in the sense that if the firing rate of neuron i is high, then ϕ is large, hence larger stimulus is required to cause i to fire. If the firing rate is low, the magnitude of ϕ is small (close to 1) and less stimulus, depending upon the value of V , is required.

2.3.2 The Threshold Function

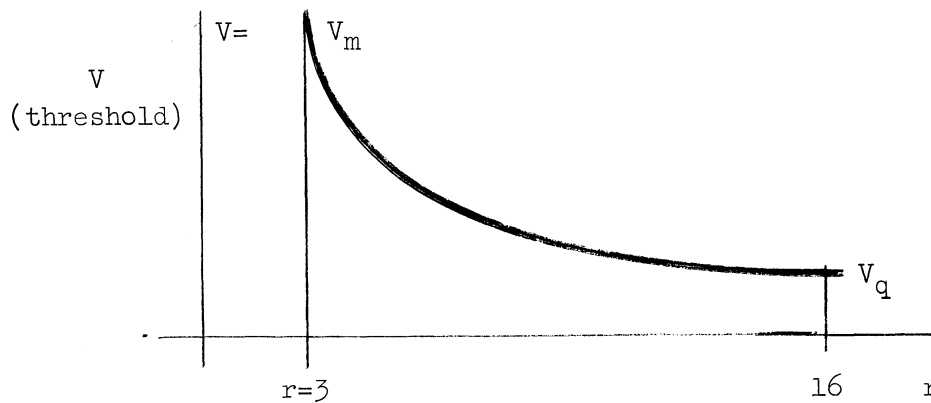
From 2.2 one sees that the threshold value, $R_i(t)$, of neuron i is that value which corresponds to the recovery state r_i of neuron i ; that is, r_i = the number of time steps since neuron i fired. Each neuron i of the network has associated with it a value of r_i , depending on its immediate firing history. Thus, if $\delta_i(t) = 1$ (i.e., neuron i fired at time t), then $r_i(t) = 0$; if $\delta_i(t-10) = 1$, and $\delta_i(t-9) = 0, \dots, \delta_i(t-1) = 0, \delta_i(t) = 0$ (i.e., neuron i fired at $t-10$ and did not fire again up to and including time t), then $r_i(t) = 10$. Each time neuron i fires, r_i is set to zero. Each time it fails to fire, it is incremented by 1, i.e., $r_i(t) = r_i(t-1) + 1$. r_i has a maximum of 16, further incrementation fails to change it—i.e., $16 = 16 + 1$. The function $V(r_i(t))$ which gives the value $R_i(t)$ is universal over the net, that is all neurons i conform to it. Because, at any given time t , these neurons may have distinct values of $r(t)$, they will usually have distinct threshold values.

The absolute refractory period or period of infinite threshold, r_α , is taken to be two time steps. That is, if $\delta_i(t) = 1$ (neuron i fires at t), then i cannot fire again until $t + 3$ (until $r_i = 3$). The total number of time steps to quiescence, that is, the resting values of threshold, is 16. Thus, if neuron i fires at t , it is fully recovered (has reached the resting value) at $t + 16$.

There are at least three additional important aspects of the threshold

curve. The first is its value at $r = 3$, the second is its quiescent value—i.e., its value for $r = 16$, and the third is its functional form (i.e., exponential, quadratic, linear, etc.) especially in the recovery range $r = 5, 6, \dots, 10$. A formal derivation of the analytic expression of the threshold curve will be given later (see Chapter 3), in the order in which it was discovered. Note that the reciprocal of the recovery, $1/r$, averaged in some appropriate fashion, will correspond to the firing-rate of the neuron. For example, if a neuron fires on the average once every five times, its "average" recovery is $\bar{r} = 5$ and its firing rate = $1/5 = 1/\bar{r}$.

The threshold curve, then, has the following form, where V_m = the maximum value (for $r = 3$), V_q = the quiescent value ($r = 16$):



The functional form of this curve, the quantities V_m and V_q as well as the initial values of r for each neuron of the net, will be specified for each experiment. The quantity V_q is important because it defines the least amount of input stimulus (synapse-value) which may fire the neuron. In the first experiments, prior to its analytic derivation, the threshold curve was assumed to be an exponential curve of the form $V = ae^{r-3} + b$ where a and b are constants (> 0). The reason for this assumption is two-fold: (a) in the physiological situation, the cell-body is a membrane which may be as-

sumed to have electric properties similar to the axonal membrane and (b) the recordings off of real neurons of the recovery to quiescent values for their cell potentials look to be of an exponential nature. As we shall see later, (b) is nearly correct; (a), being as it is a tenuous inference, we might expect not to be wholly true.

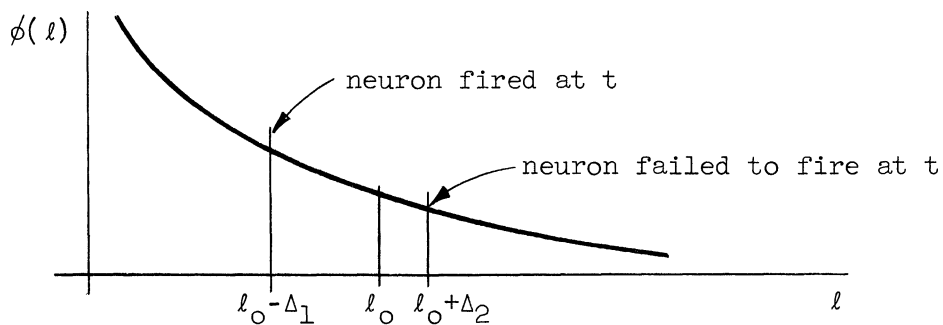
2.3.3 The Fatigue Function

As already mentioned, the fatigue value $F_i(t)$ serves to modulate the threshold value $R_i(t)$ of neuron i and hence modulates the firing rate of i . The desired effect of the fatigue function is as follows: given the neuron in a fully recovered state, that is, the threshold value is near V_Q and the fatigue value is 1, then suppose inputs are presented to the neuron so as to cause it to fire at a fairly high rate (above the background rate f_b). Then, gradually over a period of 50 to 100 time steps the fatigue value, i.e., $\phi(l)$, of the neuron increases in such a fashion as to cause the firing rate of the neuron to drop back to f_b and keep it there as long as the given inputs are present. Suppose next the inputs themselves drop off so that at the most they would cause the neuron to fire at f_b . Then, its fatigue value, $\phi(l)$ decreases slowly back to 1 so as to preserve approximately the average firing rate of f_b . Intense activity of the neuron, that is, firing at near maximal rates, produces a more abrupt increase in ϕ , whereas sudden drop-off in activity, that is, firing at very low rates ($<f_b$) produces a more abrupt decrease in ϕ .

The fatigue value $F_i(t)$ of neuron i is determined by the fatigue function ϕ from the fatigue level $l_i(t)$ of neuron i at time t , $F_i(t) = \phi(l_i(t))$. The function ϕ is universal for all neurons of the net and similar remarks for the variation in threshold values among the neurons of the network apply to the fatigue values as well. The fatigue value is used, as has been indi-

cated, as a multiplicative factor of the threshold value for the given neuron. ϕ is a monotonically decreasing function of l with $\phi \geq 1$. The larger the ϕ , the larger the product $R \cdot F$. Thus, neuron i may be fully recovered, $r_i(t) = l\phi$, and $R_i = V(r_i(t)) = V_q$, but ϕ may be so large that $R_i(t) \cdot F_i(t) = V_q \cdot \phi(l_i(t))$ is much greater than V_m . Fatigue is rendered ineffective by setting $\phi(l) = 1$ for all l . Then $R_i(t) \cdot F_i(t)$ always equals $R_i(t)$. Note that the fatigue value has no effect on the absolute refractory period ($\phi \cdot \infty = \infty$).

The quantity l for a given neuron varies incrementally from 0 to 32 with $1/32$ as the smallest possible increment. The manner of variation is the following: Suppose the neuron has fatigue level l_0 at time t . Then, if the neuron fires at t , l_0 is decremented by a quantity Δ_1 , i.e., $l_0 \rightarrow l_0 - \Delta_1$. If it does not fire at t , then it is incremented by a quantity Δ_2 , i.e., $l_0 \rightarrow l_0 + \Delta_2$. This is illustrated below:



In general, $\Delta_1 > 0$, $\Delta_2 > 0$, and $\Delta_1 > \Delta_2$. Decrementation below 0 and incrementation above 32 have no effect, i.e., $0 - \Delta_1 = 0$, $32 + \Delta_2 = 32$.

Δ_1 and Δ_2 are extremely important numbers since in terms of them is expressed a crucial parameter of the net, namely the firing rate at which a neuron experiences no net change in fatigue level. Thus, if a neuron is firing at this rate—call it f_b —then over an interval of length T time steps, say, there is no net change in the l for that neuron and we have—recalling that $f_b T$ is the number of times a neuron fired in the given inter-

val and $(1 - f_b)T$ the number of non-firings--

$$\Delta_1 f_b T - \Delta_2 (1 - f_b) T = 0 .$$

Solving for f_b gives

$$f_b = \frac{\Delta_2}{\Delta_1 + \Delta_2} .$$

This quantity, f_b , already mentioned, is called the background firing rate or the nominal system average. It will be treated in detail later on. Note that given f_b , one can determine Δ_1 and Δ_2 (up to a constant multiple $k > 0$ which may be chosen as 1) and, conversely, given Δ_1 and Δ_2 , f_b is uniquely determined. f_b plays an important role in Crichton's theory [7].

The functional form of the fatigue curve, the numbers Δ_1 and Δ_2 (or, f_b), as well as the initial value of l for each neuron of the net, will be specified for each experiment. The form of the curve is clearly of the greatest importance since it, together with the numbers Δ_1 and Δ_2 , determine the recovery rate of a neuron as well as its fatiguing rate. The desired properties of this curve have been outlined above. The rationale for this will be given in the next chapter. For the early experiments, an exponential curve of the form $\phi = ae^{-bl} + c$ was used. However, as will be discussed in detail, in the order of its discovery, this form will not work and, in fact, ϕ must be a double-valued, hysteresis function.

2.3.4 The Synapse Value Function

Suppose a neuron j sends one directed connection to another neuron i . As we have seen (1.3.3), to each such directed connection at time t is associated a positive number, the synapse level, $\lambda_{ji}(t)$. Just as with the recovery states and fatigue levels, λ is used to determine a value, the synapse value, S , by means of a functional relationship. λ has a range from 0 to 15.

It is incremented according to the "law of effect" as follows: suppose the connection $j \rightarrow i$ has the synapse level λ . Then, if j fired at $t - 1$ and i fired at t , $\lambda_0 \rightarrow \lambda_0 + 1$, with probability $U(\lambda_0)$. If j fired at $t - 1$ and i did not fire at t , then $\lambda_0 \rightarrow \lambda_0 - 1$, with probability $D(\lambda_0)$. Otherwise, $\lambda_0 \rightarrow \lambda_0$ —i.e., no change. If $\lambda_0 = 0$, no further decrementation is allowed; if $\lambda_0 = 15$, no further incrementation is allowed. The statement " $\lambda_0 \rightarrow \lambda_0 + 1$ ($\lambda_0 - 1$) with probability $U(\lambda_0)$ ($D(\lambda_0)$)" means that if λ_0 is to be increased (decreased)—depending upon whether j fired at $t - 1$ and i at t , etc.—then the incrementation takes place with probability $U(\lambda_0)$ ($D(\lambda_0)$). In general, U and D are assumed to be uniform over all values of λ , i.e., for $\lambda_0 = \lambda_1$, $U(\lambda_0) = U(\lambda_1)$, $D(\lambda_0) = D(\lambda_1)$, etc., with the exception that $U(15) = 0$ and $D(0) = 0$. Note that the incrementations with probability $U(\lambda_0)$ or decrementations with probability $D(\lambda_0)$ form independent trials, e.g., if the synapse level from j to i is λ_0 and that from k to l also equals λ_0 , both j and k fired at $t - 1$, and both i and l fired at t (hence λ_{ji} and λ_{kl} are both candidates for incrementation), then the probability $U(\lambda_0)$ is consulted independently in each case.

The numbers U and D are of great importance, especially in light of the theory developed by Crichton mentioned above. Like the numbers Δ_1 and Δ_2 of the preceding section, U and D are related to the nominal system average, f_b . The reason is quite simple: Assume that the rate of change up of a synapse is proportional to U , say $= kU$, likewise that the amount of change down is proportional to D , say $= kD$. f_b is again defined as that firing rate for which no net change in λ between A and B will occur, assuming for the moment that neurons A and B are firing randomly and independently at the rate f_b . If this is the case, then f_b^2 will represent the probability that the firing of A at $t - 1$ is followed by the firing of B at t ("success"), likewise

$f_b(1 - f_b)$ is the probability that a firing of A at $t - 1$ is followed by a non-firing of B at t ("failure"), $f_b^2 T$ is the number of "successes" over a time interval of length T , $f_b(1 - f_b)T$ the number of "failures." $kUf_b^2 T$ is the net change up in the interval of length T ; $kDf_b(1 - f_b)T$ the net change down. By assumption, the difference of these is zero and

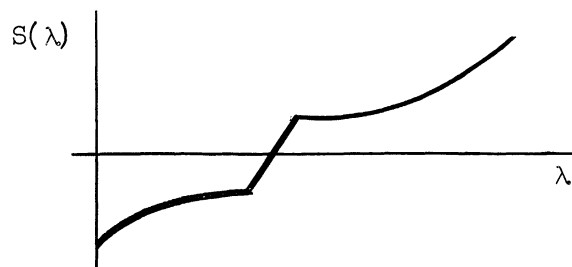
$$Uf_b^2 = D(1 - f_b)f_b$$

or

$$f_b = \frac{D}{U + D}.$$

For the initial experiments in the sequel, ad hoc values of U and D were used; more detailed discussion will be postponed to a later section corresponding to the time at which the absolutely crucial character of these parameters was made evident.

Recall that the firing or non-firing of a neuron is determined by a comparison of the sum of the synapse values on the active inputs (that is, those connections coming from neurons which fired the preceding time step) with the product $R \cdot F$ (which is infinite if the neuron has fired at one of the previous two time steps). If this sum is less than $R \cdot F$ the neuron does not fire, otherwise it does. No restriction of the synapse values has been placed. However, synapse values for small λ 's are assumed to be negative, large λ 's—positive. The negative synapse values for active input lines correspond to inhibitory inputs to the neuron. S is assumed to be a monotonic increasing function of λ , e.g.,



It turns out, just as for the fatigue function, that this type of function is inadequate and that S must also be given as a hysteresis function.

2.4 NOTE ON THE SIMULATION PROGRAM

A diagram representing the operation of the network, given an environment, the initial conditions, and the behavioral hypothesis was given earlier. A program was written for the IBM 7090 computer which simulates the operations indicated in that diagram. This program consists of four basic parts: (1) the lists which describe the state of the net at each time step. The lists are a block of reference information for (2) below and in turn consist of two parts: (a) a permanent part which is never changed in the course of a run, and (b) a volatile part which may change; (2) the net program which computes at each time step the various functions required by the model, referring to the lists for parameter values and making appropriate changes to the lists; (3) the executive and environment routine, a supervisory program which performs two functions: (a) it monitors pertinent net parameters, running time of the program, etc., and handles the appropriate output editing and (b) simulates the environment of the model—i.e., computes input and output functions, making any necessary changes to lists; (4) input-output editing and other special-purpose routines, usually slaves of the executive routine.

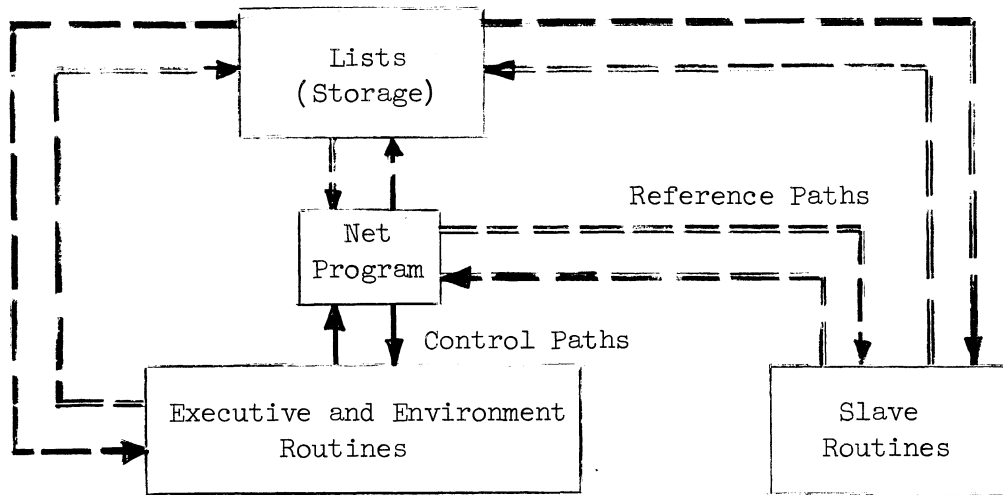
The net program seldom ever will be varied: the executive and environment routines will vary from experiment to experiment and often from run to run. Parameters in the lists will vary from run to run in general, while those lists particular to a given experiment will vary from experiment to experiment. It is the lists that determine the structure of a given net: i.e., neuron inter-connections, density of connections, etc.

Note that the executive routine contains provisions for experimenter intervention in an experimental run. Thus, the experimenter, while watching a

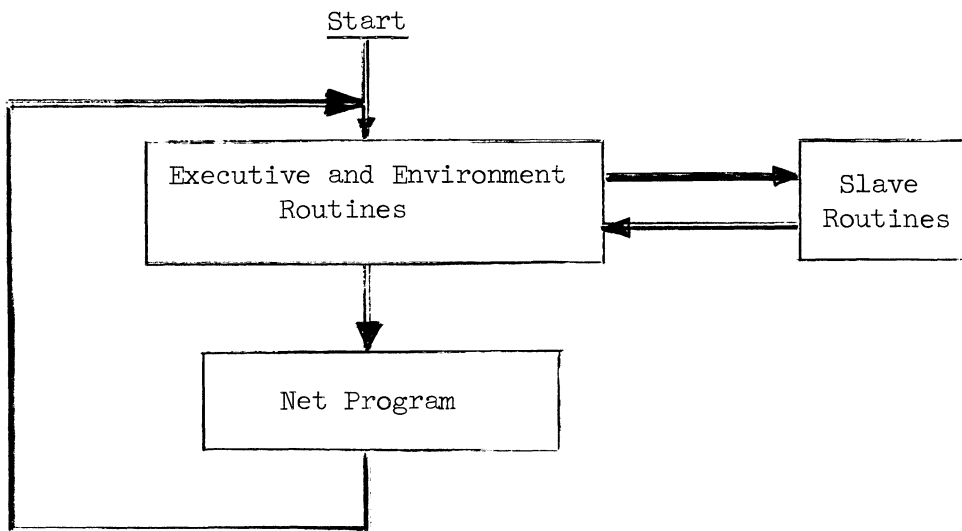
real-time display of selected functions of the network, may at any time change the display, modify parameters, store the entire state of the system for future back-up purposes, etc.

Diagrams giving the overall structure of the program and the flow of control are given below:

Structure of Program



Flow of Control



Symbols Used in Section 2.2

$S(t)$	state of the network at time t
$E(t)$	state of the environment at time t
$I(t)$	input to the network from the environment at time t
$O(t)$	output from the network to the environment at time t
F_N, F_E	state transition function of the network and the environment, respectively
$\delta_i(t) = 1$	the statement "neuron i fired at time t "
$T_i(t)$	the condition for $\delta_i(t) = 1$
$R_i(t)$	threshold-value of neuron i at time t
$F_i(t)$	fatigue-value of neuron i at time t
$S_{ji}(t)$	synapse-value of the connection from neuron j to neuron i at time t
$I_i(t)$	input to neuron i at time t from the environment
$r_i(t)$	recovery state of neuron i at time t
$l_i(t)$	fatigue-level of neuron i at time t
$\lambda_{ji}(t)$	synapse-level of the connection from neuron j to neuron i at time t
$V(r_i(t))$	threshold function, gives $R_i(t)$ as a function of $r_i(t)$, $R_i(t) = V(r_i(t))$
$\phi(l_i(t))$	fatigue function, gives $F_i(t)$ as a function of $l_i(t)$, $F_i(t) = \phi(l_i(t))$
$S(\lambda_{ji}(t))$	synapse-value function, gives $S_{ji}(t)$ as a function of $\lambda_{ji}(t)$, $S_{ji}(t) = S(\lambda_{ji}(t))$
$\rho_i(t)$	random number associated with neuron i at time t
Δ_1	fatigue-level change if $\delta_i(t) = 1$
Δ_2	fatigue-level change if $\delta_i(t) = 0$

f_b	nominal system frequency or average background frequency
$U(\lambda_{ji}(t))$	probability of change up for synapse-level $\lambda_{ji}(t)$
$D(\lambda_{ji}(t))$	probability of change down for synapse-level $\lambda_{ji}(t)$
m_{ji}	multiplicity of the connection from neuron j to neuron i

Symbols Used in the Flow-Diagram

$i = 1, \dots, N(1)$ β	"loop" to N times, starting at $i = 1$, incrementing i by 1 each time; i.e., first $i = 1$, then $i = 2, 3, \dots$, etc. up to $i = N$
$A \rightarrow B$	replace the value of A by the value of B
$P_1(t)$	the condition for incrementing $\lambda_{ji}(t)$
$P_2(t)$	the condition for decrementing $\lambda_{ji}(t)$

3. CORRELATION EXPERIMENTS, CYCLE-LESS CASE

3.1 INTRODUCTION

In the implementation of Hebb's theory, several questions may be isolated in an attempt to elucidate the nature of the cell-assembly. Perhaps the first of these concerns identification of cell-assemblies, that is, in terms of the given models, what are the criteria for cell-assembly-ness? This question is aimed at a static, structural condition and may be paraphrased as follows: suppose a model is given in which it is suspected that cell-assemblies have formed. How, then, does one identify them? The second question (which, causally speaking, should be first) is concerned with the formation of cell-assemblies: i.e., in terms of the given models, how does such a structure (as yields a cell-assembly) come into existence? This question is aimed at dynamic, structural changes and goes hand-in-hand with a third: what are the stability conditions, in the given models, for cell-assemblies? To make this last question more meaningful, the informal description of cell-assembly given in 1.2 is augmented as follows: One may regard a cell-assembly as a union of a large number of reverberatory circuits (in the Lorente de N^o sense of the term), any several of which may be active for a very brief period of time and interrelated so that while any one of the circuits may be rapidly extinguished (within 1/100th of a second in the physiological situation), yet for a much greater period of time (several seconds or longer) the structure as a whole is active in the sense that at least one of the component circuits is active. That is, within a given cell-assembly there are a number of alternate pathways which perform the same function. Therefore, the stability question for such a structure is absolutely crucial; yet, this character of the cell-assembly accounts for the fact that the loss or damage of part of a fully-

developed cell-assembly need not impair its overall function, thus for the seemingly small effect in some cases of brain damage upon learning ability and memory. (This is part of Hebb's dual trace memory mechanism and accompanies his postulate of synapse growth since the reverberatory activity would assist to retain memory temporarily while at the same time it would facilitate the long-run growth changes necessary for permanent memory (see [3], pages 60-78, in particular, p. 62).

Thus the cell-assembly gets us away from a strict dependence (in the cortex) upon individual neurons. Yet for its growth and development the cell-assembly depends upon the law of effect (Hebb's neurophysiological postulate) and upon the availability of neurons which can be "recruited" to the assembly when they act in synchronization with it and likewise which can be dropped out of the assembly (fractionation) when they fall into disuse.

The ability or non-ability of the models to allow recruitment of neurons to an assembly or fractionation of neurons away from it, then poses a fourth question which is taken as the starting point of this study: do the neurons of the models have the ability to be recruited into an assembly when presented with the same input patterns and, dually, to fall away through disuse? This question leads, as shall be shown in the next section, to simple networks which are extremely useful for studying the behavior of single neurons and small groups of neurons.

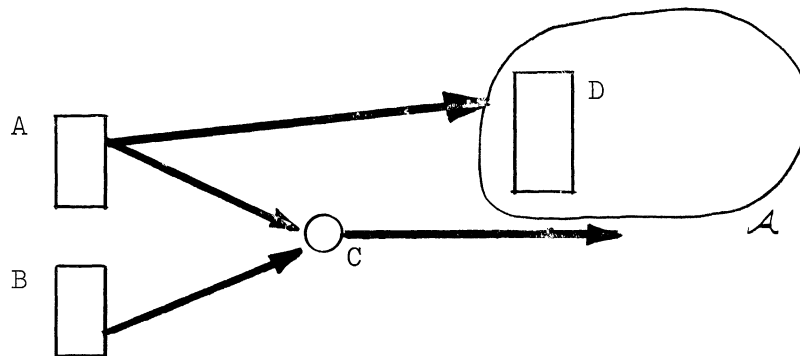
Crichton, in the appendix to his thesis [7] has discussed the stability of cell-assembly-like structures, called by him "semi-autonomous subsystems," and some results of his analysis will be referred to later on.

3.2 CORRELATION

The behavior of a neuron of the model depends upon its input history (which includes synapse value changes on the input lines) and upon its in-

ternal state changes (threshold, fatigue). To determine the response of a given neuron to a particular input pattern, one has to take into consideration the effect of this pattern upon the internal state changes of the neuron and the relationship of this pattern to any other inputs the neuron may have. Basically, therefore, the behavior of a neuron may be regarded as being determined by some function over the totality of its inputs.

Consider now a situation in which recruitment might occur. Let C be an uncommitted neuron of the system and suppose it is presented with a patterned input from a source A of neurons. (A might be, for example, a set of neurons of area 17, reflecting a direct sensory input from the retina.) Lump all the other inputs to C into a group B. Now it might be that A directly affects a system of neurons D, which I will assume form part of a cell-assembly \mathcal{A} . The synapse values from the neurons of A to C will be, by assumption, low initially.



Likewise, the synapse values from A to D are assumed to be high. If, as a result of repeated application of the input from A, the synapse values from A to C rise and become high, then the neuron C is a good candidate for recruitment into the cell-assembly \mathcal{A} . Whether it is recruited or not depends, of course, upon its relationship to other neurons of the system. It may merely continue to operate in parallel to the assembly \mathcal{A} . In fact C could become part of a system of neurons which would tend to suppress, via inhibitory con-

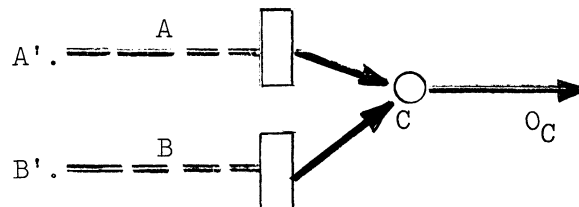
nections, an antagonistic assembly B . In any case, therefore, the question of when C would "correlate" with A in its firing arises. Here "correlate" means that the synapse values of A to C are high and that C tends to follow the same firing pattern as do the neurons of A . Therefore, whether C correlates with A or not depends critically upon the relationship between the firing patterns of A and B .

3.3 NETWORK CONFIGURATIONS FOR THE FIRST STAGE

In this section, first a general overview of the type of experiments which are to be carried out in this chapter is given, then second, a specification of the networks, consonant with the abstract development of the network equations, is carried out.

3.3.1 Overview

The general configuration of neurons that is to serve as the basis for the first part of this study is the following:

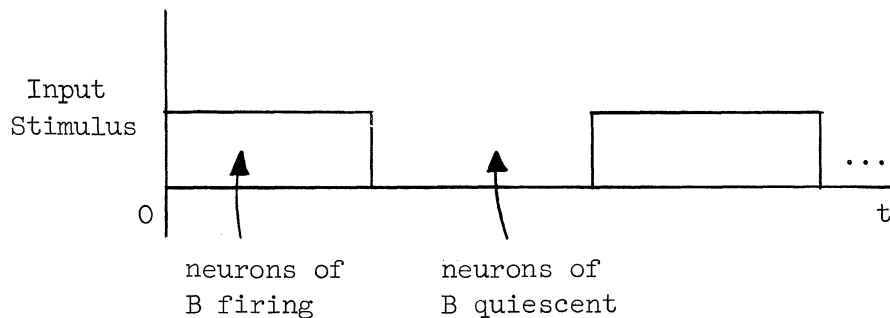


A and B are sets of neuron, C is a single neuron. Each neuron of A and B sends a connection to C . There are no other connections between neurons of A and B and C —i.e., no cycles. The neurons of A and B are assumed to be driven from stimulus sources A' and B' . From the patterns on the input lines A to C and B to C and the initial states of C , the output pattern O_C may be determined.

The sizes of the sets A and B , the particular patterns which they supply

to C, the initial states, the net parameters—all these are to be specified by the particular experiment at hand. Thus, A and B may consist of a single neuron each or A may have N neurons and B have none, etc. One can readily see then how it is possible to study the behavior of C as a function of a wide range of possible inputs and at the same time study the response of C "in isolation," as it were, given different settings of the basic net parameters.

A model situation with which we will be concerned in this chapter is that in which group A essentially provides "back-ground noise" to C, while group B provides patterned inputs of various sorts. One example of this is that where the neurons of B fire within a periodic envelope as follows:



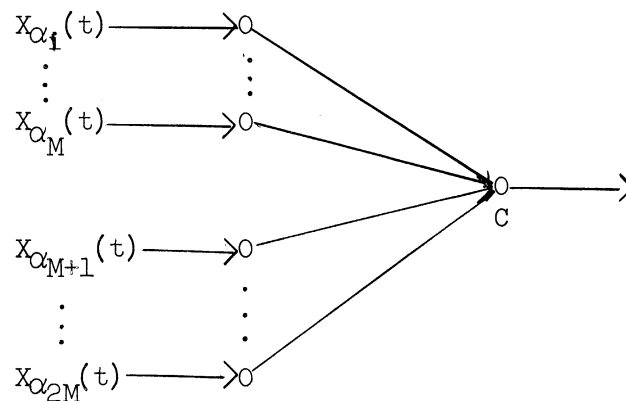
Questions such as what are the lengths of the "on" and the "off" periods in relation to neuron parameters, what are suitable firing rates of the neurons of B in the "on" and the "off" periods, etc., immediately arise and become of the greatest importance. The next step would be to have both A and B providing similar patterns such as this but out of phase, then to ask how C depends upon the phase difference, etc.

3.3.2 Specification of the Networks for the First Stage

The models of interest consist of $N = 2M + 1$ neurons (where N is the size of the network). The N neurons are partitioned into two groups of M neurons each and one group of one neuron. The former two groups will be

designated by A and B respectively, the single neuron by C. Each neuron of A and B respectively sends exactly one directed connection to neuron C. C, therefore, has $2M$ inputs. The output of C goes to the environment. The environment provides the neurons of A and B with inputs of the following type: Letting $\alpha_1, \dots, \alpha_M$ be the neurons of A and $\alpha_{M+1}, \dots, \alpha_{2M}$ be those of B, then to each α_i is associated a probabilistic stimulus $X_{\alpha_i}(t)$. At time t , independently of $X_{\alpha_i}(t+k)$ for all $k = \pm 1, \pm 2, \dots$, and with probability f_{α_i} , $X_{\alpha_i}(t) = 1$; with probability $1 - f_{\alpha_i}$, $X_{\alpha_i}(t) = 0$. If $X_{\alpha_i}(t) = 0$, neuron α_i is not effected. If $X_{\alpha_i}(t) = 1$, α_i is provided with an input stimulus ($I_{\alpha_i}(t)$ in the network equations $T_{\alpha_i}(t)$) which is always greater than $R_{\alpha_i}(t) \cdot F_{\alpha_i}(t)$ unless, of course, α_i is absolutely refractory (i.e., if $\delta_{\alpha_i}(t-1) = 1 \vee \delta_{\alpha_i}(t-2) = 1$). α_i has no other inputs. Notice that the probability f_{α_i} approximates the actual firing rate of α_i , that is, $f_{\alpha_i} \cdot T$ is the expected number of firings of α_i over a time interval of length T . Specification of the probabilistic vector $X_{\alpha_i}(t)$, $i = 1, \dots, 2M$, then determines the "vector" of frequencies f_{α_i} of the neurons α_i which comprise the total input set to neuron C. In each particular experiment, the vector $X_{\alpha_i}(t)$ will be specified in complete detail.

The connection-scheme, complete with the input vector $X_{\alpha_i}(t)$, has the following form:



The distinction between A and B is only for the purpose of allowing two sub-vectors of $X_{\alpha_i}(t)$, $i = 1, \dots, 2M$ to be applied, i.e., $X_{\alpha_1}(t), \dots, X_{\alpha_M}(t)$ and $X_{\alpha_{M+1}}(t), \dots, X_{\alpha_{2M}}(t)$. (Note: This network is obtained by specifying the $m_{\alpha_i C}$'s, $i = 1, \dots, 2M$, to be 1's and all others to be zero out of the set of $N^2 + N$ possible interconnections within the given set of N neurons.)

3.3.3 Network Functions, Initial Conditions, Environment

The threshold, fatigue, and synapse-value functions together with the parameters associated with them, such as Δ_1, Δ_2, U and D , etc., will be specified separately in each of the following experiments.

The initial conditions comprise specification of the following values:

1. $\lambda_{\alpha_i C}(0)$, $i = 1, \dots, 2M$
2. $r_{\alpha_i}(0)$, $i = 1, \dots, 2M$ and $r_C(0)$
3. $l_{\alpha_i}(0)$, $i = 1, \dots, 2M$ and $l_C(0)$
4. $I_{\alpha_i}(0)$, $i = 1, \dots, 2M$

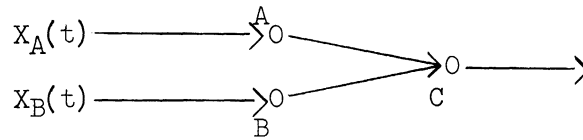
The $I_{\alpha_i}(0)$'s are assumed to be all equal and constant over all time, and so large that except when the α_i are absolutely refractory, they always cause α_i to fire when $X_{\alpha_i} = 1$. Thus, the initial values $r_{\alpha_i}(0)$ and $l_{\alpha_i}(0)$ are not so important. Yet the initial values of r_C and l_C clearly are important for, for example, if $l_C(0)$ is at the minimum, then neuron C starts out fully fatigued and may fail to respond to initial inputs for some period; whereas if it is fully rested, that is $l_C(0)$ is near the maximum, then C will most likely respond to the initial inputs.

The function of the environment in these experiments is, at each time step, to operate the probabilistic vector $X_{\alpha_i}(t)$, $i = 1, \dots, 2M$ and to observe the output of neuron C.

3.4 THREE-NEURON EXPERIMENTS

In the first series of experiments the schema of 3.3.2 was specialized

to $N = 3$ and $M = 1$, that is, three neurons, two of which, A and B, send one connection each to the third, C. The probabilistic vector $X_{\alpha_i}(t)$ reduces to $(X_A(t), X_B(t))$ and the corresponding probabilities f_{α_i} become f_A and f_B :



The general hypothesis for this series, stated formally, is the following:
HL. Given the three-neuron configuration, then for some appropriate selections for the network functions V , ϕ , and S and appropriate initial conditions, neuron C will tend to correlate with neuron B in the sense that as t becomes sufficiently large,

$$\begin{aligned} \lambda_{BC}(t) &\gg \lambda_{AC}(t) \quad \text{and} \\ \delta_B(t) = 1 &\implies \delta_C(t+1) = 1 \\ \delta_B(t) = 0 &\implies \delta_C(t+k) = 0 \end{aligned}$$

for some range of the rates f_A and f_B with $f_B < f_A$. (For " \implies " read "is followed by".)

This hypothesis merely says that neuron B eventually gains control over neuron C and that neuron A loses control over C. The motivation for this hypothesis is that the slow input, neuron B, may be regarded as the information carrying line, while the fast one might be regarded as a noisy line. One then would expect the neuron to correlate with the information-bearing line and not with the noisy one, at least under suitable conditions. For example, over the rapid staccato of a pneumatic hammer operating out-of-doors one might well hear a periodic knocking on the door.

3.4.1 Experiment 1

In this experiment, neurons A and B were presented with a constant stim-

ulus $I_A = I_B$ with probabilities f_A and f_B respectively, where according to hypothesis, $f_A > f_B$. The experiment was run for a variety of settings of these probabilities. The threshold and synapse-value functions, $V(r)$ and $S(\lambda)$, were chosen as indicated in Figures 1 and 2 while the fatigue function was taken to be identically 1, $\phi(l) = 1$. Notice that the range of the threshold curve is from $V_m = 100$ to $V_q = 1$ and that its form is exponential from $r = 3$ to $r = 11$ and is constant, equal to one, from $r = 11$ to $r = 16$; likewise the range of $S(\lambda)$ is from -15 for $\lambda = \lambda_{\min} = 0$ to $+15$ for $\lambda = \lambda_{\max} = 15$ and $S(\lambda)$ is a linear function with slope 2. The synapse-level probabilities were set to 0.1, that is $U(\lambda) = D(\lambda) = 0.1$ for all λ except, of course, $U(15) = 0 = D(0)$. The input stimuli $I_A(t)$ and $I_B(t)$ were both set to the constant value of 100; this is always sufficient to cause A or B to fire whenever $X_A(t)$ or $X_B(t) = 1$, unless A or B is absolutely refractory. The initial conditions for each run were $\lambda_{AC}(0) = 15 = \lambda_{BC}(0)$ and $r_A(0) = r_B(0) = r_C(0) = 16$.

The results of fifteen separate runs over a range of probabilities f_A and f_B are shown in Table 1. In this table, the following quantities are given for each run: (1) the values of f_A and f_B , (2) the length of the run (total number of time steps used in the run), (3) the terminal values of λ_{AC} and λ_{BC} , and (4) the number of times neuron i fired, N_i , for $i = A, B, C$. Detailed histories of the synapse-level changes for two typical runs are shown in Figure 3.

3.4.2 Analysis and Comment on Experiment 1

In this experiment, the choices for the threshold and synapse-value curves were made more or less arbitrarily. That is, the basis for these choices was not formal, rather was intuitive. Thus, the threshold curve was taken to be exponential over the range $r = 3$ to $r = 11$ and flat otherwise. The choice of the form of the curve was based on physiological grounds as

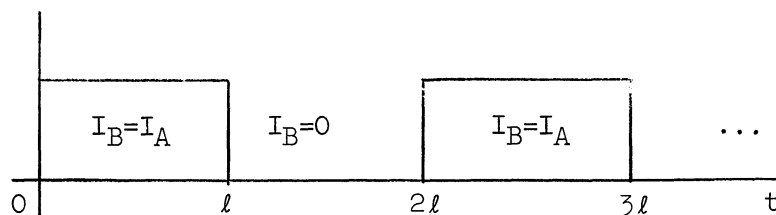
mentioned earlier, the choice for the particular values of the curve was based on the consideration that the main range of operation of neuron C, as driven by A or B, lie in the range $r = 6$ to $r = 10$ (approximately)—that is, in the mid-range of recovery-values. The flat portion was to allow C to be driven with minimum stimulus and hopefully encourage development of λ_{BC} . Likewise, the choice of $S(\lambda)$ was ad hoc, using a curve balanced between positive and negative values "for a starter." In order to accentuate the effects of the threshold and synapse-value curves, however, the fatigue function was set to the identity. The values of U and D were chosen to attenuate the growth or decay of λ_{AC} and λ_{BC} and again were starting values.

A glance at the terminal values of λ_{AC} and λ_{BC} will suffice to show that the results of this experiment are inconclusive. Sometimes C correlates with B (runs 1, 2, 9, 11, 14), sometimes with A (runs 3-8), sometimes with neither (runs 12 and 15) with no apparent reason. Moreover, it is not clear that in this case there should be a preference for C to correlate with A or B, since neither of the inputs is structured in any way—thus C is being asked to discriminate between two completely random input sequences which differ only in their relative frequencies. Therefore, this experiment was abandoned for the case in which neuron A continues to present C with a random input sequence, but B now presents C with a periodic input, thus an input signal with structure to which C should respond selectively.

For reasons to be mentioned shortly, Experiment 1 would not be expected to be successful in any case. The motivation for including it here is mainly historical as well as to illustrate some of the specific problems that arise in implementing models of the kind considered in this paper. One interesting phenomenon should be noted, however. That is, even with the retarding probability $D = 0.1$, the synapse-levels drop rather rapidly. This suggests the need for a positive bias in the $U(\lambda)$'s and $D(\lambda)$'s.

3.4.3 Experiment 2

In this experiment, neuron A was presented with a constant stimulus I_A with probability f_A whereas neuron B was presented with a periodically interrupted stimulus I_B which equals I_A on a set of intervals $t = 2kl$ to $t = (2k + 1)l$ for $k = 0, 1, 2, \dots$ where l is the length of the interval, and equals 0 on the complementary intervals. In the intervals in which $I_B = I_A$, I_B is again presented with probability $f_B = f_A$. The intervals in which $I_B = I_A$ are called the "on-periods" for B, those in which $I_B = 0$ are called the "off-periods."



As in Experiment 1, the threshold curve was taken as in Figure 1, with slight variation in one case, while again $\phi(l) \equiv 1$. However, variation was introduced in the choice of the synapse-value curve and the probabilities $U(\lambda)$ and $D(\lambda)$. A variety of runs were performed for various settings of the functions $V(r)$ and $S(\lambda)$ and values of f_A and l . The runs performed are discussed separately below while their results are presented in Table 2. In this table, the following quantities are displayed for each run: (1) the values of l and f_A , (2) the length of the run, (3) the terminal values of λ_{AC} and λ_{BC} , and (4) the number of times neuron i fired, $i = A, B, C$. Detailed histories of the synapse-level changes for several typical runs are shown in Figure 5.

Run 1. The functions $V(r)$ and $S(\lambda)$ were taken as in Experiment 1 (Figures 1 and 2). $I_A = 140$ ($= I_B$ in the on-periods). The initial conditions were $\lambda_{AC}(0) = \lambda_{BC}(0) = 8$, $r_A(0) = r_B(0) = r_C(0) = 16$.

Runs 1b-1f involve the same network functions and parameters as Run 1 and the same period, but the probability f_A is varied.

Runs 2a, 2b. Functions $V(r)$ and $S(\lambda)$ as in Figures 1 and 2, except incremented by 2 throughout; e.g., if $V_0(r)$ is the function of Figure 1, then the threshold function used in Run 2 is $V_0(r) + 2$, etc. $I_A = 180$ ($= I_B$ in the on-periods). $U(\lambda)$ and $D(\lambda)$ as in Table 3. Initial conditions as in Run 1. The run was performed twice: that is, done once, then with initial conditions restored, repeated. Since the random-number generating procedure used to determine the vectors $(X_A(t), X_B(t))$ was not re-initialized, the results of repeated runs like this need not be identical.

Runs 3a, 3b through Runs 6a, 6b. Exactly as Runs 2a, 2b except that the functions $V(r)$ and $S(\lambda)$ are incremented by two in going from Run 2 to Run 3, again from Run 3 to Run 4, etc. Thus, the threshold function for Run 6 is $V_0(r) + 10$ where $V_0(r)$ is the $V(r)$ in Figure 1.

Runs 7a, 7b through 11a, 11b. For these runs $V(r)$ was taken as 100 times the function of Figure 1, that $V(r) = 100 V_0(r)$. I_A was taken as 10,000. Again, $U(\lambda)$ and $D(\lambda)$ are those of Table 3. For Run 7, the synapse-value curve was taken as $S(\lambda) = 1500 + 8/9 \cdot (S'_0(\lambda) - 1500)$ where $S'_0(\lambda) = 100 S_0(\lambda)$, that is 100 times the curve of Figure 2. That for Run 8 was taken as $S(\lambda) = 1500 + 8/10 (S'_0(\lambda) - 1500)$, ... , for Run 11 as $S(\lambda) = 1500 + 8/13 (S'_0(\lambda) - 1500)$. $S'_0(\lambda) = 100 S_0(\lambda)$, $S_0(\lambda)$ is the curve of Figure 2. These curves are given in Figure 4.

3.4.4 Analysis and Comments on Experiment 2

As the results of Runs 1a, 1b, ... , 1e, 1f show, given the network functions of Experiment 1, no clear pattern of success occurs; e.g., Run 1a is bad, both λ_{AC} and λ_{BC} are low, yet Run 1c is good, $\lambda_{BC} \gg \lambda_{AC}$. It was sus-

pected that a positive bias was necessary in the synapse-value function and in the probabilities $U(\lambda)$ and $D(\lambda)$; this prompted the values for U and D given in Table 3 and the schemes for biasing $S(\lambda)$ as used in Runs 2-5 and 7-11 (see in particular, Figure 4). Again, the selections of these particular values were largely ad hoc. As no clear picture of success emerged from this procedure, it became clear that the experimental hypothesis could not possibly hold in these models for a three-neuron network: For B to gain control over C , with both S_{BC} and S_{AC} set initially equal to moderate values, B must fire initially with some regularity in unison with A in order to cause C to fire (at the same rate as B); however, this is an unlikely event since the probability of joint firing of A and B is $f_A \cdot f_B$, which in Run 1 would be $1/16$ in B 's on period, 0 in the off-period.

This undesirable situation is remedied by replacing the single neurons A and B by groups of neurons A and B so that the probability of firing for any neuron of B is much greater than $f_A \cdot f_B$, although f_A and f_B are still the rates of the individual neurons of A and B

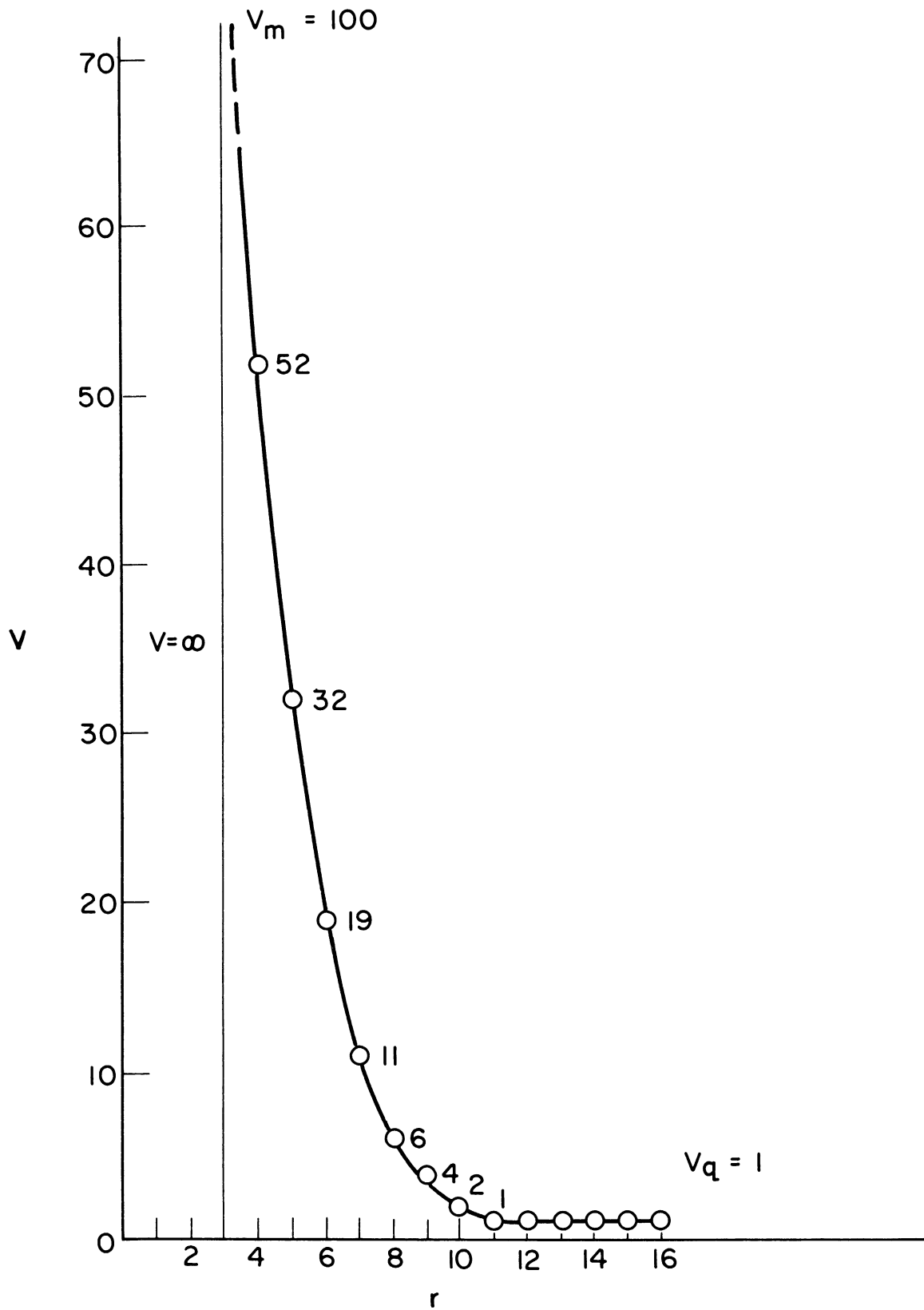


Figure 1. Threshold Curve for First Series. $V_m = 100$, $V_q = 1$. Form of curve is exponential from $r = 3$ to $r = 11$, linear and constant for $r = 11$ to $r = 16$.

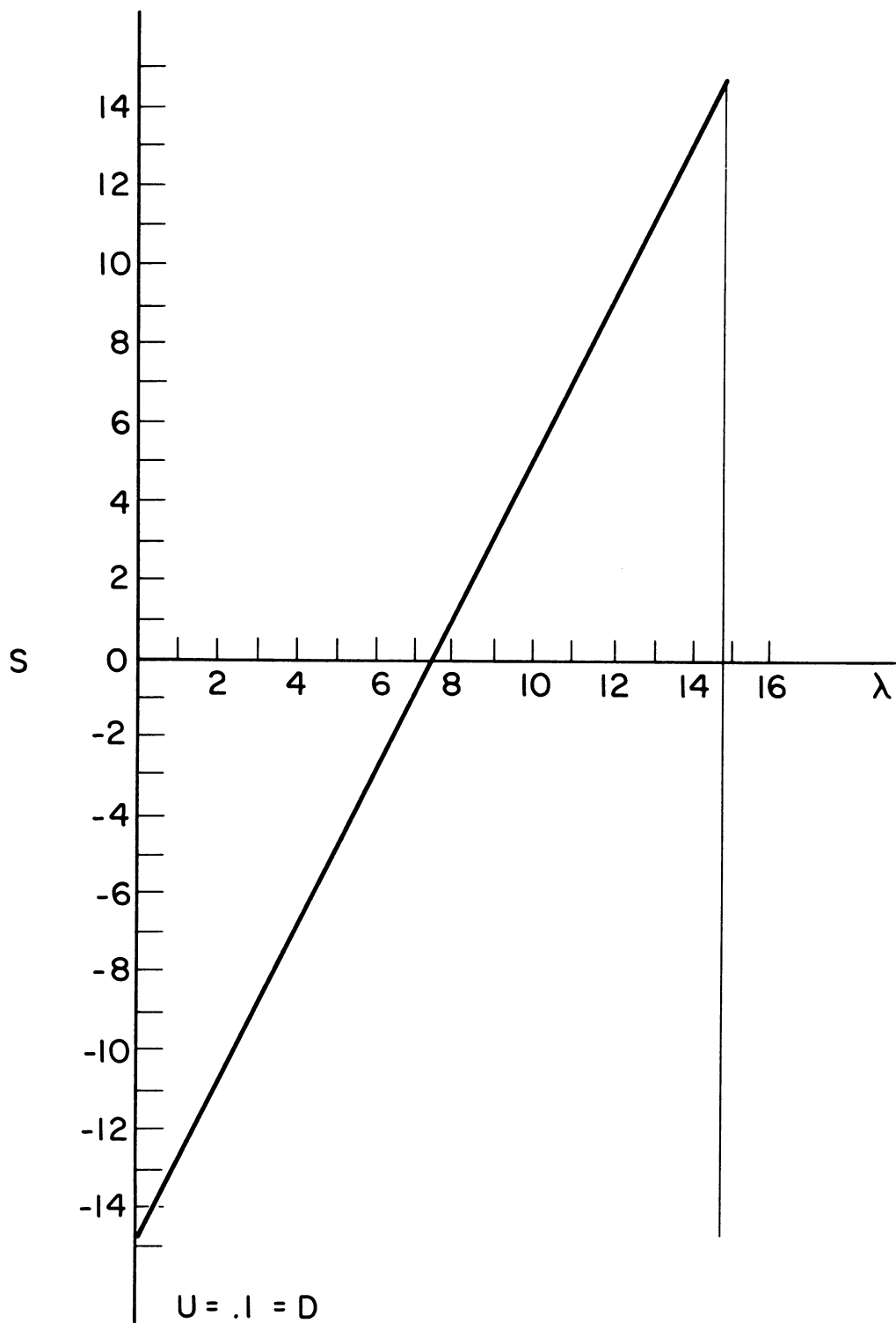


Figure 2. Synapse Value Curve for First Series. $\lambda_{\min} = 0$, $\lambda_{\max} = 15$, $S_{\min} = -15$, $S_{\max} = +15$. Form of curve: linear with slope 2.

TABLE 1
RESULTS OF 15 RUNS FOR EXPERIMENT 1

Run No.	f_A	f_B	Length of Run (timesteps)	λ_{AC} (final)	λ_{BC} (final)	N_A	N_B	N_C
1	1/4	1/6	10,000	0	15	1658	1256	698
2	1/6	1/9	10,000	0	15	1208	891	588
3	1/10	1/15	10,000	15	0	790	559	591
4	1/12	1/18	10,000	15	0	698	474	539
5	1/16	1/24	10,000	15	0	543	365	447
6	1/7	1/9	10,000	9	0	1084	891	707
7	1/8	1/10	10,000	15	0	977	809	682
8	1/9	1/11	10,000	15	0	891	725	623
9	1/10	1/12	10,000	0	15	790	666	562
10	1/11	1/13	10,000	10	15	743	631	745
11	1/6	1/9	5,000	0	15	635	444	296
12	1/7	1/10	5,000	0	0	563	452	14
13	1/8	1/11	5,000	0	14	513	405	255
14	1/9	1/12	5,000	0	11	437	351	257
15	1/10	1/13	5,000	12	14	420	321	400

Run No. 5			Run No. 12		
Time Step No.	λ_{AC}	λ_{BC}	Time Step No.	λ_{AC}	λ_{BC}
1	9	9	1	9	9
99	10	9	9	9	8
331	11	9	30	8	8
671	11	8	35	8	7
678	11	7	44	8	6
694	11	6	224	7	6
882	12	6	227	6	6
914	11	6	237	6	5
968	12	6	248	5	5
1020	13	6	253	5	4
1181	13	5	270	4	4
1196	14	5	283	3	4
1381	15	5	326	3	3
1454	15	4	327	2	3
1493	15	3	390	1	3
1506	15	2	398	0	3
1589	15	1	403	0	2
1645	14	1	634	0	1
1749	14	0	899	0	0
1805	15	0			
2538	14	0			
3256	15	0			
5959	14	0			
5989	15	0			
7814	14	0			
8583	15	0			

Figure 3. Histories of Synapse-Level Change for Runs 5 and 12 of Experiment 1. The synapse-levels are shown only when one of the two changes; thus for $t = 1$ through $t = 98$, $\lambda_{AC} = 9$, $\lambda_{BC} = 9$, then at $t = 99$, λ_{AC} becomes 10, λ_{BC} remains 9, etc.

TABLE 2
RESULTS OF 11 RUNS FOR EXPERIMENT 2

Run No.	f_A	l	Length of Run	λ_{AC} (final)	λ_{BC} (final)	N_A	N_B	N_C
1a	1/4	40	2,000	0	0	324	178	2
2a	1/6	60	3,000	2	1	373	192	2
2b	1/6	60	3,000	1	2	377	189	16
3a	1/6	60	3,000	13	1	386	196	202
3b	1/6	60	3,000	14	1	401	206	196
4a	1/6	60	3,000	12	13	386	192	227
4b	1/6	60	3,000	13	0	400	77	204
5a	1/6	60	3,000	0	12	360	194	74
5b	1/6	60	3,000	12	2	380	180	163
6a	1/6	60	3,000	1	3	377	194	133
6b	1/6	60	3,000	1	2	345	200	9
7a	1/6	60	3,000	2	1	373	192	2
7b	1/6	60	3,000	1	2	377	189	19
7c	1/6	60	3,000	13	2	386	196	242
8a	1/6	60	3,000	13	3	378	189	240
8b	1/6	60	3,000	14	2	396	187	251
8c	1/6	60	3,000	11	15	388	186	263
9a	1/6	60	3,000	13	4	352	178	224
9b	1/6	60	3,000	2	15	351	172	141
9c	1/6	60	3,000	2	10	357	181	107
10a	1/6	60	3,000	15	8	397	182	263
10b	1/6	60	3,000	2	15	358	192	172
10c	1/6	60	3,000	15	2	394	190	248
11a	1/6	60	3,000	14	13	342	185	264
11b	1/6	60	3,000	11	11	351	184	258
11c	1/6	60	3,000	12	5	324	192	240
1b	1/4	40	2,000	2	1	324	178	2
1c	1/5	40	2,000	2	11	282	145	84
1d	1/6	40	2,000	2	15	247	129	86
1e	1/7	40	2,000	1	2	240	112	40
1f	1/8	40	2,000	4	13	207	102	80

TABLE 3

λ	$D(\lambda)$	$U(\lambda)$
0	0	.2
1	.005	.2
2	.01	.2
3	.02	.2
4	.04	.2
5	.06	.2
6	.08	.2
7	.1	.2
8	.1	.2
9	.1	.18
10	.1	.16
11	.1	.14
12	.1	.12
13	.1	.1
14	.1	.08
15	.1	0

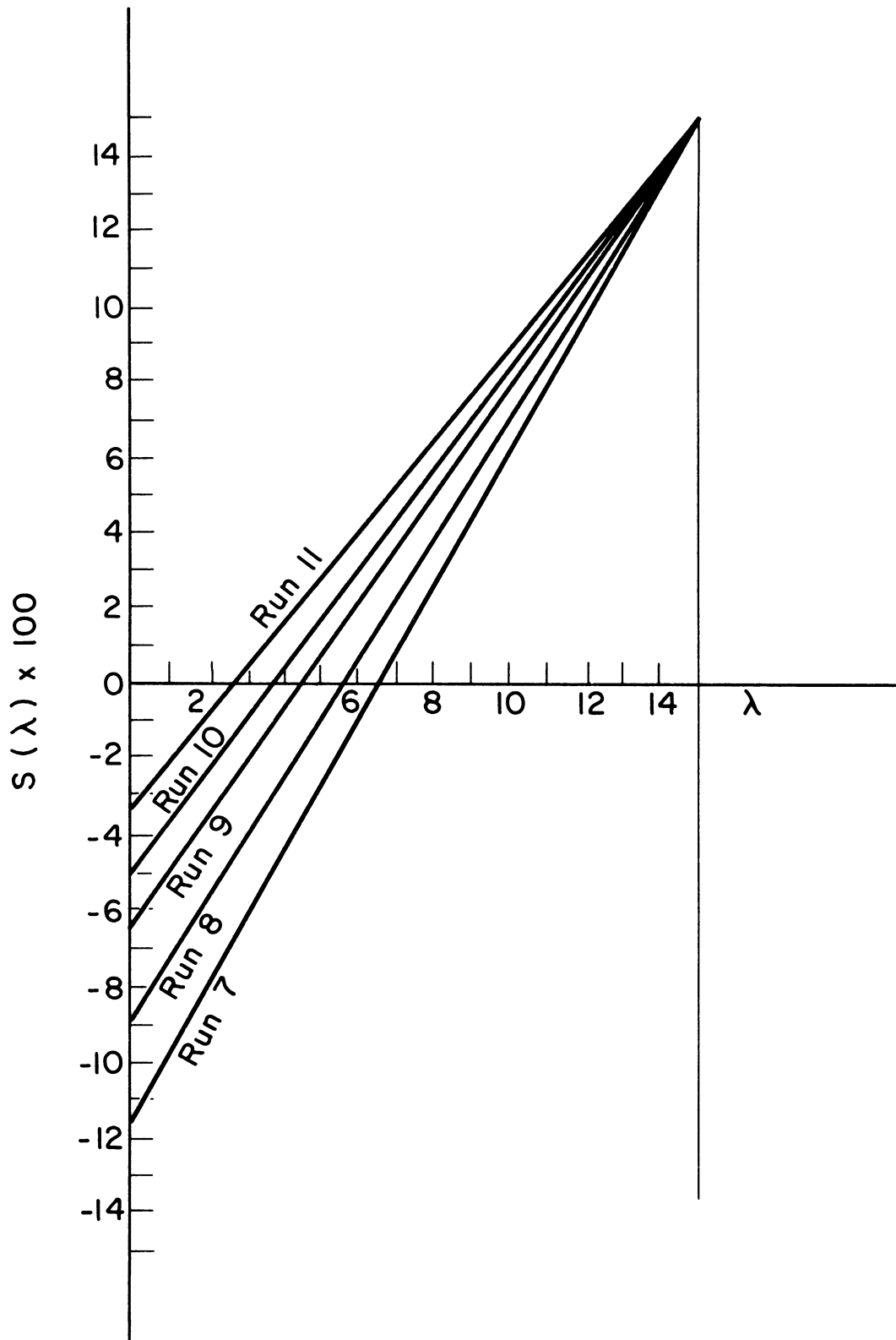


Figure 4. Synapse-Level Curves for Runs 7-11 for Experiment 2.

Run 7a			Run 5a			Run 5b		
t*	λ_{AC}	λ_{BC}	t*	λ_{AC}	λ_{BC}	t*	λ_{AC}	λ_{BC}
1	8	8	1	8	8	1	8	8
12	8	9	4	8	9	44	9	8
15	8	8	63	7	9	104	10	8
19	7	8	169	6	9	114	11	8
21	7	7	237	5	9	141	11	9
28	6	7	256	4	9	228	12	9
255	6	6	266	3	9	264	12	10
335	5	6	280	3	10	337	11	10
383	5	5	377	3	11	379	11	9
541	4	5	403	3	10	385	11	10
837	3	5	514	3	11	440	10	10
852	3	4	627	3	12	579	9	10
979	3	3	645	3	11	587	10	10
1452	3	2	653	2	11	598	11	10
1667	2	2	720	1	11	606	11	9
2079	1	1	749	1	10	744	11	8
			767	1	11	755	11	7
			894	1	12	774	12	7
			1099	1	11	804	11	7
			1233	1	13	813	10	7
			1450	1	12	827	11	7
			1490	0	12	839	12	7
			1565	0	11	881	11	7
			1572	0	12	934	12	7
			1603	0	11	962	12	6
			1608	0	12	1202	11	6
			1620	0	11	1307	12	6
			1730	0	10	1312	11	6
			1825	0	9	1333	11	5
			1842	0	8	1343	11	4
			1962	0	9	1443	11	3
			1981	0	10	1444	10	3
			2067	0	11	1588	10	2
			2207	0	12	1655	11	2
			2292	0	11	1675	12	2
			2303	0	12	1686	11	2
			2530	0	11	1691	12	2
			2569	0	12	1753	11	2
			2783	0	11	1858	12	2
			2788	0	12	1877	11	2
						1898	12	2
						1927	11	2
						1952	12	2
						2029	11	2
						2076	12	2
						2135	11	2
						2212	12	2
						2282	11	2
						2360	12	2
						2443	11	2
						2583	12	2
						2680	11	2
						2832	12	2
						2909	11	2
						2925	12	2
						2960	11	2
						2964	12	2

*t indicates the time step at which the change took place.

Figure 5. Histories of Synapse-Level Changes for Runs 7a, 5a, and 5b.

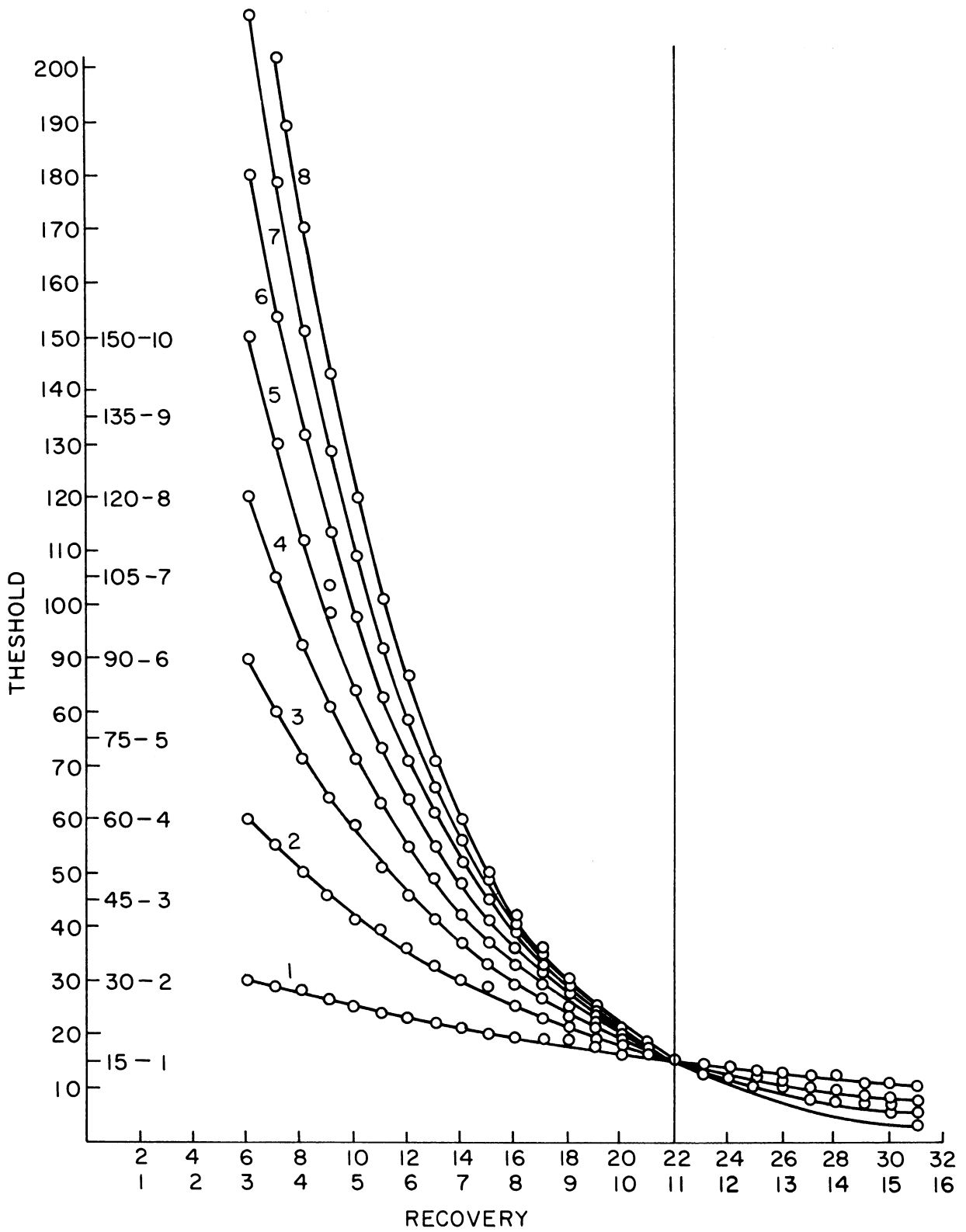


Figure 6. Threshold Curves for Experiment 3, Runs 1-8.

	Run No.							
	1	2	3	4	5	6	7	8
λ_{A_1C}	2	0	0	0	0	0	0	0
λ_{A_2C}	0	1	0	1	0	0	1	0
λ_{A_3C}	1	0	2	1	6	0	0	1
λ_{A_4C}	1	12	0	0	0	0	1	0
λ_{B_1C}	15	15	15	15	15	14	15	15
λ_{B_2C}	15	15	15	15	15	15	14	15
λ_{B_3C}	15	15	15	14	14	15	15	14
λ_{B_4C}	15	15	15	14	15	15	15	15
N_{A_1}	335	335	360	331	332	333	330	344
N_{A_2}	310	325	332	335	333	336	320	347
N_{A_3}	344	345	329	322	339	344	333	321
N_{A_4}	337	327	322	332	334	315	347	319
N_{B_i}	344	326	330	319	332	339	330	317
N_C	360	338	242	250	234	228	226	220

Figure 7. Terminal Synapse-Levels, etc. for Runs 1-8 of Experiment 3. Run j used threshold curve j of Figure 5.

	Run No.							
	1	2	3	4	5	6	7	8
λ_{A_1C}	0	0	14	0	0	0	0	0
λ_{A_2C}	0	0	0	0	0	0	0	0
λ_{A_3C}	0	0	0	0	0	0	14	0
λ_{A_4C}	0	0	0	0	4	0	0	0
λ_{B_1C}	9	6	5	3	0	0	2	13
λ_{B_2C}	0	12	3	1	1	11	0	1
λ_{B_3C}	14	9	2	6	14	0	13	3
λ_{B_4C}	0	0	0	4	2	4	0	7
N_{A_1}	344	329	335	343	335	340	317	337
N_{A_2}	347	330	345	341	349	313	333	329
N_{A_3}	357	337	332	336	345	313	324	338
N_{A_4}	359	326	321	338	344	338	322	322
N_{B_1}	163	183	172	170	157	174	182	177
N_{B_2}	167	160	175	175	172	169	186	153
N_{B_3}	177	172	170	168	159	170	164	162
N_{B_4}	178	164	167	163	164	177	168	172
N_C	123	166	170	120	158	114	168	109

Figure 8. Terminal Synapse-Levels, etc. for Runs 1-8, Experiment 4.

TABLE 4

SYNAPSE-LEVEL PROBABILITIES FOR RUNS 9-16, EXPERIMENT 4

Run	$U(\lambda)$	$D(\lambda)$	Ratio D/U (Approx.)
9	.2	.1	1/2
10	.2324	.09221	1/3
11	.2701	.07406	1/4
12	.3138	.06373	1/5
13	.3647	.05485	1/7
14	.4237	.04720	1/9
15	.4924	.04062	1/12
16	.5721	.03476	1/16

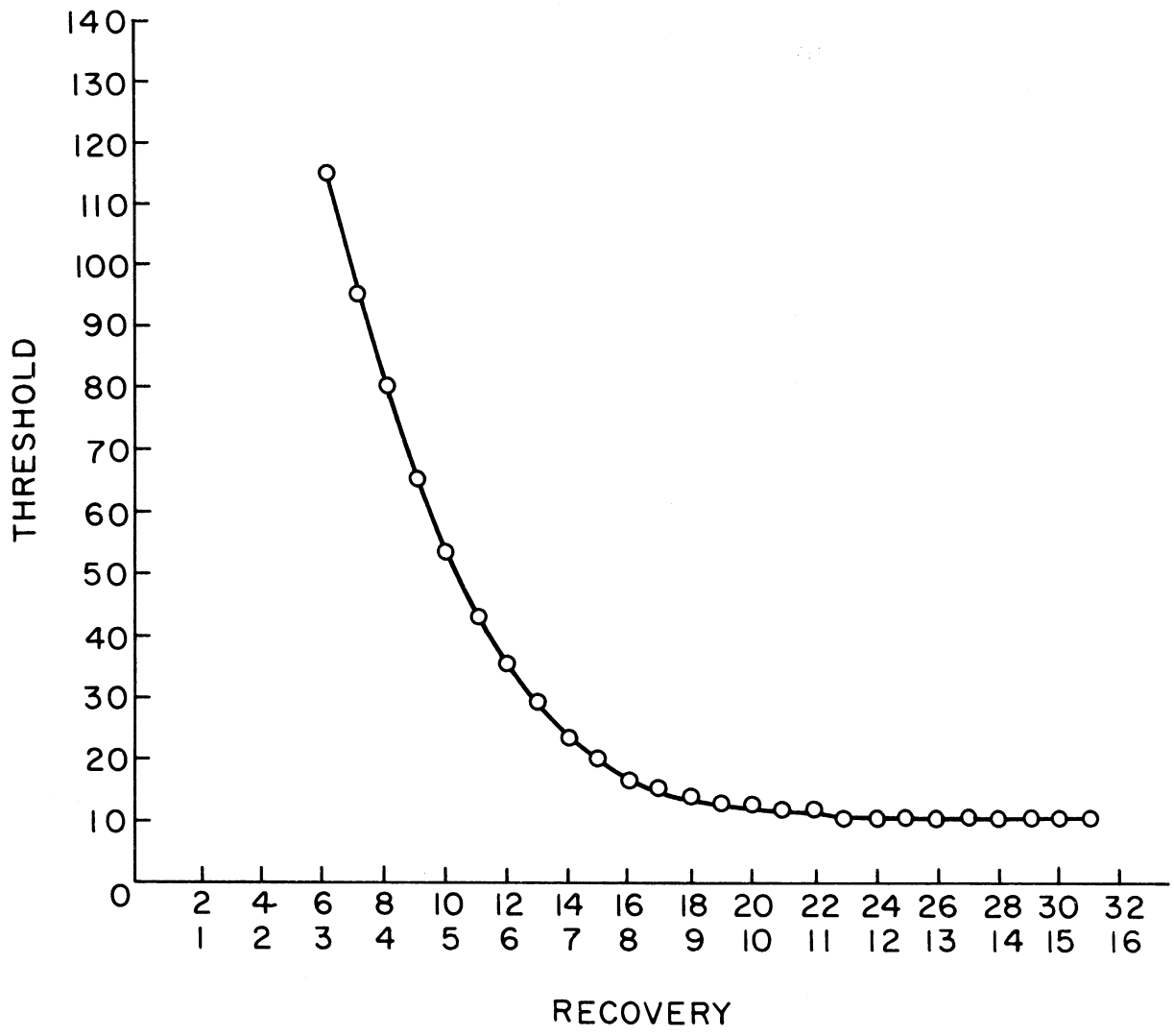


Figure 9. Threshold Curve for Runs 9-16, Experiment 1.

	Run No.							
	9	10	11	12	13	14	15	16
λ_{A_1C}	0	0	0	0	15	15	14	15
λ_{A_2C}	0	0	0	0	9	14	15	15
λ_{A_3C}	0	0	0	0	15	14	14	15
λ_{A_4C}	3	0	0	0	15	15	15	10
λ_{B_1C}	0	0	5	15	9	15	15	13
λ_{B_2C}	0	0	0	12	13	14	15	15
λ_{B_3C}	0	0	0	6	15	15	15	11
λ_{B_4C}	0	0	15	8	13	15	11	15
N_{A_1}	344	329	335	343	335	340	317	337
N_{A_2}	347	330	345	341	349	313	333	329
N_{A_3}	357	337	332	336	345	313	324	338
N_{A_4}	359	326	321	338	344	338	322	322
N_{B_1}	163	183	172	170	157	174	182	177
N_{B_2}	167	160	175	175	172	169	186	153
N_{B_3}	177	172	170	168	159	170	164	162
N_{B_4}	178	164	167	163	164	177	168	172
N_C	125	9	107	91	272	264	288	263
Ratio D/U (Approx.)	1/2	1/3	1/4	1/5	1/7	1/9	1/12	1/16

Figure 10. Final Synapse-Levels, etc. for Runs 9-16, Experiment 4.

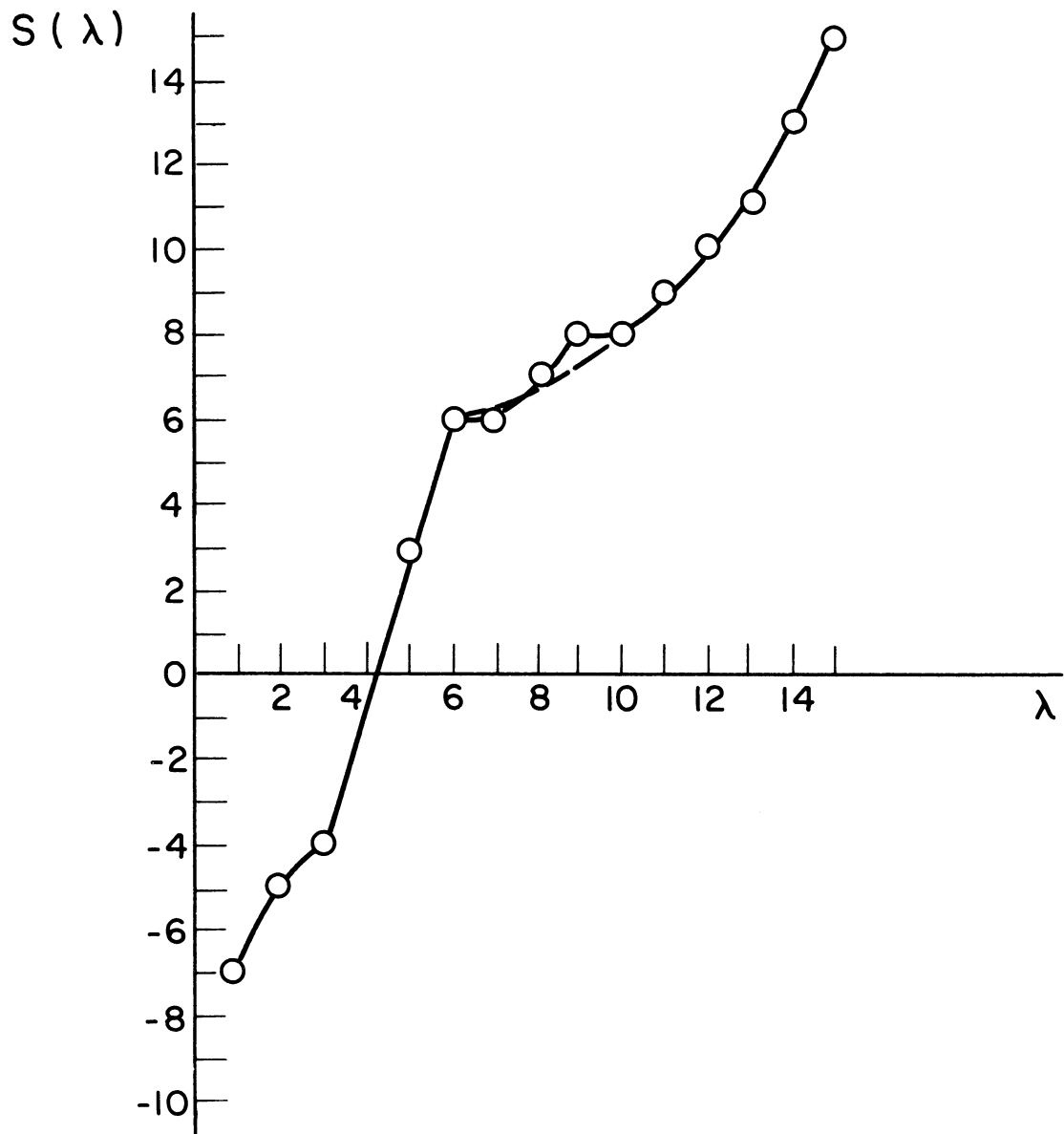


Figure 11. $S(\lambda)$ for Runs 17-32, Experiment 4.

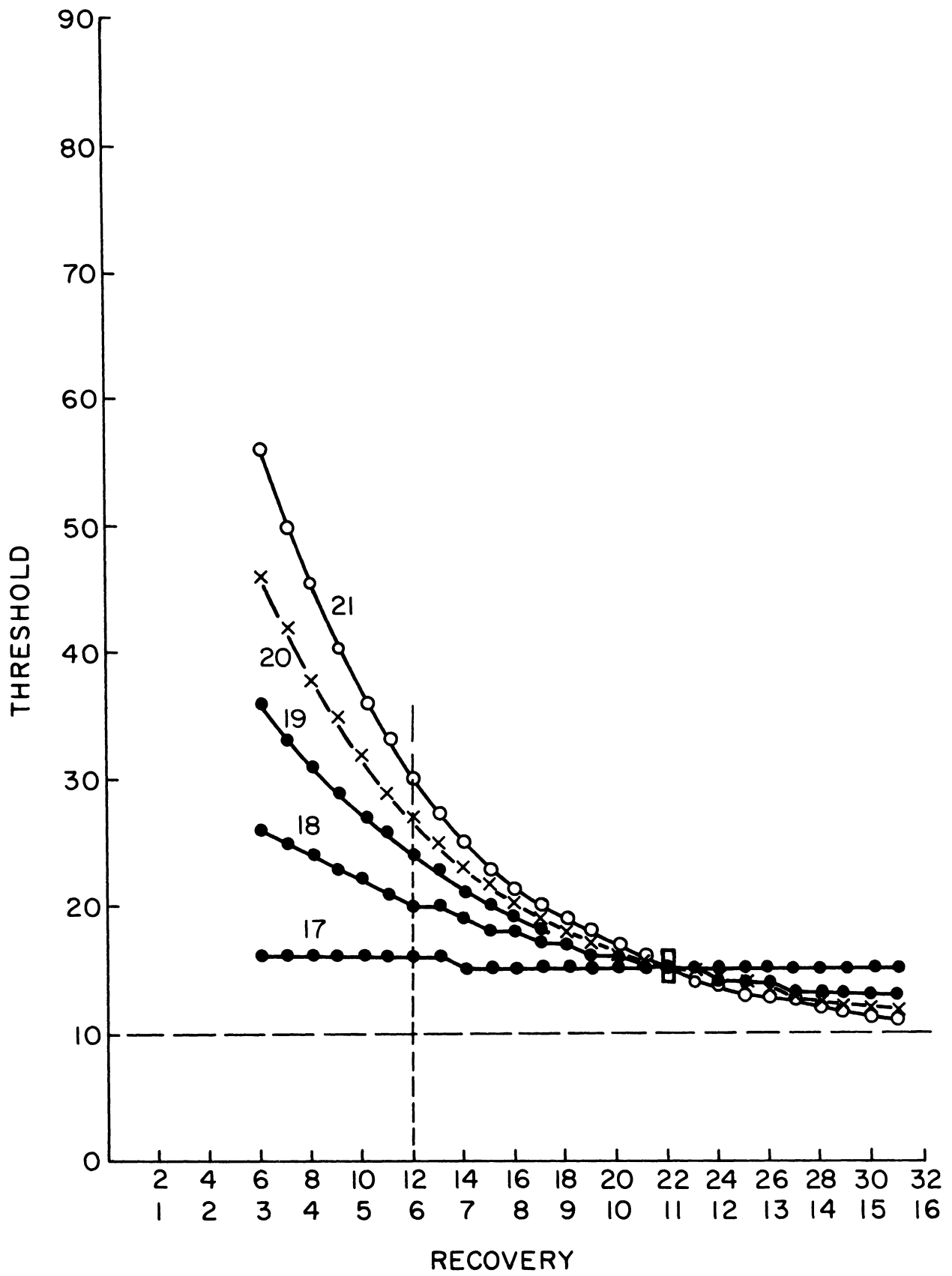


Figure 12. Threshold Curves for Runs 17-21 and 25-29, Experiment 4.

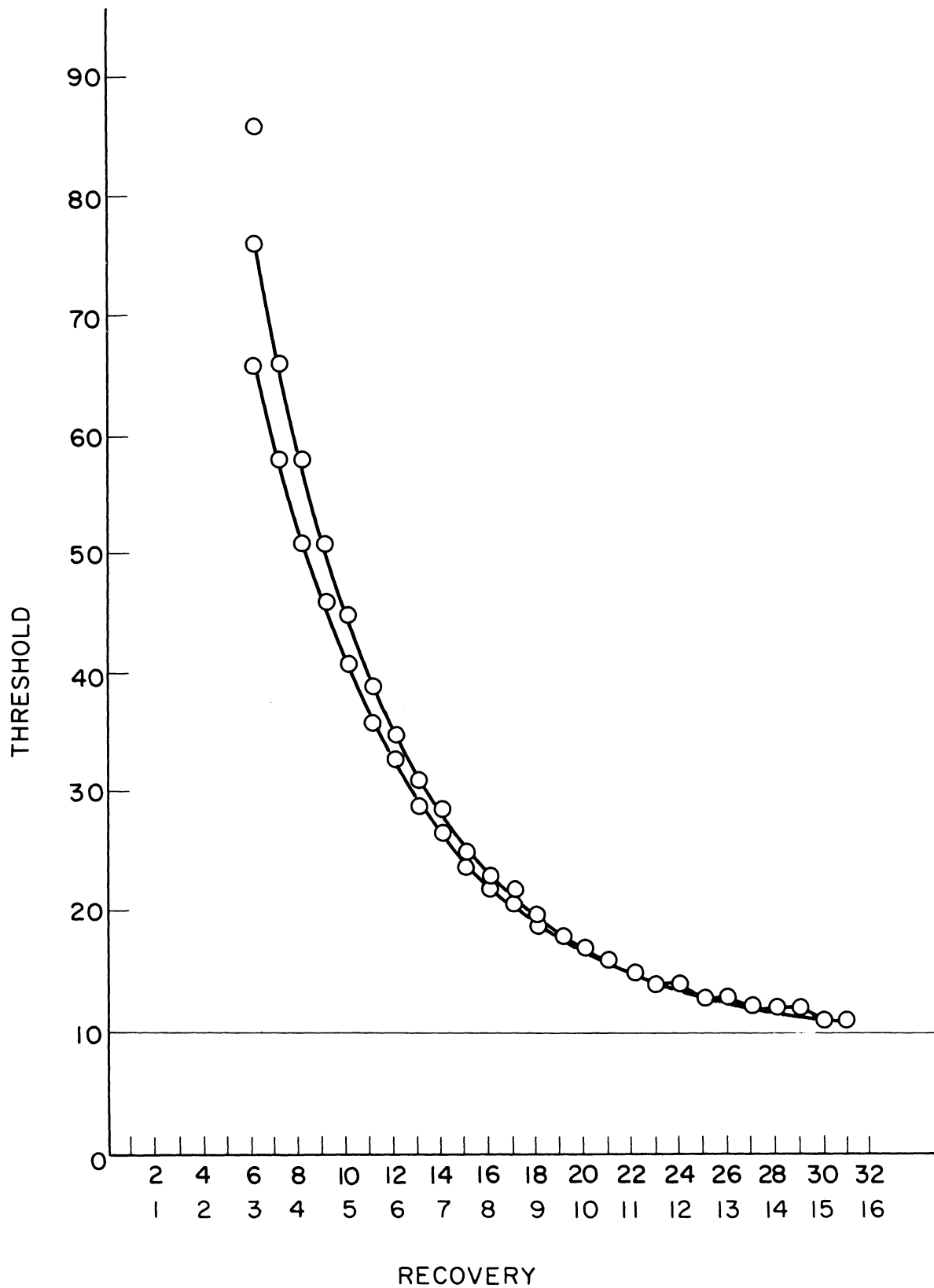


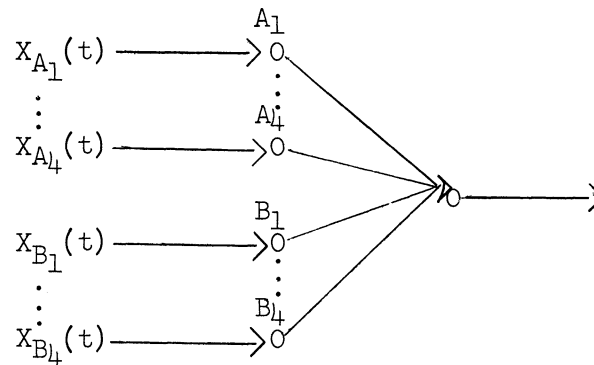
Figure 13. Threshold Curves for Runs 22-24 and 30-32, Experiment 4.

	Run No.							
	17,25	18,26	19,27	20,28	21,29	22,30	23,31	24,32
λ_{A1C}	14 15	15 13	15 15	15 0	13 13	15 13	2 6	0 3
λ_{A2C}	15 15	10 11	15 15	0 0	15 14	14 15	1 7	13 15
λ_{A3C}	14 14	15 15	13 13	3 0	4 9	10 0	15 13	11 6
λ_{A4C}	15 15	11 14	15 15	15 4	15 11	0 3	7 0	15 1
λ_{B1C}	15 15	15 15	10 10	11 15	15 15	15 5	13 10	13 12
λ_{B2C}	15 15	14 14	8 8	15 10	9 13	10 13	14 7	6 1
λ_{B3C}	15 15	14 14	15 15	14 14	15 15	15 9	14 15	11 15
λ_{B4C}	15 15	15 13	14 14	14 13	14 14	9 15	13 12	8 7
N_{A1}	344 344	329 329	335 335	343 343	335 335	340 340	317 317	337 337
N_{A2}	347 347	330 330	345 345	341 341	340 340	313 313	333 333	320 320
N_{A3}	357 357	337 337	332 332	336 336	345 345	313 313	324 324	338 338
N_{A4}	359 359	326 326	321 321	338 338	344 344	338 338	322 322	322 322
N_{B1}	163 163	183 183	172 172	170 170	157 157	174 174	182 182	177 177
N_{B2}	167 167	160 160	175 175	175 175	172 169	169 169	186 186	153 153
N_{B3}	177 177	172 172	170 170	168 168	159 159	170 170	164 164	162 162
N_{B4}	178 178	164 164	167 167	163 163	164 164	177 177	168 168	172 172
N_C	409 400	331 318	267 266	238 134	255 233	227 191	220 196	206 188

Figure 14. Results of Runs 17-32, Experiment 4. (The paired runs differed only on the initial conditions on the λ 's.)

3.5 NINE-NEURON EXPERIMENTS

Following the suggestion made at the end of the last section, the schema of 3.3.2 was specialized to $N = 9$ and $M = 4$, that is, a network consisting of nine neurons, two groups of which, A and B, have four neurons each, all of which send one connection to the single neuron C. The vector $X_{\alpha_i}(t)$ becomes $(X_{A_1}(t), \dots, X_{A_4}(t), X_{B_1}(t), \dots, X_{B_4}(t))$ where $A_i \in A$ and $B_i \in B$, $i = 1, \dots, 4$. The probabilities f_{α_i} become f_{A_i} and f_{B_i} , $i = 1, \dots, 4$:



The hypothesis for the three-neuron experiments H1 is modified as follows:

H2. For appropriate selections of the network functions V , ϕ , and S and appropriate initial conditions, neuron C will tend to correlate with group B in the sense that as t becomes sufficiently large,

$$\lambda_{BC}(t) \gg \lambda_{AC}(t)$$

and

$$\text{for all } i, \quad \delta_{B_i}(t) = 1 \implies \delta_C(t+1) = 1 \text{ a.a.}$$

and

$$\text{for all } i, \quad \delta_{B_i}(t) = 0 \implies \delta_C(t+1) = 0 \text{ a.a.}$$

for some range of the rates f_{A_i} and f_{B_i} with $f_{B_i} < f_{A_j}$ for $i, j = 1, \dots, 4$.

The notation

$$\lambda_{\overline{BC}}, \lambda_{\overline{AC}}$$

implies some sort of average (e.g., the mean) over the $\lambda_{B_i C}$'s and $\lambda_{A_i C}$'s respectively. This leaves room for a specific $\lambda_{B_i C}$ being less than some $\lambda_{A_j C}$. "a.a." (almost always) implies that the condition occurs with a high probability, but not with probability 1. Thus it leaves room for the events $\delta_{B_i}(t) = 1 \implies \delta_C(t+1) = 0$ and even—for some i — $\delta_{A_i}(t) = 1 \implies \delta_C(t+1) = 1$.

This hypothesis says that group B eventually takes over control of neuron C in the sense that the synapse-levels from B to C become high (in average-value), those from A to C become low, thus ensuring most of the time the firing of C at $t + 1$ when one or more of the neurons of B fires, while the neurons of A seldom cause C to fire.

3.5.1 Experiment 3, Synchronous Case

In this experiment, the vector $X_{\alpha_i}(t)$ was taken in the following fashion: $X_{A_j}(t) = 1$ with probability f_A independently of $X_{A_j}(t)$, $j \neq i$, and independently of $X_{A_j}(t+k)$, for all j , $k = 1, 2, \dots$, likewise, the X_{A_i} are treated independently of the X_{B_i} . Thus, the neurons of A fire randomly and independently at the rate f_A . (The values of $I_{A_i}(t)$ and $I_{B_i}(t)$ were set to the constant value of 1000.) For the neurons of B, $X_{B_i}(t) = 1$ for all $i = 1, \dots, 4$ with probability $f_B = f_A$. Thus, the neurons of B fire in synchrony.

This experiment does not quite conform to H2 since f_B is not less than f_A ; however, it does provide some insights, as will be seen.

The threshold curves for this experiment are shown in Figure 6; the synapse-value curve is that of Experiment 1: $\phi(\ell) = 1$, and $U(\lambda) = 0.2$ and $D(\lambda) = 0.1$ for all λ , except that $U(15) = 0 = D(0)$. f_A was set to $1/6$, the

length of all runs was 3000 time steps. $\lambda_{A_i C} = 15 = \lambda_{B_i C}$ and $I_{A_i} = I_{B_i} = 1000$ for $i = 1, \dots, 4$. $r_C(0) = r_{A_i}(0) = r_{B_i}(0) = 16$. The results of eight separate runs are shown in Figure 7, where the terminal values of λ 's are given together with the number of times each neuron fired. Run j uses threshold curve j of Figure 5, $j = 1, \dots, 8$.

3.5.2 Analysis and Comment on Experiment 3

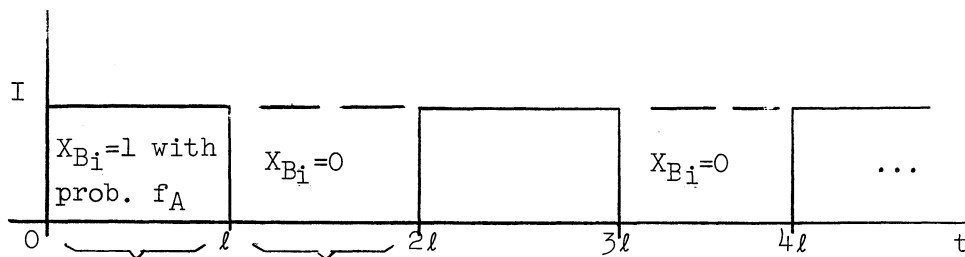
It is evident from Figure 7 that all the runs performed were successful, that is, $\lambda_{BC} \gg \lambda_{AC}$. There are a few points that should be noted, however. (1) the terminal results are given after running sufficiently long that no reverse-trends seem to arise; (2) while all the N_{B_i} clearly must be equal, yet the $\lambda_{B_i C}$ are not necessarily so, since the probabilities $U(\lambda)$ and $D(\lambda)$ are consulted independently for each i in $\lambda_{B_i C}$ (see 2.3.4); (3) the threshold curves are set so that a total stimulus of 15 may fire neuron C for $r_C \geq 11$; thus, initially any neuron may fire C. The steepest of the threshold curves, number 8, will allow four neurons whose synapse-values to C are maximal (15) to fire C for $r_C \geq 7$. The effect of the steepening threshold curves is to decrease the total activity of neuron C, i.e., N_C decreases; (4) (Not shown in Figure 7) the steepening threshold curves tended to accelerate the rate at which λ_{AC} decreased; (5) in Run 2, one of the $\lambda_{A_i C}$'s ($\lambda_{A_4 C}$) remained high (although it was still decreasing when the run terminated).

Thus, Experiment 3, under the conditions for which it was performed, was successful. However, the significance of this success is not clear; little was actually demonstrated about the network parameters that was not already clear. Moreover, the condition of synchrony is so strong that the experiment almost had to work for any reasonable selection of the parameters. Therefore, further exploration with it was abandoned in favor of the more interesting

case described in the next section where, somewhat as in Experiment 2, the neurons of group B fire randomly and independently only in an "on-period" and are silent otherwise.

3.5.3 Experiment 4, Asynchronous Case

The relationship of the firing rates of the neurons of group A and B was chosen as follows: as in Experiment 3, $X_{A_i}(t) = 1$ with probability f_A . However, $X_{B_i}(t) = 1$, with probability f_A independently of the $X_{A_i}(t)$'s and of $X_{B_j}(t+k)$, $k = 0, 1, 2, \dots$ for $j \neq i$, only for t in the intervals $[2kl, (2k+1)l]$ for $k = 0, 1, 2, \dots$, where l is again the length of the interval. Such intervals are called the on-periods for the neurons of B. On the complementary intervals (off-periods), $X_{B_i}(t) = 0$. The I_{A_i} were taken equal to a constant $I (= I_{B_i}$ when $X_{B_i}(t) = 1$).



In the runs to be described, l was taken as 60, f_A as $1/6$. Notice that f_{B_i} averaged over a full-cycle (120 time steps) is $1/2 f_{A_i}$; that is, the neurons of B fire one-half as often as those of A, thus the hypothesis H2 is completely satisfied.

The runs for this experiment were designed to gain further information about the form of the threshold and synapse-value curves and to derive workable values of $U(\lambda)$ and $D(\lambda)$.

Runs 1-8. The network functions and initial conditions were chosen exactly as in Experiment 3, using again the threshold functions of Figure 5. The

stimulus pattern was as described above. The results of these runs are displayed in Figure 8, with the same interpretation of symbols as in Figure 7.

Comment on Runs 1-8. As seen from Figure 8, these runs were not successful. They do indicate one thing, however; that is a tendency for the λ 's to plunge to zero. It is not shown in the figure that the λ_{BC} 's dropped more slowly than the λ_{AC} 's. Thus, the probabilities $U(\lambda)$ and $D(\lambda)$ are suspect. However, the threshold curves were deficient in that they are all too high to allow any single neuron of B, even with maximum synapse-value, to fire C if $r_C < 11$. Thus, a sort of upper bound on the firing rate of C is established, reducing the number of favorable situations $\delta_{B_i}(t) = 1$ & $\delta_C(t+1) = 1$ for incrementation of λ .

Runs 9-16. The purpose of these runs was to test the behavior of the system, given the same initial conditions and network functions each time, for a series of different values of $U(\lambda)$ and $D(\lambda)$. The threshold curve used for these runs is given in Figure 9, $U(\lambda)$ and $D(\lambda)$ for each run in Table 4. These probabilities were chosen so that the ration $U(\lambda)/D(\lambda)$ varied from 2:1 to 15:1 in seven equal steps. The linear $S(\lambda)$ curve of Experiment 1 was used. Length of the runs was 3000 time steps. The initial conditions were as in Runs 1-8 except that $\lambda_{A_i C}(0) = 10 = \lambda_{B_i C}$, $i = 1, \dots, 4$. The results are shown in Figure 10.

Comment on Runs 9-16. Of these runs, Run 12 was the most successful, although (not shown in Figure 10), $\lambda_{B_3 C}$ and $\lambda_{B_4 C}$ were still decaying when the run was terminated. These runs clearly illustrate how sensitive the network is to the settings of $U(\lambda)$ and $D(\lambda)$. Thus, the ratios $D/U = 1/7, 1/9$, etc., are clearly too strong—all the λ 's rise; whereas the ratios $1/2, 1/3$, and possibly $1/4$ are too small—most, if not all, of the λ 's decrease to 0.

Runs 17-32. The synapse-level probabilities $U = .3138$ and $D = .06373$ of Run 12 were taken as tentative workable values for these quantities. The curve for $S(\lambda)$ was given the non-linear form of Figure 11. This particular form was chosen to bias changes upward when λ is large, downward when λ is very small, and to provide for gentle transitions in the midrange of λ . Runs 17-24 were done using the threshold curves of Figures 12 and 13 with the initial conditions $\lambda_{A_i}C(0) = \lambda_{B_i}C(0) = 12$ and all recovery-states at maximum. Runs 25-32 are identical except for the initial conditions $\lambda_{A_i}C(0) = \lambda_{B_i}C(0) = 10$. The results of these runs are given in Figure 14.

Comment on Runs 17-32. The most successful were Runs 20 and 28; the overall results were somewhat disappointing, however. See the following section for a discussion of the difficulties and possible solutions.

3.5.4 Analysis and Comment on Experiment 4

The hypothesis H2 was confirmed for two cases (Runs 20 and 28) using threshold curve 20 of Figure 12, the synapse-value curve of Figure 11, the values 0.3138 and 0.06373 for $U(\lambda)$ and $D(\lambda)$ respectively, and two different initial values for $\lambda_{A_i}C(0)$ and $\lambda_{B_i}C(0)$ ($i = 1, \dots, 4$). The failures, however, outnumber the successes and, consequently, several serious questions arise:

(1) Presumably one choice of the threshold curve should be universal; that is, one particular curve should work for a variety of initial conditions; yet in this experiment the results seemed keenly dependent on the form of the curve so that a relatively small change in initial conditions produced a sharp change in the final results. Likewise, the threshold curves used here were geared to the eight-input schema and would not work for larger or smaller numbers of inputs. This brings up the second question:

(2) The absence of a non-trivial fatigue function seems unrealistic. That is, a neuron that fires at a high rate for a period of time would be expected to become fatigued. Thus, fatigue would produce a dampening effect on the behavior of the neurons. Also, the appropriate fatigue function could adjust the threshold curve to a varying number of inputs, thus answering the objection of (1) and ensuring the existence of a universal threshold curve.

(3) Bearing in mind the dampening effect of fatigue, it appears that the relationship of firing of the neurons of group A with those of group B is too stringent; that is, in the on-periods, the neurons of B fire at the same rate as those of A and C is being required to discriminate between the two solely on the basis of one of them being shut off periodically! This suggests that the firing-rates of the neurons of B should be fairly high in the on-period, while the rate of the neurons of A should be lower throughout. With a suitable fatigue function, however, the neurons of B would become damped toward the end of the on cycle; likewise during the off cycle they would rest (their fatigue-values tend back to 1). The neurons of A would fire at a rate producing little or no fatigue. This rate would form a type of background frequency for the models.

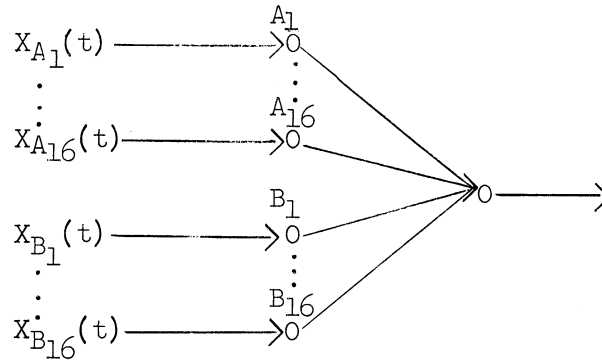
These considerations led to the series of experiments described in the following sections in which some of the results of the analysis of Crichton's thesis [7] were introduced.

3.6 THIRTY-THREE NEURON EXPERIMENTS

3.6.1 General

For this series of experiments, the size of the networks was increased from $N = 9$ to $N = 33$ and $M = 16$, that is, the schema of 3.3.2 becomes a thirty-three neuron network consisting of the two groups A and B of neurons of sixteen neurons each and the single neuron C which receives one connection from each neuron of A and B. The vector $X_{O_1}(t)$ becomes $(X_{A_1}(t), \dots, X_{A_{16}}(t))$,

$X_{B_1}(t), \dots, X_{B_{16}}(t)$ where $A_i \in A$ and $B_i \in B$ $i = 1, \dots, 16$; likewise the f_{α_i} become f_{A_i} and f_{B_i} , etc:



Recalling the remarks made at the end of the last section, the basic hypothesis H1 becomes the following:

H3. For appropriate selections of the network functions V , ϕ and S and appropriate initial conditions, neuron C will tend to correlate with group B in the sense that as t becomes sufficiently large,

$$\lambda_{\overline{BC}}(t) \gg \lambda_{\overline{AC}}(t)$$

and

$$\text{for all } i, \quad \delta_{B_i}(t) = 1 \implies \delta_C(t+1) = 1 \quad \text{a.a.}$$

and

$$\delta_{B_i}(t) = 0 \implies \delta_C(t+1) = 0 \quad \text{a.a.}$$

over some range of the rates f_{A_i} and f_{B_i} such that $f_{B_i} > f_{A_j}$ ($i, j = 1, \dots, 16$) locally but where the f_{A_i} 's and f_{B_i} 's have a common average, f_b over the interval $[0, \infty]$. "Locally" means that over certain sufficiently small time intervals the relationship $f_{B_i} > f_{A_j}$ holds.

The intent of the hypothesis is, given that the neurons of B are periodically interrupted as in Experiment 4, that the f_{B_i} 's be greater than the f_{A_j} 's within the on-periods of the neurons of B, equal to or less than the f_{A_i} 's in the off-periods of B, but that the average values of the frequencies

over large time intervals be close to the common value f_b . f_b will be called the background rate of the system.

3.6.2 Some Theoretical Considerations

In the appendix to his thesis [7], Crichton discusses the stability of systems of neurons which he calls "semi-autonomous subsystems." These are networks of neurons which may correspond in a limited way to the cell-assemblies of Hebb's theory. In his development in which, unlike the approach of this paper, he is concerned with the statistical properties of a very large set of neurons, he makes a number of assumptions, two of which are relevant to the experiments of this chapter: (1) the neurons of the system fire aperiodically, randomly and independently of one another, and (2) all neurons tend in their firing to a common average rate f_b . This f_b he calls the nominal system average. From his arguments he derives some bounds on the threshold curve (to be discussed later) and some important relationships between the fatigue increments, Δ_1 and Δ_2 , and the probabilities of synapse-level change, $U(\lambda)$ and $D(\lambda)$. The gist of his argument is this; that the role of the fatigue function must be to drive the neurons of the system to the frequency f_b ; thus, if a neuron falls below f_b in its firing rate, then the fatigue should decrease so as to bring the rate back up to f_b ; likewise, if its rate exceeds f_b , fatigue should increase so as to bring the rate back down to f_b . Firing at the rate of f_b , there is no net change in fatigue. This last condition implies that $f_b = \Delta_2 / (\Delta_1 + \Delta_2)$ since then $Tf_b\Delta_2 - T(1-f_b)\Delta_1$ must be zero, where T is the length of the time-interval under consideration (see 2.3.3). Similarly, the condition for no net change in synapse-level becomes $f_b = D(\lambda) / (U(\lambda) + D(\lambda))$.

One further relation that he gives is useful: Consider two neurons A and C with a connection going from A to C, where A and C fire aperiodically

at the rates f_A and f_C respectively. Therefore, the expected rate of increase in λ_{AC} per time step is $f_A \cdot f_C$, and the expected rate of decrease is $f_A(1-f_C)$. Recalling that $D/(U+D) = f_b$ from which $U/D = (1-f_b)/f_b$, one sees that $U = K(1-f_b)$, $D = Kf_b$ for some constant $K > 0$. U and D correspond to the rate of increase and the rate of decrease of a connection and $f_A f_C K(1-f_b)$ to the expected rate of increase in λ_{AC} per time step, $f_A(1-f_C)Kf_b$ to the rate of decrease in λ_{AC} per time step. Therefore, the expected net rate of increase in λ_{AC} per time step is

$$f_A f_C K(1-f_b) - f_A(1-f_C)Kf_b = Kf_A(f_C - f_b) . \quad (F)$$

This is positive, i.e., λ_{AC} is increasing, if $f_C > f_b$ (f_A , f_C , and f_b are all assumed positive or zero), negative, i.e., λ_{AC} is decreasing, if $f_C < f_b$ and zero if $f_A = 0$ or $f_C = f_b$. This relation (F) Crichton gives as the fundamental formula for trends in synapse-levels.

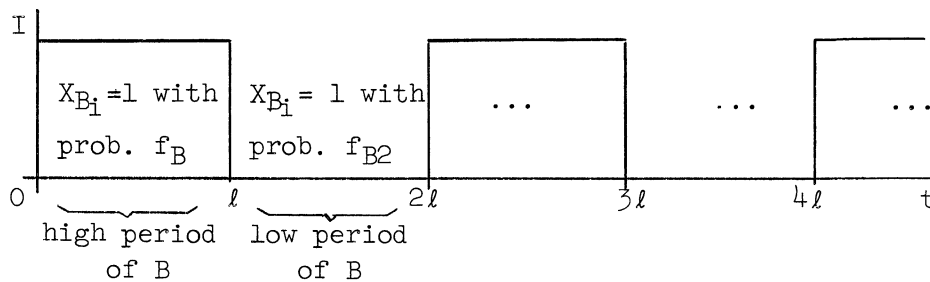
These relationships provide very useful guides and will be referred to in the following. However, a few points should be noted: (1) In the current experiments, the assumption of independence of firing of the neurons does not hold. As N increases, however, one would expect it to become more plausible. The validity of Crichton's analysis therefore increases with N in the present situation. (2) Although his theory yields fruitful relations between Δ_1 , Δ_2 , $U(\lambda)$, and $D(\lambda)$ and is useful in analyzing trends in synapse-levels, yet, beyond the bound mentioned it says nothing about the form of the threshold, fatigue, and synapse-value functions.

It should be noted that the rates for zero change in synapse-levels and zero change in fatigue-level need not be identical, that is $D/(U+D)$ may equal f_{b1} and $\Delta_2/(\Delta_1+\Delta_2)$ may equal f_{b2} . In fact, in the first group of experiments to be described below, before the condition $f_b = D/(U+D)$ and the relationship (F) were discovered by Crichton, $D/(U+D)$ was 1/6! (This occurred partly

through failure of the author to digest the import of his analysis, which was being developed about the same time as these experiments were conceived. Such are the pitfalls of the experimental approach!)

3.6.3 Experiment 5, Fatigue Curve Tests

The relationship of the firing rates of the neurons of groups A and B was chosen as follows: the neurons of A are assumed to fire at the rate f_A ; that is, $f_{A_i} = f_A$ for $i = 1, \dots, 16$. As in Experiment 4, the neurons of B fire periodically so that $X_{B_i}(t) = 1$ with probability f_{B1} in the intervals $[2k\ell, (2k+1)\ell]$ and $f_B = f_{B1} > f_A$. In the complementary intervals, unlike the previous case, $X_{B_i}(t) = 1$ with probability $f_{B2} \ll f_A$. f_{B1} is called the high frequency of group B, f_{B2} the low frequency. The intervals in which f_{B1} applies correspond to the on-periods in Experiment 4.



In the runs to be described, f_A was taken to be $1/10$, $f_B = 1/5$, $f_{B2} = 1/1000$. ℓ , the length of the high-period of B, was taken as 200 time steps. Notice that in the high period, $f_{B_i} = 2f_{A_i}$, but that over a complete cycle, f_{B_i} averages very closely to f_{A_i} ($1/2(1/5+1/1000) \approx 1/10$). The threshold and synapse-value curves used are shown in Figure 15. $U(\lambda) = 0.3138$ and $D(\lambda) = 0.06373$. The quantities Δ_1 and Δ_2 were set to $5/8$ and $1/16$ respectively (thus $f_b = 1/11$ and is close to f_A). Initially, all neurons were rested, that is, $l_C(0) = l_{A_i}(0) = l_{B_i}(0) = 31$ and $r_C(0) = r_{A_i}(0) = r_{B_i}(0) = 16$,

and $\lambda_{A_i C}(0) = \lambda_{B_i C}(0) = 8$ for $i = 1, \dots, 16$. The input stimuli I_{A_i} & I_{B_i} were taken to be the constant value 1000. Each run performed was terminated after 2100 time steps.

Four runs were performed for this experiment, each one using a separate fatigue curve from Figure 16. (Run i used curve i of the figure, $i = 1, 2, 3, 4$.) Letting $S_A(t)$ be the value of the sum of the $S_{A_i C}(t)$ and $S_B(t)$ that for the $S_{B_i C}(t)$, the terminal results of these runs may be stated as follows:

Run 1: $S_A = 32, \quad S_B = 22$

Run 2: $S_A = 40, \quad S_B = 80$

Run 3: $S_A = 22, \quad S_B = 53$

Run 4: $S_A = 14, \quad S_B = 79$

(Note: In this and following experiments, it was found to be more convenient to refer to the synapse-value, S , rather than to the synapse-level λ .)

The statistics S_A and S_B do not reflect the dispersion of the $S_{A_i C}$ or $S_{B_i C}$, which in fact, was considerable. In each run the number of negative S 's was about the same as the number of positive. The detailed history of synapse-value changes for Run 4 are shown in the Appendix. Also, in all four runs, the firing rate of C was less than or equal to f_A (see 3.6.2) and thus the condition (F) would predict no (uniform) increase in synapse-levels.

3.6.4 Analysis and Comments on Experiment 5

The purpose of this experiment primarily was to obtain a good starting setting for the fatigue function $\phi(l)$. Of the four runs performed, the fourth was the most successful in the sense of the hypothesis. However, even there, of the $S_{B_i C}$, five were still decreasing when the run was terminated, five were increasing, and only six were stable, though not large. Therefore, Run 4 was repeated, but allowed to run for 10,000 time steps. The results were:

Run 5: $S_A = -129, \quad S_B = -89$

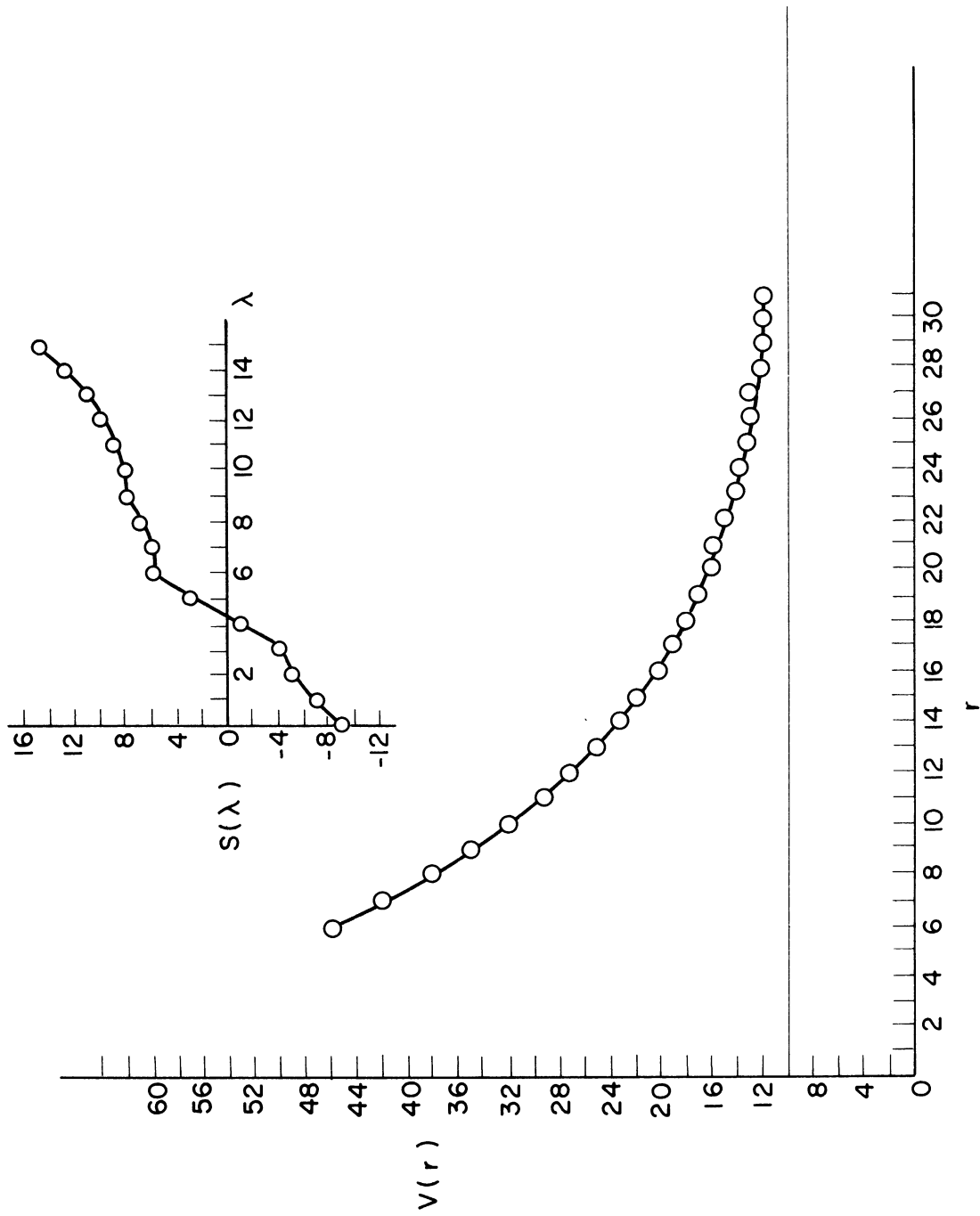


Figure 15. Threshold Curve and Synapse-Value Curves for Experiment 5.

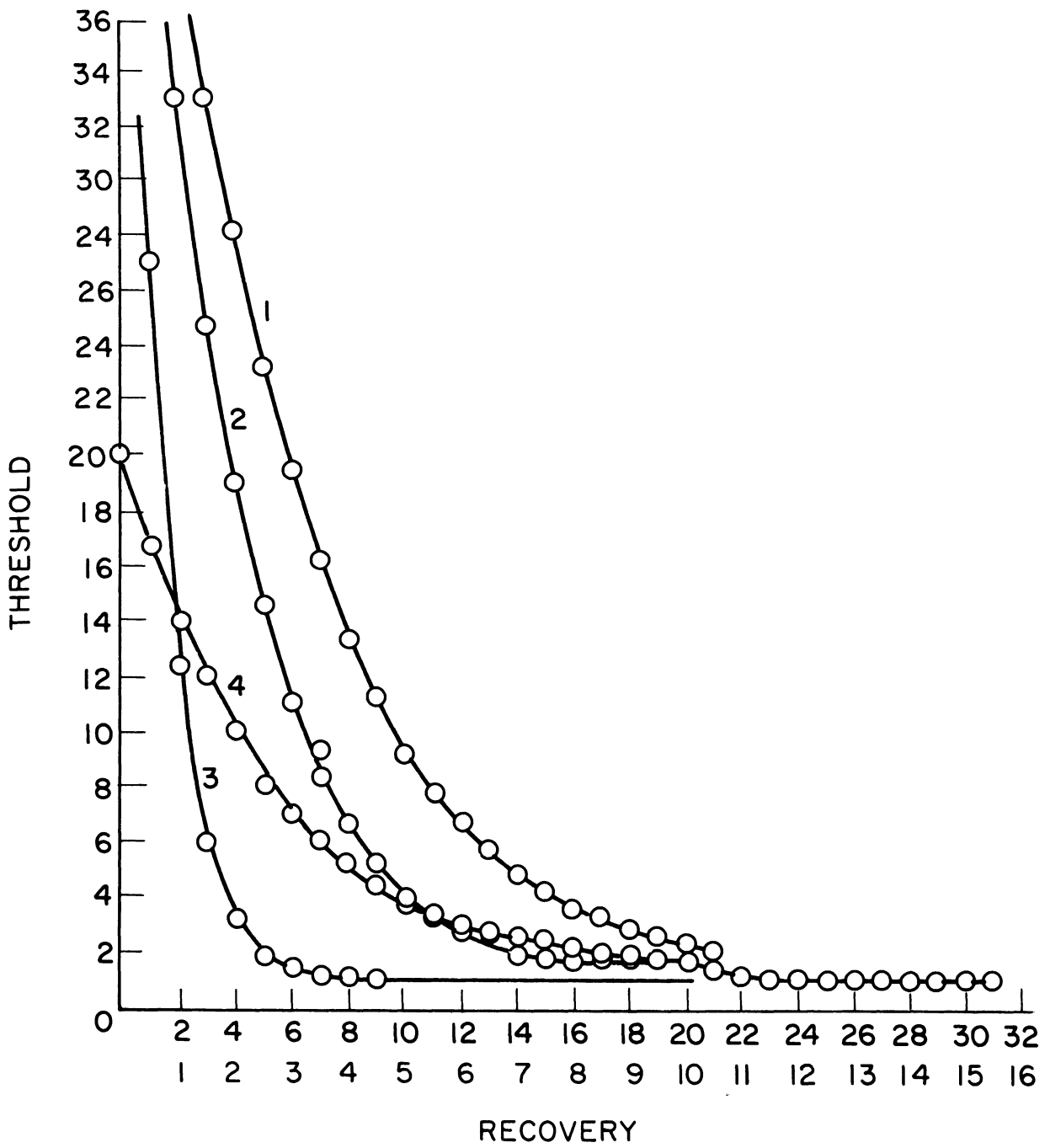


Figure 16. Fatigue Curves for Runs 1-4, Experiment 5.

S_A had decayed to zero by 2351 time steps, S_B by 7801. Thus, though this run was a failure, yet at least the S_{B_i} 's decayed at a lower rate than the S_{A_i} 's and were not as negative as the S_{A_i} 's. To test the rate of decay of the synapse-values per se, a variant run of length 10,000 time steps was carried out in which l was made equal to 10,000, thus the neurons of A were firing at the rate of $1/5$ over the interval 0-10,000, the neurons of B at the rate of $1/10$ over the same interval. The results were:

Run 5': $S_A = -144$, $S_B = -144$

(-144 is the minimum for the sum of the S's). The rate of decay to zero for S_A was 2951 time steps, for S_B it was 1251.

These results suggested a series of initial trend studies in which the initial behavior of the S's could be studied in detail and in which the parameters Δ_1 , Δ_2 , U, and D and the fatigue curve could be varied and the effect of this variation studied. This series comprises Experiment 6.

3.6.5 Experiment 6, Initial Trend Studies

The purpose of this experiment was to examine in detail the effect of varying the parameters Δ_1 , Δ_2 , U and D upon the initial developments of the synapse-values. The experimental arrangement was identical to that of Experiment 5; that is, $l = 200$, $f_A = 1/10$, $f_B = 1/5$, all synapse-levels were set to 8 initially, etc. The threshold, fatigue, and synapse-value functions used are given in Figure 17.

Four separate tests were conducted, the results of which are summarized below. Each run was terminated at the end of 500 time steps. Notice that the initial value of the sums of synapse-values, S_A and S_B , for the synapse-value curve given, is 112 (= 16×7 where $S(8) = 7$ is the value of $S(\lambda)$ for $\lambda = 8$).

Run 1. The parameter values taken were $\Delta_1 = 10/16$, $\Delta_2 = 1/8$, $U = 0.3138$, and $D = 0.06373$. $\Delta_2/(\Delta_1+\Delta_2) = 1/11$.

Result: The final synapse-values were (using the same notation as for Experiment 5):

$$S_A = 115, \quad S_B = 120.$$

Neuron C fired at a slightly higher rate than the neurons of A or B.

Run 2. Δ_1 and Δ_2 were taken as in Run 1, however, U and D were both set to 0.5.

Results: Both S_A and S_B tended rapidly to the minimum value of -144 for the sum of the synapse-values; however, the sum S_B decreased at a lower rate than that of S_A . Neuron C fired at a much lower rate than that of the neurons of A or B.

Run 3. The parameter values taken were: $\Delta_1 = 1.0$, $\Delta_2 = 1/8$, $U = 0.3138$, and $D = 0.06373$. Notice that $\Delta_2/(\Delta_1+\Delta_2) = 1/9$ and $D/(U+D) \approx 1/6$.

Results: The final synapse-values are

$$S_A = 121, \quad S_B = 127.$$

The firing rate of neuron C was about 30% greater than that of the neurons of A or B.

Run 4. The parameter-values taken were: $\Delta_1 = 1.0$, $\Delta_2 = 3/16$, U and D as in Run 3. $\Delta_2/(\Delta_1+\Delta_2)$ becomes $3/19$ ($\approx 1/6$).

Results: The final synapse-values were

$$S_A = 108, \quad S_B = 112.$$

Neuron C fired about twice as often as the neurons of A or B.

3.6.6 Analysis and Comments on Experiment 6

One notices that in all but one run on this experiment, the condition (F) of Crichton holds for a net increase in synapse-value, since the firing

rate of neuron C was greater than that of the neurons of A or B and consequently was greater than the background rate of $1/10$. (This experiment was performed before the author realized that $\Delta_2/(\Delta_1+\Delta_2)$ and $D/(U+D)$ must also be equal to $f_b = 1/10$. However, this situation does not alter the conclusions drawn here.) The anomalous run, Run 2, simply allowed too much change downward in λ and could hardly have been expected to work (that is S_{B_iC} rise, S_{A_iC} fall to about 0); yet it was instructive in that it showed a certain sluggishness on the part of the λ_{B_iC} to move downward even with the high value of $D = 0.5$. The remaining runs, Runs 1, 3, and 4 were favorable in their outcomes, but just barely so. One would expect, in light of the remark above about condition (F) being fulfilled for the case of an increase in $\lambda(S(\lambda))$, a stronger trend upwards of the λ_{B_iC} or at least, a stronger trend downwards of the λ_{A_iC} . Thus, a series of tests on the fatigue function and the parameters Δ_1 and Δ_2 is indicated.

3.6.7 Experiment 7. Further Tests on the Fatigue Function

The purpose of this experiment was to find a fatigue function which, together with the appropriate values of Δ_1 and Δ_2 , would accelerate the upwards trend of the λ_{B_iC} 's or at least the downwards trend of the λ_{A_iC} 's. It is certainly not essential for the hypothesis H3 to hold that the λ_{B_iC} 's all tend to the maximum or that the λ_{A_iC} 's tend to the minimum. In fact, the analysis of Crichton seems to suggest that the λ_{A_iC} 's be such that S_A be close to zero and the S_{A_iC} 's be zero or moderately positive or moderately negative, whereas the λ_{B_iC} 's should be such that the S_{B_iC} 's are strongly positive.

The role of the fatigue function seemed, at this point, so critical that the exact form of the threshold function did not seem crucial. Therefore, the threshold function was not varied throughout this experiment. Likewise, neither was the synapse-value curve. This procedure is perhaps open to ques-

tion; however it is somewhat analogous to that of finding a local maximum of a function $f(x, y, z)$ by fixing two points x_0 and y_0 and maximizing f with respect to z , then holding z and x_0 fixed, maximize with respect to y , etc. Moreover, the results thus far indicated that the problem centered more around the fatigue function than around the others.

The various tests given below utilize the same experimental arrangement as in Experiment 6 unless mentioned otherwise, that is:

1. all $\lambda(0)$'s are set to 8; hence, initially $S_A = S_B = 112$;
2. all $r(0)$'s are set to 16;
3. $f_B = 1/5$, but $f_A = 1/20$, $l = 200$;
4. $V(r)$ and $S(\lambda)$ are given in Figure 17;
5. $\phi(l)$, Δ_1 , and Δ_2 will be specified in each test;
6. $U = 0.3138$, $D = 0.06373$

Notice that the decrease in f_A now implies that the average rates of firings of the two groups A and B differ; for A it is $1/20$, for B it is $1/10$ (f_B will be varied in the tests below). The intent was to simulate the condition in which for a period of time the average rate of B exceeds that of A, but in which over a longer interval of time, it would reduce to the system average $f_b (= 1/20)$.

Run 1. The fatigue curve tested in this run is given in Figure 17. The parameters were $\Delta_1 = 1$ and $\Delta_2 = 1/8$, thus $\Delta_2/(\Delta_1+\Delta_2) = 1/9$. Length of the run was 1000 time steps. The final values of the synapse-value sums were $S_A = 119$ and $S_B = 123$. S_A went to a maximum of 125 before decaying to 119, S_B went to a maximum of 136. The firing-rate of C was greater than that of the neurons of A or B.

Run 2. Run 1 was repeated for the fatigue curve of Figure 18; length of the run was reduced to 500 time steps. The final results were $S_A = 117$ and

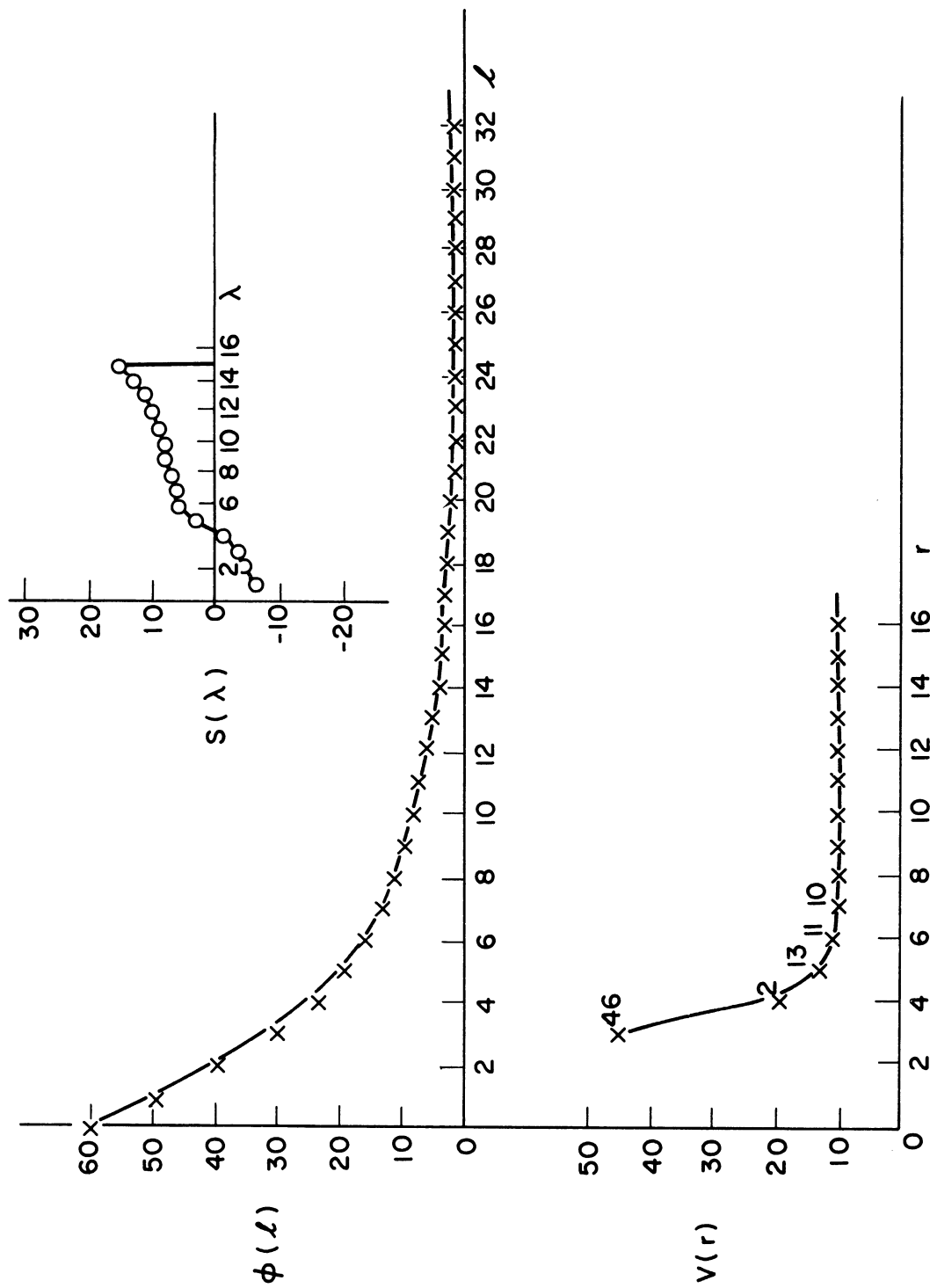


Figure 17. Threshold, Fatigue, and Synapse-Value Functions for Experiment 6 and Run 1 of Experiment 7.

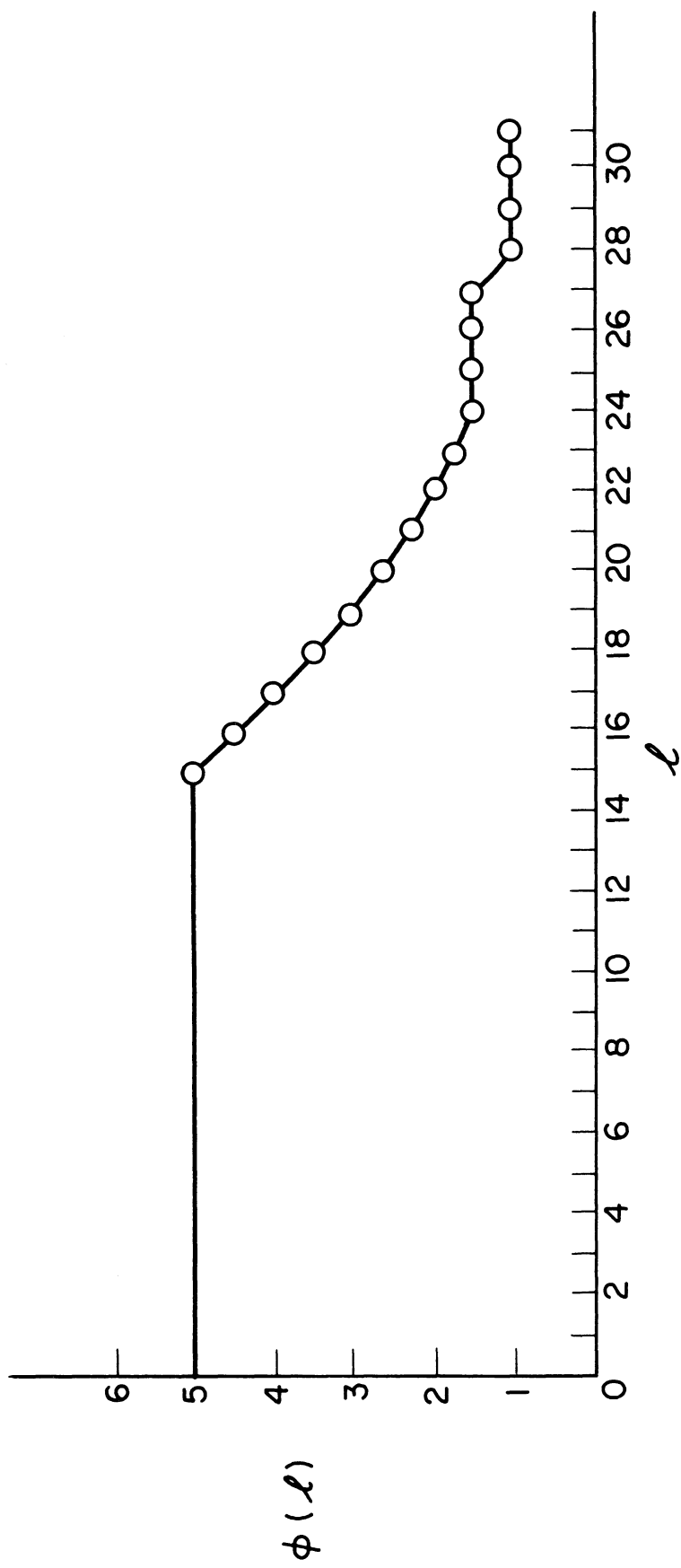


Figure 18. Fatigue Curve for Run 2, et seq., Experiment 7.

and $S_B = 135$ (these were the maximum values also). The firing rate of C was again greater than that of the A_i or B_i .

Run 3. Run 2 was repeated except that Δ_1 and Δ_2 were modified: $\Delta_1 = 1$, $\Delta_2 = 1/16$, hence $\Delta_2/(\Delta_1+\Delta_2) = 1/17$; length of run was the same as for Run 2. The final results were $S_A = 114$, $S_B = 122$. Neuron C fired at a lower rate than the neurons of B, but greater than those of A.

Run 4. Again Run 2 was repeated, now for the values $\Delta_1 = 5/8$, $\Delta_2 = 1/16$, $\Delta_2/(\Delta_1+\Delta_2) = 1/11$, C fired at a higher rate than the A_i or B_i and the final synapse-value sums were $S_A = 121$ and $S_B = 127$.

Run 5. Run 2 repeated for the values $\Delta_1 = 3/4$ and $\Delta_2 = 1/16$, $\Delta_2/(\Delta_1+\Delta_2) = 1/13$ and terminated after 10,000 time steps. The results were: $S_A = -71$ and $S_B = 68$. C fired at a lower rate than the B_i , higher rate than the A_i .

Run 6. $U = 0.5402$, $D = 0.04502$, $\Delta_1 = 3/4$, $\Delta_2 = 1/16$, so that $D/(U+D) \approx 1/12$, $\Delta_2/(\Delta_1+\Delta_2) = 1/13$, otherwise like Run 2. This run was terminated after 10,000 time steps with $S_A = 201$, $S_B = 229$. C fired at a slightly lower rate than the B_i , at a much greater rate than the A_i .

Run 7. Exactly like Run 6 except that f_B was changed to $1/6$ and f_A to $1/13$. After running for 10,000 time steps, $S_A = 142$ and $S_B = 223$. C's firing rate was greater than that of the B_i by about 5%.

Run 8. For this run a variation in the firing patterns of groups A and B was introduced in which in the high periods of B, the neurons of B fired at the rate $1/6$ and the neurons of A fired at the rate of $1/20$, while in the low-period of B, the neurons of B fired at the rate of $1/1000$ and the neurons of A at the rate $1/13$. Otherwise, all parameters remained the same as in Run 7.

After 10,000 time steps, $S_A = 156$, $S_B = 228$ and the firing rate of C was again about 5% greater than that of the B_i .

3.6.8 Analysis and Comments on Experiment 7

The most successful of the eight runs above (excluding from consideration for the moment Run 8) was Run 7. It should be noted that for this run,

$$\frac{D}{U+D} \approx \frac{\Delta_2}{\Delta_1+\Delta_2} = f_b = \frac{1}{12}$$

as the theory of Crichton requires. Notice that the average firing rate of the A_i 's and the B_i 's is $1/12 \approx f_b$. Thus, this experiment seems to be in accord with the theory. Run 8, which purports to be a slight generalization where A and B fire in phase at alternately high and low frequencies, likewise seems to conform to the theory.

There are, however, some disturbing signs: namely, that again the effect of the fatigue-function is not very sharp and there is far too much variation among the $\lambda_{A_i C}$'s and $\lambda_{B_i C}$'s. Moreover, the entire range of the fatigue-level is not used, primarily just the values in the steep portion of the curve. Finally, the selection of the threshold and synapse-value functions still did not seem quite satisfactory.

The results of this experiment prompted much reflection about the nature of the threshold, fatigue, and synapse-level functions, with the result that a derivation was obtained for the form of the threshold curve. This is discussed in the next section.

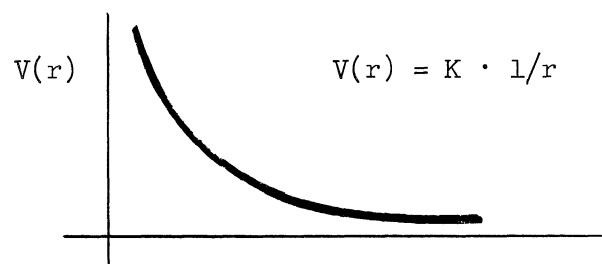
3.7 COMMENTS ON THE NETWORK FUNCTIONS V , ϕ , AND S

3.7.1 The Threshold Function, $V(r)$

Consider a neuron C with N inputs A_i , $i = 1, \dots, N$ and suppose that the input neurons are all firing randomly and independently at the rate f .

What then are the desired properties of the threshold function as far as controlling the firing rate of C is concerned? Over a short time interval, so that the effects of the fatigue function do not enter in to complicate the procedure, we may assume that the threshold curve should be such that C fires at the rate f also. (We could, of course, assume that it fires at some rate $f_1 \neq f$, but here let us restrict attention to the case $f_1 = f$.) Assume now that the synapse-values λ_{A_i} , $i = 1, \dots, N$, are all equal. With no other inputs, then, the threshold function $V(r)$ must be such that for $r \geq 1/f$ neuron C fires. Since the set of input neurons basically is a Bernoulli process (ignoring the effects of the absolute recovery period), the expected number of input neurons active at any time step is $m = Nf = N \cdot 1/r$. The expected input stimulus to neuron C per time step then is $mS(\lambda) = NfS(\lambda) = N \cdot \frac{1}{r} \cdot S(\lambda)$. This says that, if N is fixed, the threshold curve should vary linearly with respect to the input frequency f (at which it is assumed that C should fire also), i.e., $V(r)$ varies with $1/r$, $V(r) = K \cdot \frac{1}{r}$ where the constant K is determined by the expected amount of input stimulus per time step.

Thus, $V(r)$ seems to be a hyperbolic function r :



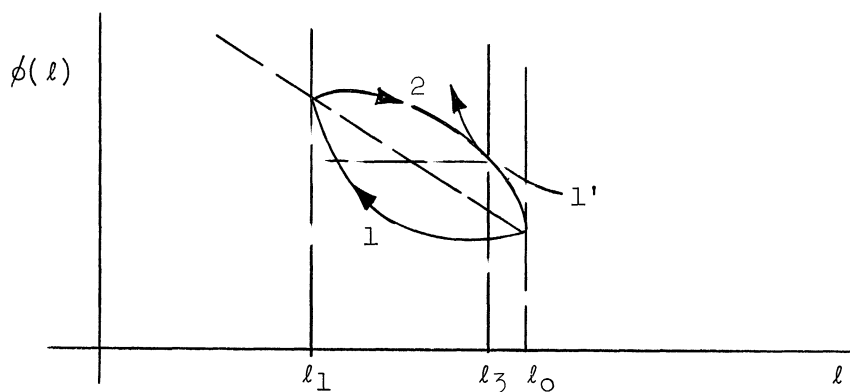
This form of $V(r)$ conforms to the bounds required in the development of Crichton's thesis [7] and to which the reader is referred for further details.

3.7.2 The Fatigue Function, $\phi(\ell)$

The main reason for the failure of the fatigue function to perform as desired in the experiments of this chapter seems to center about the fact that

l , after having been reduced to an adequately low value through decrementation by Δ_1 (hence $\phi(l)$ is large and very effective in damping firing of the neuron), the rate of recovery (back to large value of l , $\phi(l) \approx 1$) is altogether too rapid in spite of the small value of Δ_2 . Thus, in Experiment 7, Run 7, the fatigue function should be such that towards the end of the high-period of B, C is highly fatigued and does not recover so that any neuron can fire it, let alone a neuron of A, until towards the end of the off-period of B. Yet this was not the case in this experiment—C recovered quite rapidly and the neurons of A could fire it in about 20 time steps after the beginning of the low-period of B. Thus, the λ_{A_iC} 's had opportunity to develop, whereas in the ideal case they should have little or no such opportunity.

It turns out that no single-valued function $\phi(l)$ will give the desired effect. Instead, $\phi(l)$ has to be a hysteresis-type curve where, as l decreases, $\phi(l)$ increases at one rate ρ_1 and when l increases, $\phi(l)$ decreases at another rate ρ_2 . The rates ρ_1 and ρ_2 in fact should be such that ρ_1 increases as l decreases, ρ_2 likewise decreases as l increases. Pictorially, this is as follows:



Suppose l is decreasing. Then, $\phi(l)$ follow curve 1. If then the neuron ceases to fire, instead of recovering along curve 1, it recovers along curve 2. If at point l_3 the neuron fires, then instead of following curve 2, it follows

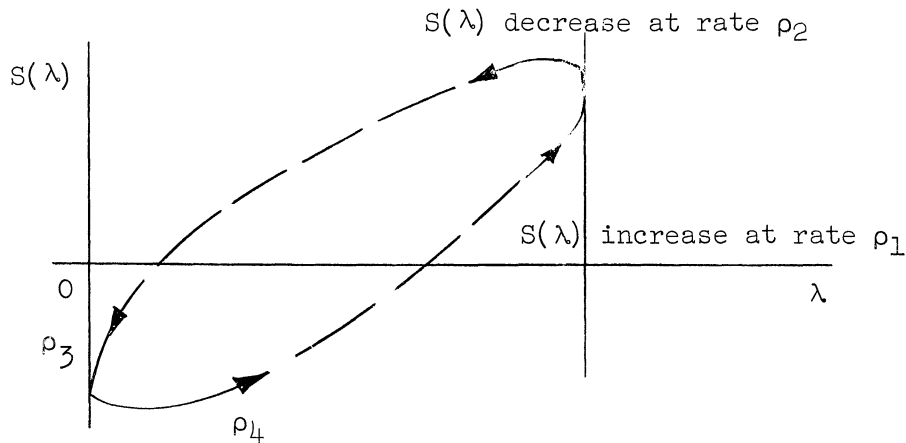
curve 1, which is 1 "moved-over" so to speak. (This picture is deceiving since "curves" 1 and 2 are really rates, and there is no actual shifting of curves.)

Given this type of fatigue function, the fatigue-value of a neuron would increase initially gradually, then progressively more as the firing-rate of the neuron increased until this rate would be suppressed for a period of time in which the fatigue-value would very gradually decrease, after which the neuron would resume firing at a lower rate. This type of function would almost guarantee success in the experiments of this chapter.

This form of $\phi(l)$ is readily implemented and is to be used in the next series of experiments.

3.7.3 The Synapse-Value Function, $S(\lambda)$

Similar remarks to those concerning fatigue can be made about the function $S(\lambda)$. In this case, one wants $S(\lambda)$ to increase, for large λ , at a rate ρ_1 and decrease at a lower rate ρ_2 ; likewise for small λ , $S(\lambda)$ should decrease at a rate ρ_3 and recover at a lower rate ρ_4 :



This means that a synapse-level of a connection would gradually at first, then more rapidly later as the activity of the connection increased, build-up, to decay slowly at first, then more rapidly later as the activity subsides.

Again, in retrospect, this form of $S(\lambda)$ would strengthen the results of this chapter. This form of $S(\lambda)$ is easily implemented in the model and will be used in the next series of experiments.

4. CONCLUSION

A series of experiments on simple, cycle-less neural networks was carried out. A number of problems regarding the nature of the network functions V , ϕ , and S arose and, using the experimental results as a guide, suggestions were made for their resolution. The analysis of Crichton [7] was demonstrated, with reservations, for the small networks considered. For certain input conditions the threshold curve was shown to be a hyperbolic function of the recovery.

The next series of experiments will test the ideas of the last section. (The author had hoped to conclude this series in time for this report, but unfortunately failed to do so.) Following this, a series of experiments is planned in which progressively more complicated feedback among the neurons is introduced.

APPENDIX

DETAILED HISTORY OF SYNAPSE-VALUE
CHANGES FOR RUN 4, EXPERIMENT 5

SYNAPSE VALUES GROUP A
(Slow Group)

Time Step*	S _{A1C}	S _{A2C}	S _{A16C}	S _A
1	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	112
51	7	8	6	7	7	7	8	8	7	6	7	7	6	7	7	7	112
101	7	8	6	6	7	7	8	8	7	6	8	7	7	7	7	7	113
151	8	8	6	6	7	7	8	8	7	6	8	7	6	7	7	8	114
201	8	8	6	6	8	6	8	9	7	6	8	7	6	7	7	8	115
251	8	8	6	6	7	6	8	9	6	6	8	7	6	7	6	7	111
301	8	8	6	3	6	6	8	9	3	6	9	8	6	6	6	8	106
351	8	8	3	6	6	6	8	9	3	3	9	8	6	7	6	7	103
401	8	8	3	3	6	6	8	9	1M	4M	9	8	6	7	6	7	89
451	8	8	3	6	6	6	8	10	1M	5M	9	8	6	7	3	7	89
501	8	8	3	6	3	6	8	10	4M	5M	9	7	6	7	3	6	81
551	8	8	1M**	6	1M	6	8	10	4M	5M	9	7	7	7	3	6	74
601	8	8	1M	3	1M	6	8	11	4M	5M	9	7	7	7	3	6	72
651	8	8	1M	3	4M	6	8	11	5M	5M	9	7	7	6	1M	7	64
701	8	8	1M	3	4M	6	8	11	5M	5M	9	7	7	6	1M	6	63
751	8	8	1M	3	4M	6	8	13	5M	5M	10	6	7	6	3	6	69
801	8	7	1M	1M	4M	6	8	13	5M	5M	10	6	7	7	1M	6	61
851	8	7	1M	4M	5M	3	8	13	5M	5M	9	6	6	7	1M	6	52
901	8	8	1M	4M	5M	6	8	11	5M	7M	9	7	6	7	4M	6	50
951	8	8	4M	4M	7M	6	8	11	5M	7M	8	7	6	7	5M	3	40
1001	8	8	1M	5M	7M	6	8	11	5M	9M	10	8	6	7	5M	6	46
1051	8	8	1M	5M	9M	3	8	11	7M	9M	10	8	6	7	7M	6	37
1101	8	8	4M	7M	9M	3	8	11	9M	9M	11	8	6	7	9M	6	29
1151	9	8	4M	7M	9M	1M	8	11	9M	9M	13	8	6	7	9M	6	28
1201	9	8	4M	7M	9M	1M	8	11	9M	9M	15	8	6	7	9M	6	30
1251	9	8	5M	7M	9M	1M	8	11	7M	9M	13	8	6	7	9M	6	29
1301	9	9	5M	7M	9M	1M	8	11	9M	9M	13	9	6	7	9M	6	29
1351	9	10	5M	7M	9M	1M	8	11	9M	9M	13	9	6	6	9M	3	26
1401	8	11	7M	7M	9M	1M	8	11	9M	9M	13	9	6	6	9M	3	24
1451	9	11	7M	7M	9M	1M	8	11	9M	9M	13	9	6	6	9M	3	25
1501	9	11	7M	7M	9M	4M	8	13	9M	9M	15	9	6	6	9M	3	26
1551	9	11	7M	9M	9M	4M	8	13	9M	9M	15	9	6	6	9M	3	24
1601	9	10	7M	9M	9M	4M	8	13	9M	9M	15	8	6	6	9M	3	22
1651	10	10	7M	9M	9M	5M	8	13	7M	9M	15	9	6	6	9M	3	25
1701	10	11	9M	9M	9M	4M	8	15	9M	9M	13	9	7	6	9M	3	24
1751	10	11	9M	9M	9M	4M	7	13	9M	9M	13	9	6	6	9M	6	23
1801	9	13	9M	9M	9M	4M	7	15	9M	9M	13	10	6	6	9M	6	27
1851	9	13	9M	9M	9M	4M	7	15	9M	9M	15	10	6	6	9M	6	29
1901	9	11	9M	9M	9M	4M	6	15	9M	9M	15	10	6	6	9M	3	23
1951	8	11	9M	9M	9M	4M	6	15	9M	9M	13	10	6	6	9M	3	20
2001	8	11	9M	9M	9M	5M	6	15	9M	9M	13	11	6	6	9M	4M	13
2051	9	11	9M	9M	9M	5M	6	15	9M	9M	13	11	6	6	9M	4M	14

*Synapse values printed every fifty time steps.

**M indicates minus.

SNYAPSE VALUES GROUP B

(Fast Group)

Time Step*	S _{B1C}	S _{B2C}	S _{B16C}	S _B
1	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	112
51	8	7	7	7	6	6	7	7	7	7	7	6	7	7	7	7	7	7	111
101	10	7	7	8	7	6	7	8	7	8	7	6	7	7	7	7	8	117	
151	11	6	7	7	6	6	7	8	7	9	7	6	7	7	7	9	9	117	
201	11	6	7	6	6	6	6	6	8	9	7	6	6	7	6	8	8	111	
251	11	6	7	6	6	6	6	6	8	9	7	6	6	7	6	8	8	111	
301	11	6	7	6	6	6	6	6	8	9	7	6	6	7	6	8	8	111	
351	11	6	7	6	6	6	6	6	8	9	7	6	6	7	6	8	8	111	
401	11	6	7	6	6	6	6	6	8	10	7	6	6	7	6	8	8	112	
451	10	6	7	6	3	6	6	1M	7	11	8	1M	7	7	6	8	8	96	
501	10	7	7	6	6	6	6	6	7	11	8	3	8	8	6	8	8	113	
551	10	7	7	6	6	6	6	6	8	9	8	3	8	8	6	8	8	112	
601	10	8	6	6	6	3	6	3	9	9	9	6	8	7	6	8	8	110	
651	10	8	6	6	6	3	6	3	9	9	9	6	8	7	6	8	8	110	
701	10	8	6	6	6	3	6	3	9	9	9	6	8	7	6	8	8	110	
751	10	8	6	6	6	3	6	3	9	9	9	6	8	7	6	8	8	110	
801	10	8	6	6	6	3	6	3	9	9	9	6	8	7	6	8	8	110	
851	13	9	3	3	6	1M	6	6	8	8	8	6	8	6	6	8	8	103	
901	15	9	6	3	3	1M	6	6	8	10	9	6	8	6	6	7	7	107	
951	15	8	6	1M	1M	1M	6	6	8	10	9	6	8	7	6	6	6	98	
1001	15	8	6	3	1M	1M	6	6	9	10	8	6	7	6	6	6	6	100	
1051	15	8	6	1M	1M	1M	6	6	9	10	8	6	7	6	6	6	6	96	
1101	15	8	6	1M	1M	1M	6	6	9	10	8	6	7	6	6	6	6	96	
1151	15	8	6	1M	1M	1M	6	6	9	10	8	6	7	6	6	6	6	96	
1201	15	8	6	1M	1M	1M	6	6	9	10	8	6	7	6	6	6	6	96	
1251	15	8	6	1M	4M	3	3	7	10	9	8	7	6	6	6	6	6	95	
1301	11	8	6	1M	4M	3	1M	6	10	9	8	7	6	6	6	7	7	87	
1351	13	8	6	4M	4M	6	3	6	10	9	7	8	3	6	6	7	90	90	
1401	13	8	6	5M	4M	6	3	6	11	8	6	8	1M	8	6	6	85	85	
1451	13	8	6	5M	4M	6	3	6	11	8	6	8	1M	8	6	6	85	85	
1501	13	8	6	5M	4M	6	3	6	11	8	6	8	1M	8	6	6	85	85	
1551	13	8	6	5M	4M	6	3	6	11	8	6	8	1M	8	6	6	85	85	
1601	13	8	6	5M	4M	6	3	6	11	8	6	8	1M	8	6	6	85	85	
1651	13	8	7	7M	5M	3	3	6	11	10	3	8	3	10	6	6	85	85	
1701	13	8	7	9M	5M	3	3	6	11	9	6	8	3	10	7	7	87	87	
1751	13	9	7	9M	7M	6	3	6	13	8	6	6	6	10	7	7	91	91	
1801	15	8	8	9M	7M	7	3	1M	11	8	6	6	6	9	6	7	83	83	
1851	15	8	8	9M	7M	7	3	1M	11	8	6	6	6	9	6	7	83	83	
1901	15	8	8	9M	7M	7	3	1M	11	8	6	6	6	9	6	7	83	83	
1951	15	8	8	9M	7M	7	3	1M	11	8	6	6	6	9	6	7	83	83	
2001	15	8	8	9M	7M	7	3	1M	11	8	6	6	6	9	6	7	83	83	
2051	13	8	7	9M	9M	7	3	1M	11	7	7	6	7	8	6	8	8	79	

*Synapse values printed every fifty time steps.
 **M indicates minus.

REFERENCES

1. Burns, B. Delisle. The Mammalian Cerebral Cortex. London, Edward Arnold (publishers) Ltd. (1958).
2. Eccles, J. C. The Neurophysiological Basis of Mind. Oxford, Clarendon Press (1953).
3. Hebb, D. O. The Organization of Behavior. New York, John Wiley and Sons, Inc. (1949).
4. Milner, P. M. "The Cell Assembly: Mark II." Psych. Review, 64, 4, 242-252 (1957).
5. Sharpless, S. K. and Halpern, L. M. "The Electrical Excitability of Chronically Isolated Cortex Studied by Means of Permanently Implanted Electrodes." Electroenceph. Clin. Neurophysiol., 14: 244-255 (1962).
6. Rochester, N., Holland, J. H., Haibt, L. H., and Duda, W. L. "Tests on a Cell Assembly Theory of the Action of the Brain, Using a Large Digital Computer." Transactions on Information Theory, IRE. IT-2, No. 2, Sept. 1956.
7. Crichton, J. W. Doctoral Dissertation. The University of Michigan, 1964.
8. _____, and Holland, J. H. "A New Method of Simulating the Central Nervous System Using an Automatic Computer," Technical Memorandum 2144-1195-M, The University of Michigan, March 1959.
9. _____, and Finley, M. R. Programmed Simulation of Nerve Nets. (Letter to John H. Holland, Summer 1961.)
10. _____. "Requirements on a Function for Computing the Number of Fibers Going from a Given Set of Neurons to a Given Neuron. (Internal Note, Logic of Computers Group), 1961.

UNIVERSITY OF MICHIGAN



3 9015 02826 7758