Publication No.: 9438

MINSKY, Marvin Lee. THEORY OF NEURAL-ANALOG REINFORCEMENT SYSTEMS AND ITS APPLICATION TO THE BRAIN-MODEL PROBLEM.

Princeton University, Ph.D., 1954
Mathematics

Please Note: Find page numbered as 1-8 at the end of film copy.

University Microfilms, Inc., Ann Arbor, Michigan

THEORY OF NEURAL-ANALOG REINFORCEMENT SYSTEMS AND ITS APPLICATION TO THE BRAIN-MODEL PROBLEM

By
Marvin Lee Minsky

A D I S S E R T A T I O N

presented to the

Faculty of Princeton University

in candidacy for the Degree

of Doctor of Philosophy

Recommended for Acceptance by the
Department of Mathematics
December 1953

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

POOR COPY

Chapter 1

Introduction

The purpose of this paper is to develop a new approach to the "brain-model" problem. The following questions illustrate the most emportant problems in this field.

(1). What are the physiological processes through which animals perform such activities as "learning", "memorization", "recognition", "attention", "reasoning", etc.?

(2). How can these activities of so-called "sentient" organisms be duplicated in systems which we can actually construct? Is it possible to describe a system capable of behavior of humanoid complexity, yet simple enough in its physical structure that it can be understood? Can we describe such a system which is in addition, sufficiently resilient that its functioning can be maintained in the face of extensive injury, as in the case of the animal nervous system?

We begin by considering the properties of single neurons, and of simple sets of interconnected neurons ("nets"), with the objective of investigating phenomena which may be of importance in much larger nets and in the brain itself. The results of this analysis are then applied to the study of very large, "random", nets, and finally, to certain assemblies of very large nets. It is these assemblies which are our "brain models".

Each "brain model" is formed of a small number of very large "random" neural nets with a small number of channels connecting these nets; each connection is a large set of fibres which run between a pair of specified nets. In addition to these connections there must be a set of "output" or "motor" channels and a set of "input" or "sensory" channels which run between the brain model and its environment.

Now the "quality" of a brain can hardly be evaluated except in terms of the relation between it and its environment. Our initially disorganised "random" model must be able to raise itself from its initial chaotic state to a higher degree of internal organization, and this organisation must be measured in terms of the extent to which the brain can learn to deal with its environment. This latter capacity must itself be evaluated by some measure; for animals it seems natural to use as measure the ability to maintain the internal (physiological) state within some "normal" range, i.e., the ability to survive.

It will be shown that when our models are placed in an environment, and assigned (in a way described later) a range of "normal" internal states, that they will acquire a level of organization that can be compared only to that of the highest animals. Furthermore, this capacity for self-organization will in general, not be lost after injury, unless the injury is such as to change the gross topology of the system, e.g., if one of the basic nets is removed entirely, or if one of the gross connections is entirely destroyed. Thus these methods provide an approach to the problems of (2) above.

In addition, these models also provide an approach to the problems of (1) above. For the models are so constructed that they resemble the brain-net only in its higher organization, but also on the level of physical structure. Most of the properties of the "cells" of the theory are based on properties established for neurons in the experimental literature. A few of the properties may only be described as plausible. This situation is inevitable with the present state of information about the nervous system; while some new properties have to be assumed here, they have been made as simple, plausible, and few as seems possible. In every such case the assumptions are based on a plausible analogy with other biological situations. The geometry of the nets is also based on real biological data, in this case on the evidence of neuroanatomy. The viewpoint that the basic ingredients of the system can be taken as large unorganised "random" nets can be justified in several ways. The nets of the brain, for the most part, appear quite disorderly at the level of interconnections between cells. (There are exceptions to this, but they are usually confined to regions associated with certain special activities, and need not be considered in a theory of this generality.) There is no evidence of anything like the critically orderly connections of a modern computer. As the power of the microscope is reduced, order is perceived, and for the gross brain a pattern of a small number of discernable "regions" and distinguishable bundles of connections can be seen. The evidence provided by surgery and neuropathology support this picture. Also, it can be argued that the organization of biological structures in general cannot be too complex, without some process of self-organization; recent estimates (Quastler 1953, 263 ff) of the information carried by the genetic determiners may mean that tissues can be organized only along general plans; there is not enough information to determine many individual connections (unless it were done in regular patterns, which it is not.). Finally, as many of the results of this paper are to a high degree independent of the exact connection structure (on the "local" or "microscopic" level within a single random net), it is not necessary for us to specify this structure to any large extent.

In order to analyse the "behavior" of the brain models, it is necessary to introduce a number of "learning-theoretic" notions. In this paper, the most prominent of these ideas are those of "reinforcement operator" and "reinforcement process". Because these ideas are basic to the present analysis, and are, in addition, very closely related to a well-developed body of contemporary psychological theories, a separate chapter is devoted to this study. In a reinforcement process, the reactions of a system to external stimuli are originally a matter of chance. But the result, or immediate consequence, of each reaction is given a valuation, and this valuation determines the form of an operator which is applied to the system. If the valuation is "high", the effect is to raise the dependability of the associated reaction. Thus we have a sort of "trial and error" process.

It turns out that certain assemblies of random nets are capable of realizing this kind of process, if the valuation and reinforcement operator are controlled by an external "trainer". Then an important step is taken in showing that in assemblies of a very few nets, an "internal" or "secondary" reinforcement system can be made to evolve entirely within the net system; starting with a very simple primitive valuation system (such as is the basis of simple "reward-punishment" schemes of animal training) these assemblies reinforce themselves whenever any of a small distinguished set of stimuli occur. They learn to apply reinforcement also to behavior patterns which lead to the occurrence of these stimuli, as well, and can organize themselves to do this on higher and higher levels. It can then be seen that such systems, which initially have very little organization, evolve complex behavioral patterns which exploit the structure of their environment (or any environment which contains an appropriate degree of regularity) so as to force the occurrence of environmental events which have a high valuation in the reinforcement structure that has evolved within the system. Thus the system displays behavior which has, undisputably, the characteristics of both "goal-oriented" and "need-oriented" motivations. By relating the initial, primitive, valuation to the "internal physiological state" of the system, in such a way that reactions which bring this "state" toward its ~~x~~ "normal" value, the overall evolution of the system will be made to tend toward the establishment of behavioral patterns which are effective in satisfying the "physiological needs" of the system.

In additionato the highly developed reinforcement systems acquired by these assemblies, another process occurs which exhibits the features of what might be called "simple associative learning". (The system thus provides a model for theories of the "contiguity" group of contemporary theories of learning, and perhaps indicates how the controversy between the "reinforcement" and "contiguity" schools may have to be resolved.) Assemblies with no more than three or four nets are shown to have the capacity for organization into much more advanced activity; they are capable of "considering" alternative actions, making an estimate of the consequences of eahh alternative (using previously acquired information about the regularities of the environment) and preforming or rejecting actions on the basis of such an estimate. There is no evident limit to the degree of complexity of behavior that may be acquired by such a system. Their development will depend to a great extent on the environment (including here the physical body) in which it is embedded, the sensory and motor channels with which it is provided, and on the early experiences to which it is subjected.

valuation to the "internal physiological state" of the system, in such a way that reactions which bring this "state" toward its ~~s~~ "normal" value, the overall evolution of the system will be made to tend toward the establishment of behavioral patterns which are effective in satisfying the "physiological needs" of the system.

In additionate the highly developed reinforcement systems acquired by these assemblies, another process occurs which exhibits the features of what might be called "simple associative learning". (The system thus provides a model for theories of the "contiguity" group of contemporary theories of learning, and perhaps indicates how the controversy between the "reinforcement" and "contiguity" schools may have to be resolved.) Assemblies with no more than three or four nets are shown to have the capacity for organization into much more advanced activity; they are capable of "considering" alternative actions, making an estimate of the consequences of each alternative (using previously acquired information about the regularities of the environment) and preforming or rejecting actions on the basis of such an estimate. There is no evident limit to the degree of complexity of behavior that may be acquired by such a system. Their development will depend to a great extent on the environment (including here the physical body) in which it is embedded, the sensory and motor channels with which it is provided, and on the early experiences to which it is subjected.

The paper is organised as follows:

The present chapter is a general introduction.

Chapter 2 is devoted to the study of the logical structure of neural nets. The central question is that of what kinds of behavior can be obtained from nets which contain only cells which satisfy certain postulates. While some of the results obtained here are more general than is actually reqüired for the sequel, they throw some light on the question of neural inhibition. The biological data is so sparse in this area that it seemed appropriate to exploit the mathematical aspects of the problem, so as to best utilize the available information.

Chapter 3 discusses the neurophysiological basis for the systems of this theory. While there is very little information available about the properties of the cells of the central nervous system, what information exists strongly indicated that the nerve impulse mechanism is the same as for peripheral nerve. However, for reasons discussed in chapter 3, it is likely that there is a large "noise" component for activity at the interneural junctions. Accordingly, it is assumed that the properties of the junctions of our nets are like the excitability properties for peripheral nerve (which are well-established) except that a probabilistic uncertainty is attached to the classical notion of excitability "threshold". Thus the artificial axioms of chapter 2 are replaced by a set of biologically very plausible postulates.

Chapter 4 develops some of the learning-theoretic notions that will be required. "Reinforcement process" and "reinforcement operator" are defined, and a set of abstract "behavioral models" are examined. Each model has an abstract "environment" and we determine the extent to which each model can exploit the structure of its environment. The notion of a simple "global" reinforcement process (which resembles the systems of contemporary "reinforcement theories of learning") and a related notion of "local" reinforcement are contrasted, and it is seen that the "global" notion does not provide a natural description for complex mechanisms. Therefore the "local" concept is used henceforth. A machine, the SNARC, has been constructed which realizes a local reinforcement operator, and its structure and behavior is described.

Chapter 5 is an analysis of the activity of closed re-entrant "cycles" of neurons. It is shown that under the neural postulates of chapter 3, certain special forms of neural activity, called F-active patterns, will be distinguished, in the behavior of a random net, by their peculiar persistence properties. In the absence of "noise" (defined here as pulses arising from outside the F-active pattern) the lifetime of such a pattern is extremely long, but in the presence of "noise" they are peculiarly fragile, and are destroyed, created, and mutated rapidly. It is seen that this fragility in the presence of noise is not dependent on the structure of the underlying net. The linking and mutual interference of these patterns are discussed, and their growth and destruction examined. It is shown that if it is not assumed that all the cells of the net have the same properties (and they certainly do not in the brain) then the net may be regarded, from the viewpoint of the theory of F-activity, as composed of more or less distinct "interval spectrum domains" and that an F-active pattern must remain in one such domain. Two F-active patterns in different domains cannot both survive if they happen to intersect in the net.

The chapter contributes to the general developement of the theory of neural nets in that the "time quantization axiom", traditional in much of the work in this field is not used, and it is shown that

Acknowledgement

I am gratefully indebted to Dr. G.A.Miller for his encouragement and counsel from the time this work was first considered, and to Dr. A.W. Tucker for his encouragement, criticism, and for his deft removal of obstacles.

Much of the material in this work stems from extended discussions and arguments with Drs. John McCarthy, John Nash, John von Neumann, Claude E. Shannon, John Tukey, and David L. Yarmush.

In matters related to neurology, I am deeply indebted to Drs. E.G.Butler and Marcus Singer for superlative instruction and friendly advice, and to the generosity of Dr. John Welsh who provided me with laboratory space and biological materials when I was an undergraduate at Harvard.

I am deeply grateful to my wife, Gloria, for her devoted help in matters of style and organization, as well as in technical matters.

Certain of the biological viewpoints have been influenced by valuable conversations with Drs. K.Lashley, H.Quastler, P.Yakovlev and the late Dr. Nathan Savitsky.

I wish to acknowledge the less direct but equally important influences of A.M.Casson, H.A.Forrester and Dr. A.M.Gleason.

I wish to further acknowledge the friendship and tireless assistance of Dean S. Edmonds, without whose aid and advice the SNARC could not have been constructed. And I owe much to Drs. E.B. Newman and G.A.Miller for making that project possible.

It goes without saying that none of the above can be held in any way responsible for the innumerable weaknesses of this work; in particular, I must take full responsibility for the biological statements and conjectures.

This work was supported in part by the Logistic Research Project, Department of Mathematics, Princeton University, sponsored by the Office of Naval Research.

CHAPTER 2 FINITE AUTOMATIC NETWORKS

CHAPTER 2 FINITE AUTOMATIC NETWORKS

FINITE AUTOMATIC NETWORKS

2/1 Terms like "neural network" or "nerve net" are used at present to denote the subjects of a number of theories, each of which represents an abstraction of some of the knowledge derived from contemporary neurophysiological theory. In this chapter we define a few such objects and establish some theorems on the equivalence of certain sets of axioms for such theories.

2/1.1 S. C. Kleene has defined a FINITE AUTOMATON as a generalization of some present theories. Because we have a different emphasis his system is presented in a slightly different form:

A FINITE AUTOMATON is a collection of elements called "cells" whose operation is determined by the following axioms:

F-1: TIME is "quantized" as a sequence of discrete moments (indexed by the integers).

F-2: CELLS: There are a finite number of <u>cells</u>, each of which admits of one of a finite, $\geqslant 2$, number of states at any moment.

F-3a: Two kinds of cells are distinguished; INPUT CELLS and INNER CELLS.

F-3b: The state of an INNER CELL at a time t depends on the states of all cells at time t - 1.

F-3c: The state of an INPUT CELL at a time t is said to "depend on the environment".

F-3c means that the states of the "input" cells may be any function of time (of the integers), or be arbitrarily assigned by an "operator" of the automaton.[1]

2/1.2 The dependency relation of F-3 is entirely unspecified. There is one feature of the dependency relation that is common to all so-called "neural-network" theories. It is expressed by adding the following axiom to Kleene's system:

F-3n: There are a set of "CONNECTIONS" $\{C_{ij}\}$. A connection C_{ij} is said to "originate on cell C_i", and "terminate on cell C_j". There is a connection C_{ij} only for

---

[1] NOTE: Kleene restricts the input cells to take, at each moment, one of two states called 0 ("quiet") and 1 ("firing"). As he points out, one can always construct a logically equivalent finite automaton in which each cell has just two states, at the price of a uniform expansion of the time scale. However, the present theory is definitely not oriented in a logical-algebraic direction, and, replacing the axioms by a simpler equivalent system in which each cell has just two states would be an unnatural imposition.

certain pairs of cells (C_i, C_j). No connection terminates on any <u>input</u> cell.[1] The state of a cell C_j at a time t depends only on the states, at time t - 1, of those cells C_i for which there exists a connection C_{ij}.

<u>Def.</u> <u>A system which satisfies F-1, F-2, F-3, F-3n will be called a "FINITE AUTOMATIC NETWORK"</u>.

2/1.3 A notable example of a finite automatic network is provided by the system of McCulloch and Pitts (1943). Axioms for this system are present in a form consistent with those in 1.1 and 1.2.

MP-1: MP-1 is F-1, the time quantization axiom.

---

1 NOTE: F-3n states that no connection terminates on any input cell (which follows also from F-3c, if one regards a connection with no effect as vacuous). However, in the present theory, the "environment" of a given net will often be another net, and connections from the environment net will terminate on the input cells of the given net. The distinction between "input" and "inner" cells is to be regarded as a classification of a cell's position in a subnet, in relation to an observer's specification of which cells of a larger underlying net belong to the given subnet, and is not to be taken as a distinction between inherently different kinds of cells. At a later point, certain input cells will be designated as "receptor cells" (e.g., thermal receptors), and this designation will represent an inherent difference, or "specialization" of cells.

MP-2: MP-2 is F-2 with each cell restricted to two states, 0 "quiet" and 1 "firing".

MP-3: MP-3 is F-3n with the dependency law completely specified: The law can be given as follows (in a form arranged to match the F axioms):

Each connection C_{ij} has a numerical value which is either a positive integer or minus infinity. The value can be denoted by $\not{b}C_{ij}$. Let $\not{b}C_{ij} = 0$ in the case that there is no connection ~~to~~ C_{ij}.

Let $C_j(t)$ represent the function which has value 1 if C_j fires at time t and has value 0 if this is not the case. Let $C_j(t)$ also represent the proposition "C_j fires at time t".[1]

Finally, each cell C_j has a numerical "Threshold" $\#C_j$, which is a positive integer. The dependency law can then be stated as

$$C_j(t) \equiv \left[\sum_i C_i(t-1) \cdot \not{b}C_{ij} \geq \#C_j\right].$$

Note: In the MP theory, the cells are called "neurons".

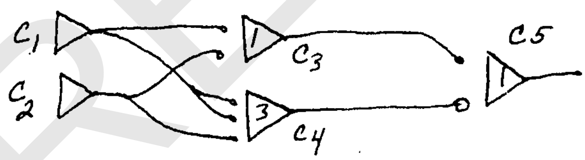
---

1 This convention will be used throughout this paper.

If $\emptyset C_{ij}$ is positive, we say that C_i has $\emptyset C_{ij}$ "endbulbs" on C_j. If $\emptyset C_{ij} = -\infty$, we say that C_i has an "inhibitory endbulb" on C_j. An example is provided to demonstrate the use of the network notation, and its description using the propositional calculus.

CELLS:	CONNECTIONS:
$\#C_3 = 1$	$\emptyset C_{13} = 1$
$\#C_4 = 3$	$\emptyset C_{14} = 2$
$\#C_5 = 1$	$\emptyset C_{24} = 1$
	$\emptyset C_{23} = 1$
	$\emptyset C_{35} = 1$
	$\emptyset C_{45} = -\infty$

DIAGRAM:



DEPENDENCY LAWS:

$$C_3(t) \equiv C_1(t-1) \vee C_2(t-1)$$

$$C_4(t) \equiv C_1(t-1) \cdot C_2(t-1)$$

$$C_5(t) \equiv C_3(t-1) \cdot \sim C_4(t-1)$$

THEOREM:

$$\therefore\ C_5(t+2) \equiv [C_1(t) \vee C_2(t)] \cdot \sim [C_1(t) \cdot C_2(t)]$$