

SOME TOPOLOGICAL ASPECTS OF MEMORY FORMATION
IN THE PROBABILISTIC NEURAL NET MODELLING OF
THE ASSOCIATION CORTEX

Maximos S. Papadopoulos

A Thesis
in
The Department
of
Physics

Presented in Partial Fulfillment of the Requirements
for the degree of Doctor of Philosophy at
Concordia University
Montréal, Québec, Canada

January 1980

© Maximos S. Papadopoulos, 1980.

ABSTRACT

SOME TOPOLOGICAL ASPECTS OF MEMORY FORMATION IN THE PROBABILISTIC NEURAL NET MODELING OF THE ASSOCIATION CORTEX

Maximos S. Papadopoulos, Ph.D.
Concordia University, 1980

The objective of this thesis is to investigate some cognitive aspects of the probabilistic neural net modeling of the association cortex. Our main focus is the mechanism of memory formation and the subsequent operation of this mechanism in the act of remembering. The association cortex is viewed as a randomly connected net that as a result of experience becomes structured. This structure is implied by two fundamental assumptions. The first assumption, due to Hebb, states that if two neurons are active together and one helps to fire the other, then adhesion of the two may occur. Neuronal assemblies are thus formed. The second assumption, due to Szentagothai, states that any two neurons in the association cortex may be connected via a small number of intermediate synaptic stations. This second assumption leads to the topological characterization of the structures constructed.

These structures represent, in mathematical terms, the physiological substrates of memory. In this thesis they are called memory domains M_i . It is found that as a result of the second assumption the Boolean overlap of two such domains is non-vanishing. It is also found that a given domain may go from a non-oscillating capability to an

oscillating one. Each such domain is found to require a threshold excitation before it can start reverberating mode. In neuronal terms this means that a minimum number of neurons within a domain must be simultaneously excited before reverberation sets in. If this minimum number belongs to the overlap of two domains the possibility of firing these two domains in close time contiguity arises. The possibility, in other words, of association emerges in this model. It is further found that the transition to a reverberating mode may be parametrized by a phase constant η , the minimum number of inputs required to fire a neuron in a given domain. A low η domain is a structure that can exhibit short term memory behaviour. In this dissertation we associate the structured domain with long term memory and we further associate the reverberating modes of this domain with short term memory. An event S is said to be remembered if it can excite through its make-up a given number of memory domains. This is possible provided its Boolean overlap with these domains is non-vanishing.

Utilizing the theory of sets we represent this by the formula

$$S = \bigcup_i (S \cap M_i).$$

ACKNOWLEDGEMENTS

The author would like to express his deepest thanks and appreciation to his advisor, Dr. P. Anninos, whose constant encouragement, guidance and active participation made this work possible. He would also like to thank the members of his examining committee for reading this thesis and offering their comments. In particular, he would like to thank Dr. Silvio Zenone for his comments on several parts of this work. The author would also like to express his appreciation to Dr. Peter Shizgal of the Psychology Department, Concordia University for his aid in neuropsychological matters, especially aspects of the Hebbian schema.

I would also like to express my appreciation to Miss Rose Elbaz for her excellent typing and to Mrs. Janice Konieczny-Bernath of Dawson College for her skillful graphics.

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF ILLUSTRATIONS	v
LIST OF SYMBOLS	ix
Chapter	
I. INTRODUCTION	1
II. SOME ASPECTS OF PRESENT DAY MEMORY RESEARCH	18
Part I: Phenomenology and the Statics of Memory	18
Part II: Some Models of Memory Processes	24
Part III: Process Aspects of Memory Creation	43
Part IV: Levels of Processing	57
III. NEUROPHYSIOLOGICAL PREREQUISITES	66
A. Neurons	66
B. Synapses	69
C. Information Processing Between Neurons	69
D. Neuronal Pools	71
E. Transmission and Processing of Signals in Neuronal Pools	76
IV. ASPECTS OF THE NEURAL BASIS OF MEMORY	89
A. Introduction	89
B. Neural Circuits of Memory	89
C. Reverberating Circuits	92
D. Consolidation	93
V. HEBB'S THEORY OF CELL ASSEMBLIES	101
A. Introduction	101
B. Mechanisms of Learning and Development	102
C. Structure and Genesis of Perceptual Learning	107
D. Divergent vs. Parallel Conduction: Attention	114
E. Achievements and Limitations of the Hebbian Schema	118
F. Conclusion	128

VI. THE PROBABILISTIC NEURAL NET: ANALYTICAL APPROACH . . .	129
A. Qualitative Aspects	129
B. Quantitative Aspects	132
C. Classification of Netlets	135
VII. THE PROBABILISTIC NEURAL NET: A SET THEORETICAL APPROACH	152
VIII. CALCULATIONS OF NET BEHAVIOUR	169
A. Introduction	169
B. Outline of the Simulation Program	169
C. Pattern of Completion	172
D. The Simple Conditioned Reflex	174
E. Topology of Memory Nets	174
REFERENCES	183

ILLUSTRATIONS

FIGURE		PAGE
1.	(a) Schematic of the connectivity matrix k_{ij} of a hypothetical neural net composed of netlets i, j A, B, and C. (b) Equivalent diagram in which arrows represent coupling between netlets	9
2.	Gross schematic of the memory system according to Atkinson and Shiffrin	27
3.	Detailed schematic of short term memory according to Atkinson and Shiffrin	31
4.	Assumed topological schematic of long term memory according to Shiffrin	38
5.	The recursive search algorithm according to Shiffrin	40
6.	Global schematic of the memory system together with its control processes according to Shiffrin and Atkinson . . .	42
7.	Example of a nodal structure: The node for the word 'tavern'	48
8.	Node for an event	50
9.	Examples of primary and secondary nodes	52
10.	Table showing different features of memory as they feature in the structural model of memory	58
11.	Effect of elaboration on recognition	62
12.	Neuron and synapse schematics	67
13.	Outline of motor and sensory transmitting areas of the cortex	73
14.	(a) Broadman and (b) Penfield and Rasmussen classification of the cerebral cortex	74
15.	Gross functional classification of the cortex according to Guyton	75
16.	Rough schematic of information flow in the cortex	77
17.	Schematic of a neuronal pool	78
18.	Possible structures of reverberating pools	80

FIGURE

PAGE

19. Pulse rates from different pools	82
20. Layering just below the pial surface	83
21. Progress of a signal through modules according to Eccles .	86
22. Triangle with macular field	111
23. Parallel and divergent transmission	116
24. Schematic of an attention process according to Hebb . . .	117
25. Examples of scanpaths	123
26. Typical decomposition of two stabilized images	127
27. Connectivity matrix after Harth and Edgar	133
28. Theoretical curves of the expectation value of the activity $\langle \alpha_{n+1} \rangle$ vs. preceeding activity α_n . Netlet parameters: $\mu^+ = 5$ $h = 0$. A and B are points of unstable and stable equilibrium respectively	136
29. Theoretical curves of $\langle \alpha_{n+1} \rangle$ vs. α_n for netlets with $h = 0.3$ and $\mu^+ = \mu^- = 10$	137
30. Theoretical curves of $\langle \alpha_{n+1} \rangle$ vs. α_n for netlets with $\mu^+ = 10$ and $h = 0$	138
31. Theoretical curves of $\langle \alpha_{n+1} \rangle$ vs. α_n for netlets with $h = 0$ and $\mu^+ = 30$	139
32. Universal curve with no inhibition	143
33. Universal curve with 30% inhibition	144
34. Universal curve with 20% inhibition	145
35. $\langle \alpha_{n+1} \rangle$ vs. α_n for a class B netlet receiving sustained excitatory input 6 (schematic). Note the shift of the three intersection points O, A, B with 45° line as a function of 6	147

36. Steady state activities for a class C ₁ netlet with sustained inputs. (a) Steady state values (corresponding to points O and B figure 35) are shown as solid lines, the metastable states (A in figure 35) as dotted lines. Arrows show direction of activity change for non-steady state values of activity. (b) Hysteresis loop generated by slow changes in θ , the fraction of active afferent fibers. Irreversible transitions are shown as dotted lines	149
37. Steady state activities for different netlets subject to sustained inputs. The afferent fibers are either all inhibitory (left side of diagram) or all excitatory (right side of diagram). (a) Steady state and metastable state values as a function of afferent activity. (b) Hysteresis loops corresponding to the curves given in (a).	150
38. Connectivity matrix with reinforced domains	155
39. Mapping of stimulus sets \vec{S}_a and \vec{S}_b onto \vec{a} and \vec{b} with $\vec{a} \cap \vec{b} \neq 0$	158
40. Set \vec{a} has 'learned' \vec{S}_a	159
41. Schematic showing the map of \vec{S}_a on \vec{b} . Note $\vec{a} \cap \vec{b} \neq 0$	160
42. Schematic of $\vec{S}_a \rightarrow \vec{a}$ and $\vec{S}_b \rightarrow \vec{b}$ with $\vec{a} \cap \vec{b} = 0$	161
43. Schematic of map of figure 42 with \vec{a} , \vec{b} reinforced	162
44. Pattern Completion	163
45. Connectivity matrix for binary association	166
46. Pattern Completion	183
47. Pattern Completion with Inhibition.	184
48. Firing of an Overlap near Threshold for Sustained Oscillation,	185

LIST OF SYMBOLS

- A The number of neurons in the net.
- r The refractory period.
- τ The synaptic delay.
- h The fraction of inhibitory neurons in the net.
- μ^+ Average number of axon branches emanating from an excitatory neuron.
- μ^- Average number of axon branches emanating from an inhibitory neuron.
- K^+ Average EPSP produced by an excitatory neuron in arbitrary units of amplitude.
- K^- Average IPSP produced by an inhibitory neuron in arbitrary units of amplitude.
- θ Firing threshold of the neurons.
- n The minimum number of EPSP's necessary to trigger a neuron.
- α_n The fraction of active neurons in the net at $t=nt$.
- σ The fraction of external incoming connections.
- M_i Memory domain i .

CHAPTER I

INTRODUCTION

A. Preliminary Remarks

The present work is an attempt to understand, via a particular model, some aspects of memory operation of our central nervous system (CNS for short). More precisely when we talk of neural modelling we mean as R. J. MacGregor and E. R. Lewis stated (1977) "The attempt to relate in nervous systems, function to structure on the basis of operation". It must be stated at the outset that a general solution to this problem does not exist. There does not exist a unique model that plausibly explains the totality of the functions of the CNS. Partial successes do happen especially when pre-wired, genetically determined circuits, are involved. The existence within the mammalian cortex of such anatomically diverse structures as the cerebral cortex, thalamic nuclei, hippocampus and reticular formation does not make one's task any easier. One approach is to simulate only one part of these structures and by adjustment of various parameters try to come down to characteristics that are realistic. Artificial neural nets, usually simulated on computers, are one of the most powerful tools at our disposal.

B. Some Terminological Preliminaries

The diversity of neural structures has, as we implied above led to a great number of different models each of which purported to show correlations to this or that aspect of neural activity. Depending on the background of the researcher one had models whose inspiration came

from an engineering, information, cybernetic or some other such matrix. Csermely (1968), for example, mentions four different classes of such models. For the purposes of this work we shall follow a newer classification due to MacGregor and Lewis (1977). This will, as we shall see, further down, make our approach to the problem more comprehensible. MacGregor and Lewis use the work "model" in at least three different senses. By conceptual models (or C-models) they mean something very much like a theory based on a set of concepts and relationships among these concepts which could provide explanations for phenomena under their purview. The Hebbian schema of memory formation would thus be a C-model. A general purpose computational model (GP-model) is a computational instrument for spelling out dynamical properties of C-models. Thus a computer simulation model based on a C-model would be a GP-model. The special computational model (SC-model) is also a GP-model. The special computational model (SC-model) is a GP-model in the sense that it represents a concrete realization of a C-model. For example the formalism of differential equations could be the GP-model for Maxwell's equations whereas the specific class of differential equations of the form

$$F = m \frac{d^2r}{dt^2}$$

would represent the SC-model for Newtonian Mechanics. In other words a GP-model might correspond to more than one C-model. The contrary can also happen. Two GP-models could correspond to one C-model. This is the case in Quantum Mechanics where Schrodinger's Wave Equation corresponds to one and Born's matrix Mechanics corresponds to another GP-model of the same C-model. The value of computational models is

that they render C-models amenable to experimental verification and elaboration. Thus the energy levels of the Hydrogen atom are a direct result of the Schrodinger Equations and cannot directly be derived from the axiomatic foundations of quantum mechanics. It cannot however be overemphasized that no computational model can exist without a meaningful C-model. Thus C-models are necessary but not sufficient conditions for the existence of GP-models.

C. Anatomical Considerations in Neural Net Modeling

The modeler's task would have been easier had it not been for the diversity of structure one finds in the brain. As Szentagothai and Arbib have stated (1975): "When looking at any piece of central nervous tissue, the beholder is immediately aware of two conflicting aspects and interpretations of structure. On the one hand, the long dendritic arborizations and their frequent intertwinement with widespread, highly irregular and diffuse terminal axon arborizations would suggest a random connectivity between the several pre and postsynaptic elements participating in any piece of neuropil. This is particularly striking in the brainstem reticular formation..... On the other hand..... other..... investigations, seem to indicate that there is an extremely strict specificity, not only with respect to what kind of neuron is contacted by any given type of axon but also concerning the exact site where the synapse is established". In addition, the existence of a relatively great number of global structures in the brain produces another dimension of difficulty.

Further, as Anninos has pointed out (1975) another consideration from the point of view of the modeler is the intriguing question as to

the type of constraint imposed by a given global structure on the connectivity pattern of its constituent parts. As an example he takes the connectivity pattern where the majority of the efferent processes point out of the given nucleus. The conclusion he draws is that apart from field effects no "cross-talk" can occur between neurons of this nucleus. He then goes on to outline a number of principles of interconnectivity. Since more than one structure exists in the CNS no unique model is possible based on a fixed parametrization of its structural components. Of particular interest to this work is the apparent lack of function-related structures in the association cortex. As Anninos points out, the cerebral cortex in general allows a greater degree in connections between neighbouring cells than between distant ones. In other words the probability of connecting with one's neighbouring cells would be greater than with more distant elements.

D. General Models of Nervous Systems

MacGregor and Lewis (1977, p. 297) divide the various existing techniques of neural modeling of large networks into two classes. The first class is the analytic and/or field technique whereas the second class lends itself to simulation by large-scale, high-speed computers. It should be pointed out however, that as we mentioned earlier one could have two distinct GP-models corresponding to the same C-model. This is the case with our efforts where consistency between the simulation model and the analytico-probabilistic approach has been shown by various workers. The germinal paper for the analytic techniques is a paper by Beurle (1956). The classical paper for computer oriented approaches is the McCulloch and Pitts paper of 1943.

Beurle's approach is to approximate the discrete distributions of neurons by a continuous one. Thus neurons were assumed to be distributed continuously with a volumetric density ζ . The axonal and dendrite fields were specified by a connectional density μ . Variations were assumed to occur in the x-direction only and no inhibition was considered and $\mu(x)$ was assumed to be of an exponential form. In other words near neighbours were favoured. The cells had threshold, synaptic delay, excitatory postsynaptic potential and membrane time constant. The proportion of cells sensitive to excitation at a given time is considered via the basic field variable $F(x,t)$. The proportion becoming insensitive or refractory per unit time constant is also considered. A switching effect is demonstrated and under certain constraints the emergence of distortion free waves of neural excitation is shown. Anninos (1975) compares this switching effect to the switching of attention of a given organism. The emergence of wave solutions as MacGregor and Lewis (1977) state is the result of the basic assumptions, which stated in a differential form read:

$$dF/dt = R\phi - F \quad (a)$$

$$dR/dt = -F. \quad (b)$$

Here R is the number of cells that are sensitive to firing at a given space-time coordinate (x,t) . ϕ represents the probability that a given cell will be excited above threshold in unit time. The first equation states that the rate of change of the number of cells that become active at a given time is equal to the number of cells sensitive to firing at that time minus the ones that are already active at this same time. The second equation is simply a bookkeeping statement to

the effect that the rate of change of the number of cells that become sensitive in a given time is the negative of the number of cells active at the same time. Combining (a) and (b) one obtains

$$\frac{d^2 R}{dt^2} + (1 - R \frac{\phi}{F}) \frac{dR}{dt} = 0. \quad (c)$$

(c) is a very difficult equation to solve since ϕ is a function of F , x , t and implicitly it involves space differentials making the equation a partial differential equation. According to Beurle if $\frac{\phi}{F} = m$, (a constant) then one obtains the travelling wave solution of

$$R = \frac{1}{m} - \frac{2B}{m} \tanh B(t-t_0)$$

$$F = \frac{2B}{m} \frac{1}{\cosh^2 B(t-t_0)}$$

where B is a propagation constant determining the rise and fall of activity. Griffith (1963, 1965), then Hoopen (1965) refined Beurle's formulation so that the net was now capable of sustained oscillation - a property warranted by neurophysiological considerations.

Wilson and Cowan (1972) have also presented a formulation in terms of field quantities building largely on the work and the methodology of Beurle (MacGregor and Lewis, 1977).

They derived coupled nonlinear differential equations to describe the dynamics of spatially localized populations of both excitatory and inhibitory neurons. These equations were used to investigate the responses of neural population to various stimuli. Of interest to our work are their results showing that under certain conditions neural populations are capable of exhibiting both simple and multiple

hysteresis phenomena the putative substrates of memory organization. A critical role in the emergence of hysteresis phenomena was played by their assumption that inhibition arises from exclusively inhibitory neurons. This is known as Dale's law. (Eccles, 1964).

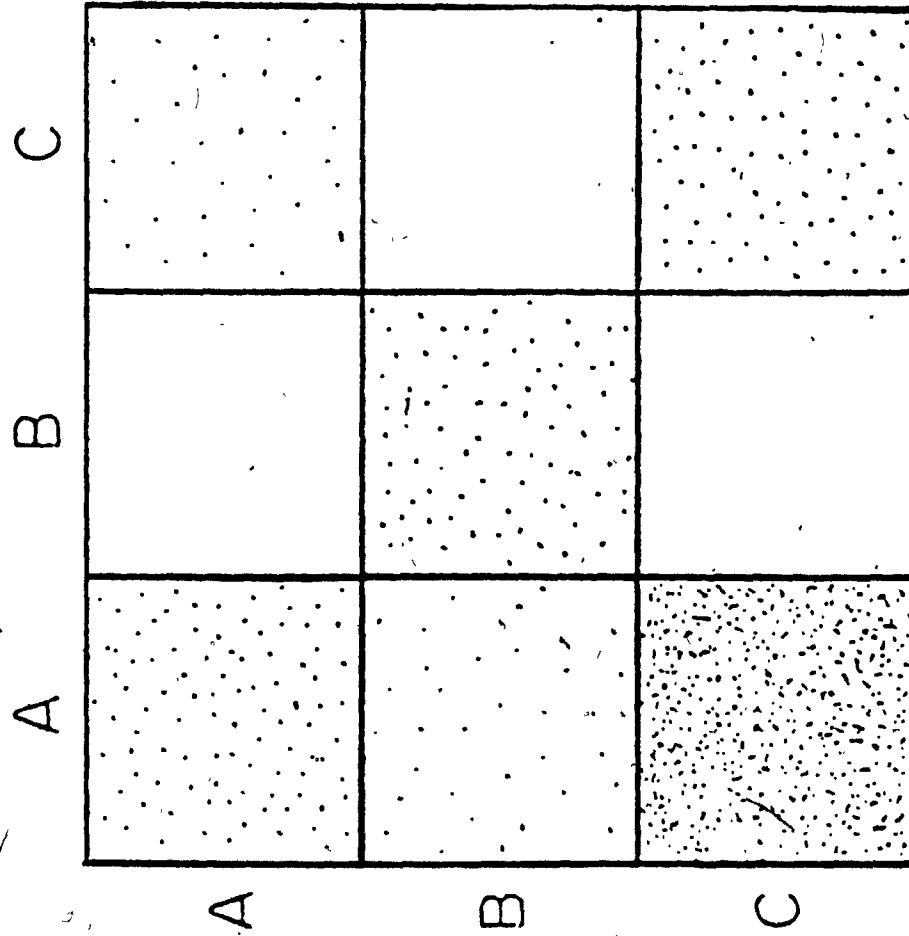
For further details on the field theoretical approach to neuronal population phenomena the reader is referred to the recent and excellent review of MacGregor and Lewis (1977). Another avenue of the analytic technique is that based on probability considerations. This approach usually considers neuronal populations not as a continuum but as a set of elements connected according to some connectivity pattern and other anatomically inspired considerations. The first to consider the dynamics of discrete probabilistic nets was Rapoport (1952) who showed that such nets are capable of accepting and subsequently propagating an excitation through the net provided initial activity exceeded some threshold value. Subsequently sustained oscillation would set in. Otherwise activity would die down. Trucco (1952) extended this work by providing conditions under which the net would exhibit ignition phenomena.

In a mathematically extensive paper, Allanson (1956) investigated the dependence of net dynamics on various parameters of randomly connected nets. Depending on the parameters, such nets could oscillate continuously, exhibit damped oscillations or remain quiescent.

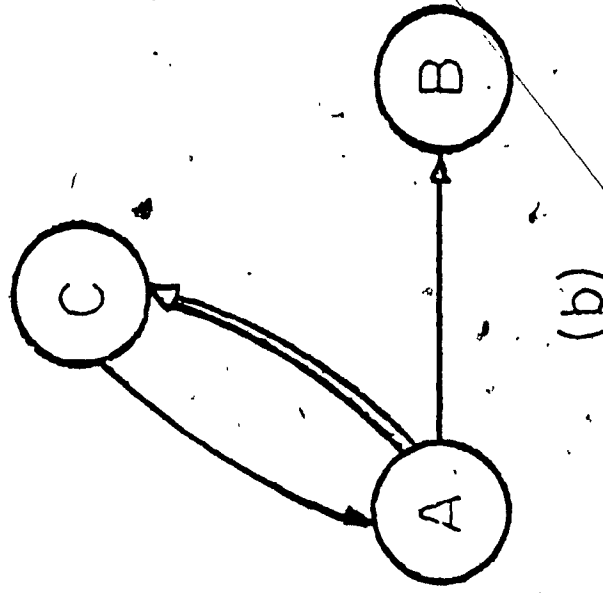
Along probabilistic lines one can also classify the analytical work of the Syracuse School led by Harth (1970). Its bases have been stated in several places (Anninos et al 1970); Csermely (1968);

Anninos and Ellul, (1971, 1974). The two fundamental assumptions are in brief: First, that the structure of the nervous system may be approximated by sets of discrete populations of randomly interconnected neurons, which are called netlets. The netlets are coupled to one another in a way which might be described as structure-in-the-large but randomness-in-the small. The second assumption concerns the importance of various dynamical (time-dependent) variables that might be derivable from the function of such a model. The dynamic variable of neural significance is taken to be the level of activity, i.e. the fractional number of neurons firing in each netlet. Details of all these assumptions as well as some of their implications may be found in Harth et al (1970) and Anninos et al (1970).

Harth and Edgar (1967) used a schematic, called the connectivity matrix $\{K_{ij}\}$ to graphically illustrate how the netlets are connected. The $\{K_{ij}\}$ is defined as follows: let K_{ij} be the change in the membrane potential of neuron i due to efferent inputs from neuron j . The quantity is referred to as the PSP. K_{ij} in other words represent in arbitrary units the net effect of neuron j on neuron i . This quantity is also called the coupling coefficient from the j th to the i th neuron. A coupling coefficient of zero implies no synaptic link between the two neurons. In general, the coupling matrix of Figure 1 (see 1a) is asymmetric. Figure 1(a) shows a representation of a hypothetical neural net. Each dot represents a non-vanishing element of $\{K_{ij}\}$. The net shown represents three netlets, labeled A, B and C. The diagonal blocks show the internal connectivities of these netlets. The randomness in the distribution of these points represents the



(a)



(b)

Figure 1(a). Schematic of connectivity matrix K_{ij} of a hypothetical neural net composed of netlets A, B, and C. (b) Equivalent diagram in which arrows represent coupling between netlets.

randomness-in-the small aspect of the model. The off-diagonal blocks represent the couplings between the netlets. The global feedback and servomechanism aspects of this model is shown in another diagram, (figure 1b). In this model the netlets were used as building blocks to simulate the design of known areas of the brain. It was thought, for example, that the neuron colonies observed by Anasuma and Sakata (1967) in the motor-sensory cortex and the columnar structures found by Hubel and Wiesel (1965) in the visual cortex of the cat were the neural substrates corresponding to these netlets. Mountcastle (1957) called them "elementary units of organization." Szentagothai in a series of papers (1969, 1972, 1973, 1974, 1975) has extended his idea of neuronal modules to the association cortex itself; a fact which is of importance to this work.

The purely analytic aspects of the model, again in summary, are as follows: (They will be restated more fully in the context and to the extent that they are of need to our work later).

One considers a set of A neurons a fraction μ of which are inhibitory. Time is quantized in terms of synaptic time delays τ . During each time interval a neuron is either "on" or "off". If the cell is "on" it can affect some other cell (provided it is connected to it) by an amount K the PSP. μ^+ , μ^- , represent respectively the number of outgoing excitatory and inhibitory connections. If the sum of excitatory and inhibitory connections exceed some threshold θ the affected cell fires on τ later. During the $\tau + 1$ interval this same cell may not fire due to refractoriness. No space or time summation is

is taken into account and there is no relative refractory period. The main dynamical variable of the model is α_n the fraction of cells activated at time τ . One proceeds on the basis of the above assumptions and derives an expression for the expectation value of activity at time $(n - 1)$ in terms of activity in time n . This is found to be

$$\begin{aligned} \langle \alpha_{n-1} \rangle \approx & (1 - \alpha_n) \cdot \exp(-\alpha_n h \mu^-) \sum_{m=0}^M (\alpha_n h \mu^-)^m / m! \\ & \times \{1 - \exp(-\alpha_n (1 - h) \mu^-)\} \sum_{\ell=0}^{n-1} (\alpha_n (1 - h) \mu^-)^\ell / \ell! \end{aligned}$$

Here ℓ is the number of excitatory postsynaptic potentials, m the number of inhibitory postsynaptic potentials and n the minimum number of excitatory inputs required to trigger a neuron which has just received an inhibitory input. Details of this derivation may be found in Anninos et al (1970). The basic behaviour of this equation is shown in figure 28.

The family of curves shown exhibit positive slopes for small values of α_n and negative for large values of α_n . Positive slope means that the net increases its excitation during the next time unit. Negative slope indicates excitation is dying down, a result to be expected due to refractoriness especially since the areas of negative slope occur when $\alpha_n \sim 0.5$. The parameter n which measures in the minimum number of excitatory inputs necessary to trigger a neuron in the absence of inhibitory inputs may be taken as a phase constant indicating the excitability properties of a given net. Under certain circumstances it will be shown in this paper that nets with high n 's become nets with low n 's as a result of synaptic facilitation.

More details about the working of the net will be given in chapter five.

As we mentioned earlier Harth uses two mutually consistent SC-models to work out the implications of his C-model. A very similar approach was followed by Smith and Davidson (1962) who following the work of Allanson (1956) attempted a probabilistic approach in their analysis of simulated neuronal networks. Both analysis and simulation indicate an ability of these nets for self sustained activity. The different types of steady-state and oscillatory activity that these networks could sustain were shown to be related to such parameters as network connectivity, latent summation, refractory period and threshold. Different levels of steady-state behaviour were found for a given net.

Further work on sustained activity and the importance of inhibition was undertaken by Griffith (1963, 1965) who showed that stable intermediate activity is possible either because one has a large scale aggregate of richly interconnected cells and/or because inhibitory connections make the intermediate stable values possible.

The first SC-model based on a digital computer seems to have been that due to Rochester et al (1956). Using an IBM 704 digital computer they simulated a probabilistic model of 69 neurons. Their C-model was the Hebbian schema only two aspects of which were included: excitatory connections and synaptic facilitation. They expected formation of reverberating nets, the so-called cell assemblies. Contrary to their expectations the total net soon reverberated and not just some determinate subset of this net. This periodic spreading of activity they called diffuse reverberations. In a second attempt, 512 neurons and inhibition were introduced. A 'distance bias' as to

connectivity pattern was also used. Thus near neighbours were more likely to be connected than distant ones. Synaptic facilitation was modified so that the sum of synaptic strengths for the entire net should remain fixed. Four receptor areas of four neurons each were designated in the net. Cell assemblies were now observed to be formed around the four receptor areas.

Farley and Clarke (1961) simulated a planar net of 36×36 neurons whose interconnections were specified by a two-dimensional probability distribution with a distance bias. Spatial and temporal summation but no inhibition were included. Whenever an element fired, its threshold was immediately set at a very high value and was subsequently allowed to decay exponentially on its resting value. An initial activity was given to the net and the resulting activity was noted in 0.5-msec intervals. Again one was looking at firing-patterns as functions of network parameters. Their results agreed with those of Beurle (1956). Interestingly enough their results also showed that under certain circumstances, such as repetitive stimulation and low threshold the net would exhibit sustained oscillations in spite of the absence of inhibitions. Farley and Clarke used in addition, another probability distribution in which all neurons had an equal probability of interconnection. Diffuse oscillations were observed that would either stop spontaneously or continue indefinitely. The physiological correlations to brain functioning are obvious. The lack of inhibition in the model is its negative feature.

E. Outline and Methodology of This Thesis

In this dissertation an attempt will be made to correlate neural net behaviour to cognitive behaviour. The basic results of this work

will be found in Papadopoulos et al (1979). This thesis may be considered as a continuation of those aspects of the work of the Syracuse School workers that directly bear on learning, memory and cognition. The associative properties of the probabilistic neural net model can be found in Harth and Edgar (1967). The basic limitation of their work is the non existence of inhibition in the net. Csermely's doctoral dissertation (1968) gave us an insight into various types of conditioning whereas Anninos' results concerning hysteresis behaviour is the first clear indication that probabilistic neural nets are capable of exhibiting phase changes as a result of learning. Our work, in this dissertation, has attempted to trace aspects of memory processes such as memory search and recognition to an underlying mechanism which in psychological terminology could be best described as generalized gestalt completion. At the neuron pool level we assume that two assemblies, for example, owing to their interconnectivity never really have a vanishing boolean overlap. Learning has the effect of increasing this boolean overlap. Consequently, triggering one assembly tends to trigger the other assembly as well. We explain this by saying that in some cases, boolean overlap, has the effect of triggering the second assembly via gestalt completion. The existence of a boolean overlap is shown to be a result of the assumption that any two neurons in the brain may be connected via finitely small numbers of synaptic relays.

Looking at correlations between our model and overt behavioural acts, as well as underlying neurological facts could be considered as an attempt to fulfill the admonitions of MacGregor and Lewis (1977)

concerning desirable attributes of a good model. This writer does not fully agree with MacGregor and Lewis with regard to his insistence that a falsifiability criterion should be included in the model. This is the well known Popperian view (1959) with which few practising scientists would agree. The Popperian precept could be shown to have some value for mature and theoretically stabilized theories. It is this writer's belief that theories trying to encompass new ground should first look for a system of concepts that could parsimoniously cover the ground in a logically consistent fashion and in such a way that they could also lead to new experimental ground. This has, certainly been the tradition in physics.

According to MacGregor and Lewis (1977, p. 354) the substantive content of brain modeling should be guided by a principle of stratification of variables. "The main assumptions in this approach are that the variables, constructs and knowledge which define the problem context of brain modeling arrange themselves naturally (or can be arranged) into strata such that individual strata identify highly interrelated concepts, and causal (or at least meaningful, and more or less direct) relationships can be identified or sought between adjacent strata". The main difficulty in establishing entirely satisfactory stratification at this stage of brain modeling seems to be the relatively incomplete development of the neurosciences and the psychological sciences. From the various models possible those models are considered good that relate adjacent strata. We think that this investigation represents such an effort. It is based on two conceptual models. The first C-model is a composite construct that emerges from

an attempt to find common implications among the various memory theories now in fashion. We undertook an investigation of some psychological theories of memory with the express purpose of tracking down this common ground. This is the subject matter of the second chapter. The third chapter presents ~~some~~ neurophysiological data which are necessary for subsequent discussion. The fifth chapter is a presentation of another C-model, namely the Hebbian Theory which may be characterized as one of the most powerful schemata whose main purpose is to correlate psychologically overt phenomena such as memory and attention to their putative neuronal substrates, the cell assemblies. The ontogenesis of cell assemblies plays a very important role in this theory. It also serves to make the connection from a higher level of integration to a lower one! As Freeman (1972) has very perspicaciously pointed out in another regard, this "search must undertake the description of neural masses in terms peculiar to their own properties, arising out of but distinct from the properties of single neurons. The concluding pages of the chapter on Hebb's Theory show where both Hebb and the memory theorists could be pointing from the point of view of the topology of cell assemblies or memory domains.

The Syracuse model may be considered in part as a special computational model based on Hebb's Theory. The next chapter of this work is a brief presentation of the Syracuse model, a set theoretic adaptation of which has been used to present some of its cognitive aspects. In another chapter we continue with the cognitive implications of this model and we present quantitative evidence to the effect that overlap and overlap growth among memory domains (or cell assemblies) is crucial to an understanding of the phenomena related to memory search as defined

by Shiffrin and other psychologists in the second chapter. Consistency is thus indicated between the various dimensions of analysis so far. We then continue on the basis of our schema to an hypothesis concerning the solution to the problem of the existence of more than one memory. We suggest that different memories may be identified with phase changes of the same neural mass under coherent excitation just as eigenstates are due to different excitations from the ground state of a given quantum system. Finally areas where this work could fruitfully be pursued further are indicated.

In **summary** this theory is a development of the work of the Syracuse-Concordia group in the direction of cognitive function studies. The new analytical tool developed is a topological category by the name of memory domain. This is a non metric topological space composed of fuzzy sets corresponding to neuronal pools. The dynamics and interaction of these domains have been studied via both analytic and simulation methods and conclusions have been reached regarding their cognitive importance. Our level of analysis is situated at the neuronal pool level as can be seen in the chart in the appendix. Our effort has shown that one can correlate the memory domain category to both the current theories and models of memory and to pertinent neurophysiological findings. In other words this work is both a consistency argument correlating various theoretical constructs and an original analytical development of the formalism of the Syracuse-Concordia School.

CHAPTER II

SOME ASPECTS OF PRESENT DAY MEMORY RESEARCH

PART I: PHENOMENOLOGY AND THE STATICS OF MEMORYA. Introduction

Present day studies in memory can be classified under at least three headings: (a) phenomenological aspects (b) statics and dynamics of systems that result in behaviour covered under (a) and finally, (c); investigations at the neurological level whose purpose is to find neural structures correlating with (a) and (b). The study of memorial processes has a history beginning with Aristotle. The literature is vast, growing and contentious. There exists a very large amount of information concerning various experiments purporting to test this or that aspect of memory. We do not propose, here, to cover or to review the whole field of memory research or, even less, to give complete account of the various issues involved. This is certainly beyond this writer's competence. Rather, what we have in mind is to first present some of the phenomenology that gives rise to the contention that a memory structure exists. Secondly enough detail of processes and putative structures giving rise to them will be given, so that our own work on the topology of memory structures can comfortably be interfaced or imbedded in the current literature. Our account is bound to be uneven for in some cases only enough detail will be given to make the account coherent. In other places we intend to go into considerable detail for it is in these areas where we must find

correspondence with our own work. Further beyond this and the following chapter where most of the current non neuropsychological work will be presented, we will still have to refer to various other results in other chapters.

B. The Behavioural Context of Memory Research

Simon (1976) in a review article, on the current status of memory research mentions eight parameters that he considers useful in parametrizing the properties of various memory units. According to him memory can be characterized and identified by, (1) the kinds of inputs that can be stored in a given system, (2) the time required to store new information in a given store, (3) the time required to access information residing in a store, (4) amount of time a given piece of information can reside in a given store, (5) the type of processes and the conditions leading to loss or inaccessibility of a given informational unit from a given structure, (6) qualitative nature of the deterioration of stored information, (7) the nature of the cues needed for accessing stored information, and (8) the form of organization of the stored information.

Two things can be stated about the above. First, Simon is abstracting from current research on memory and he therefore uses parameters that various researchers used so far. This does not mean that these are the most appropriate or that some are not more fundamental than others. Secondly, this parametrization already presupposes a memory structure of a particular kind and in some aspects it already forces the issue. (See for example the controversy over the existence of STS and the arguments of Craik and Lockhart (1972) in this regard).

Nevertheless, we feel Simon's arguments are quite useful at least from the point of view of limiting our universe of discourse to only those issues that are of interest to this work.

Simon bases his discussion on four fundamental experimental paradigms. Most current theories of memory processes, as a matter of fact, try to build structures whose explanatory area includes an explanation of these paradigms.

Paradigm 1 (Immediate Recall Experiment)

A 3 x 4 matrix made up of english letters is shown to subjects (Ss for short) tachistoscopically for 50 msec. Ss are then instructed to report what they have seen as soon as stimulus is off. Typically Ss can recall four to six letters correctly. (Sperling 1960; Simon 1974).

Paradigm 2 (Partial Report Experiment)

One uses the same experimental set up as in paradigm 1 except that now tonal cues are used to instruct Ss to report contents of a given row. Results show that for cues whose onset is near stimulus offset there is almost total recall. As the interval between stimulus offset and cue onset lengthens to about one or two seconds the results of the partial report procedure approach those of paradigm 1 (Sperling 1960).

Paradigm 3 (Delayed-Recall Experiment)

Set up is in principle like that of paradigm 1. Ss are, however, upon stimulus offset, instructed to do a distracting task such as counting backwards for about 30 seconds. There is now a definite drop in performance;

→ Ss recall about 1 or 2 items in a list of several letters. (Peterson and Peterson, 1959).

Paradigm 4 (Rote-Verbal Learning Experiment)

Ss now have as much time as they want to memorize. After a lapse of time following learning Ss are instructed to recall. Performance can be up to criterion for delays of up to several hours (Underwood 1969).

Some inferences from the Fundamental Paradigms

The most fundamental conclusion one reaches is the conclusion that memory exists since Ss can recall a given stimulus after its offset.

The partial report paradigm implies that there exists memory lasting for about half a second which contains all twelve letters of the matrix. One is led to this conclusion because Ss, during this time interval of half a second, can scan this memory store and report contents of any given cued row. Furthermore, considerable information is lost after one second.

The Immediate-Recall experiment, which resembles the behaviour of the later stages of the Partial Report Procedure, tells us that a small amount of information can be retained for a few seconds.

On the basis of these two experiments alone Simon (1976) points out that one is justified in drawing either of the two following conclusions:

Option 1: One could have a unitary memory store with capacity of twelve letters, whose forgetting rate would be very fast in the

beginning and slower in the end.

Option 2: One could have two distinct memory stores with unequal forgetting rates.

The type of errors Ss make in reporting stimuli in either of the two paradigms seem to settle the question in favour of option 2. It is found that confusions in the Immediate Recall experiment are auditory; (Sperling, 1963). Thus one reports seeing a V ("vee") instead of a Z ("zee"). In the partial report procedure confusions are of a visual character. (Sperling, 1960). These confusions are taken to imply that one is dealing with two types of stores. (Reynolds and Flagg, 1977). The memory involved in the partial report experiment is called iconic whereas the Immediate Recall memory is called Short Term Memory. In general we can conclude, (Simon, 1974) that the memory involved in the second paradigm contained information of a visual nature (the form characteristics of the stimulus) whereas the memory involved in the Immediate Recall experiment contained information about phoneme features of the stimulus. A recoding process is assumed to occur in transferring information from the iconic to the Short Term Memory.

Before leaving the above considerations we should mention the fact that with regard to the nature of the Short Term Memory vis-a-vis iconic storage and their putative differences one has data that could cloud Simon's picture considerably. Thus Reynolds and Flagg (1977) citing results due to Phillips (1974) show that one could recode information not into an aural form but into a visual form again. Phillips' data seem to indicate that whereas the icon is of large capacity and short duration there exists a "longer lasting" short term memory of smaller capacity and of larger duration, up to 9 sec or more.

Further, although iconic storage is strongly affected by the relative position of the letters, visual short term memory is not.

For the sake of completeness we should also mention that each sensory modality has its own iconic memory. Thus one talks about an echoic storage that presumably holds auditory sensory information before it is processed elsewhere. The generic name for all these memories is sensory information storage. (SIS for short).

If one takes an information processing point of view, (Norman, 1972), one can describe the above as follows:

The job of extracting the features of a sensory message in order to determine what the given message means may take more time than the duration of the actual physical signal. The sensory information storage system plays the logical role of holding the message for a sufficient time for this analysis to occur. Not only does the SIS seem to retain a good image of events that have occurred during the past few tenths of a second but as the Immediate Recall paradigm indicates there is more information there that can be extracted. This discrepancy between the capacity of the SIS and the amount of information that can be processed by later stages implies a limit on the rate of processing information from SIS to STM.

Comparison of the results of either the immediate - or delayed - recall experiments seem to indicate that there exists yet another memory. This memory created by stimuli of relatively long duration seems to survive all sorts of distracting tasks. Further no capacity limitations seems to be observed. We see this from the fact that memory exists in the rote-verbal learning experiment, that goes beyond the limits of the

memory of delayed-recall experiment. This memory is known as the Long Term Memory, (LTM for short).

PART II: SOME MODELS OF MEMORY PROCESSES

A. Introduction

As said above, most current theories of memory processes attempt to build systems which are consistent with the four fundamental paradigms. The usual practice is to build up a model incorporating features and/or parameters the modeler considers relevant. One model that more or less encompasses most of the accepted views of memory and one which has been quite influential in the last few years is that of Atkinson and Shiffrin (1968). A companion piece was written in 1969. (Shiffrin and Atkinson, 1969). In brief this model of the human memory system focuses on two aspects of memory: memory structures and control processes. Memory structures were defined as "both the physical system and the built-in processes that are unvarying and fixed from one situation to another." Control processes were defined as processes that "are selected, constructed, and used at the option of the subject and may vary dramatically from one task to another" (1968, p. 90). Schneider and Shiffrin (1977) further refined this model in 1977. We shall also have occasion to refer to this work as well, but only to the extent that it pertains to our discussion and position. Let us now take a closer look at the Atkinson and Shiffrin model.

These authors view the memory system along two basic dimensions. In one of the dimensions one has the structural features and the control processes that the subject can vary at will. The permanent features of memory which they call the memory structure included, as we said above, both physical system and the built in processes that are fixed.

"Control processes on the other hand, are selected and utilized to suit the wish of the subject. Different control strategies may be used for different situations". Atkinson and Shiffrin say that if one considers our memory system as a large computer system under the direction of a programmer, then the permanent memory, computer hardware and library functions of the system would correspond to the permanent features of memory. The programs however that the programmer can write at his console, as well as any instructions he might initiate thereon in order to control computer processing would be analogous to the control processes of the human memorial structure.

B. Structural Aspects

A second dimension of analysis, in this scheme is the division of memory into the three structural components mentioned in the beginning of this chapter. It should be pointed out at the outset, however, that Atkinson and Shiffrin have repeatedly stated (e.g. in 1969) that this division of memory into three stores does not imply that one has three distinct physiological entities that serve the purpose of being the substrates to these functions. As early as 1969 they made the proviso that STM could be a temporary activation of information stores in the LTM. They thus view memory as being made up of the sensory registers, the short term store and the long term store. Incoming sensory information enters the sensory store selected outputs of which proceed to the short term store which is the Ss working memory and which in addition must receive inputs from the long term store. Information in the STS decays completely within a period of thirty seconds unless a control process known as rehearsal maintains it there. The long term store is the repository of permanent information. The long term store "copies"

selected information from the STS. Information in LTS is of a more or less permanent duration. The basic structural division is as shown in figure 2. It should be pointed out that the form, character and coding of information in STS does not depend, necessarily, upon its form in the sensory input - if it came from there. Thus a work presented visually in the sensory register might be coded in an auditory form when it is acquired by the STS. Atkinson and Shiffrin use mainly the auditory mode of the STS which they term the audio-visual-linguistic mode. They do this because the various confusions that are contingent upon it are rather difficult to ascribe to any particular fact of this mode. They further maintain that the flow of information from store to store is under the control of the subject. By control of information flow they mean in the first place a selection of information to be handled and secondly the decision as to whether a given message or bit ought to be copied from one store to another. Information remains in a given store during and after the copying process and decays therein according to the decay rate of the particular store. After input to the sensory register the subject initiates a control process that transfers information to the STS. In certain parts of their paper they seem to imply that this process is accomplished by first scanning the SR and relating what one finds there to information in the LTS. In other instances they imply a direct, transfer. For various reasons having to do with this writer's understanding of perception we believe that it is unlikely that information enters the STS directly. Atkinson and Shiffrin further maintain that as long as information resides in the STS one has the opportunity to transfer it to the LTS under suitable control conditions. They also accommodate the view that the STM is correlated to consciousness. Further they must

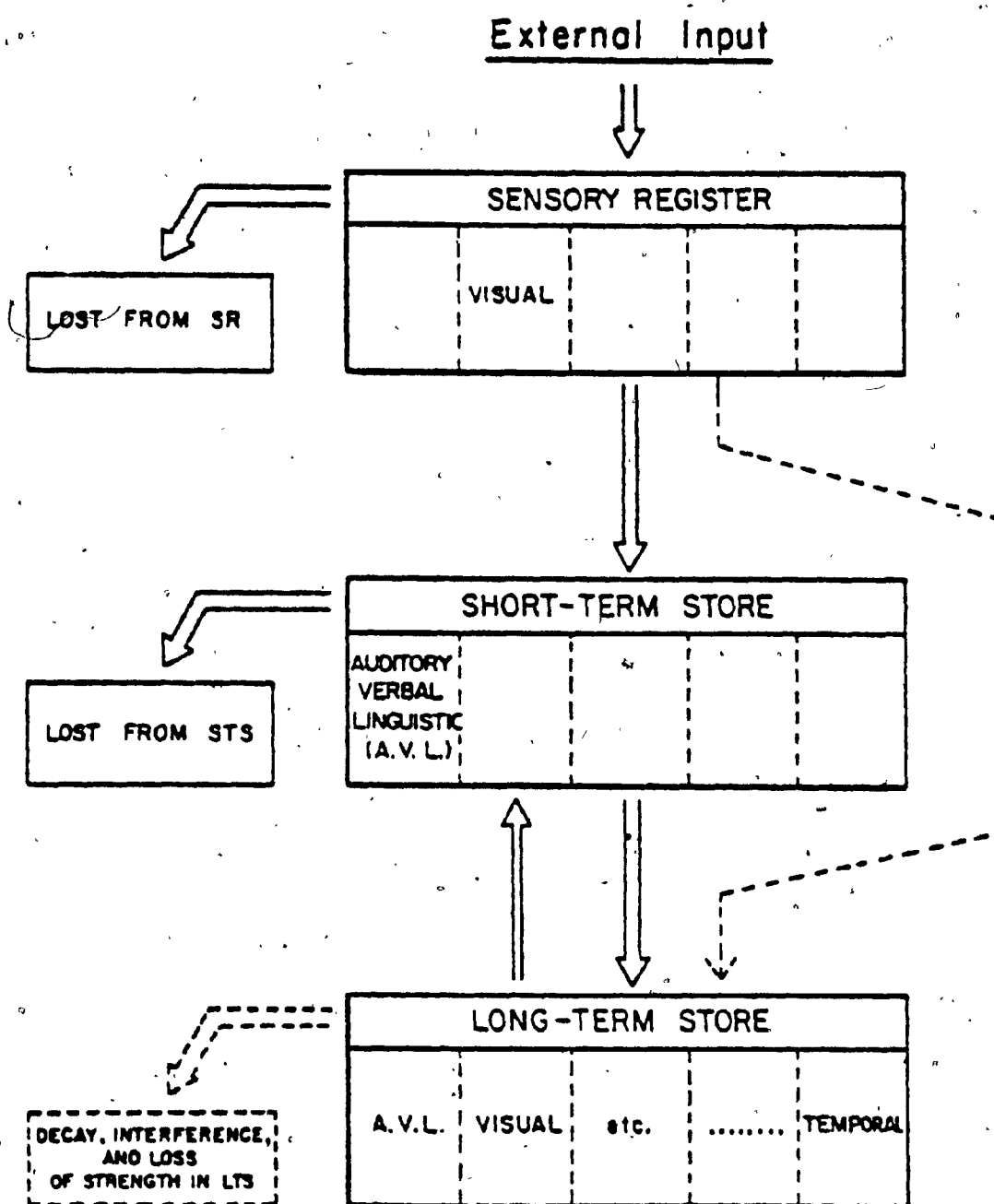


Figure 2. The basic structure of memory according to Atkinson and Shiffrin (1968)

take account of the fact that during a problem solving task such as the addition of two to two in order to obtain four, Ss must remember the steps and elements they manipulate in order to produce coherence in their mental productions. They are thus obliged to admit that the LTM sends information to STM.

C. Control Processes

We mentioned earlier that by the term "control processes" Atkinson and Shiffrin refer to those processes that are not permanent and invariant factors in a memorial process. Rather, they are transient factors under the control of the subject. Their appearance and utilization depend on such factors as instructional set, experimental task, past history of the subject etc. To make our discussion more concrete we shall examine some of the more important ones in some detail.

Control Processes in the Sensory Register

Information flows continuously from our environment into our CNS via all the senses we have. The first obvious choice is the choice of which modality to attend to. This decision is the primary control process relating to the SR. A related attention process is in the transfer to STS (direct and/or indirect) of a selected portion of information originally residing in the SR. Scanning processes within the SR can also become control processes.

Control Processes in the Short Term Store

Storage, search, retrieval, rehearsal, coding and transfer are some of the processes mentioned by Atkinson and Shiffrin. For the purposes of this work only some of these are of interest and will be

dealt with.

(i) Rehearsal: Laboratory tasks on memory often involve meaningless tasks such as memorization of nonsense syllables. In such situations rehearsal of material becomes quite important - in the same way that telephone number memorization is done through silent internal verbalization of the number. Further, as Atkinson and Shiffrin believe, as long as a given piece of information resides in the STS it can and is copied into the LTS. Believing in some form of consolidation theory Atkinson and Shiffrin claim that the more information reverberates in the STS the stronger its trace in the LTS. And finally, via rehearsal, a process that increases the lifetime of a given signal, one gives the opportunity to the coding processes of the LTS to work on the signal.

There are capacity limits for rehearsal capability. Experimentally we find that we can rehearse between 5 and 8 items at a given time. This gives rise to the notion of a rehearsal buffer. This is perceived to be some sort of a bin in STS in which there are exactly n items. Each new item entering the rehearsal buffer 'knocks out' an item already there. The maintenance and use of the buffer is a process under S's control. There are some constraints however. There seems to be some sort of upper limit of nervous energy available per unit of time. This means that if all our efforts are devoted to one thing we cannot do another. For example in setting up a maximum size buffer Ss cannot at the same time perform other activities such as solving crossword puzzles, say. In other words, a trade-off occurs between allocation of buffer size and other control processes such as coding and problem solving. Atkinson and Shiffrin believed then, that the best way of understanding memory processes was to consider information flow in and out of STS

and control processes that come into play via this flow. Figure 3 shows the rehearsal buffer and its relation to the memory system. Whatever resides in the buffer comes from elements in the STS. Presumably, they originally entered the STS either from the SR or from LTS. The arrow toward LTS is to be understood as indicating the process of building up a trace in LTS. The second arrow on the left (within the STS), indicates the "bumping off" process. Items entering the rehearsal buffer bump off some others already in it. The rejected items decay from STS.

(ii) Coding and Transfer of Information between Short and Long Term Store

Usually there is no direct transfer from SR to STS. Rather, a long term search and matching process is first assumed to occur in the LTS whenever something is presented in the SR. Once the item is identified it then is transferred to STS. For words and other familiar stimuli the process is automatic and it is executed quickly. From STS we can transfer information into our LTS. Usually this transfer is coded. By coding Atkinson and Shiffrin mean "a select alteration and/or addition to the information in the STS as a result of a search of the long term store." This change may take a number of forms, often using strong pre-existing associations already in our LTS.

In this writer's view the authors' treatment of information transfer from STS to LTS and back, because of its abstractness tends to be rather schematic. Without going into arguments concerning deep and surface structures in linguistic analysis it should be obvious that a person upon hearing a sentence such as "The sky is red" would probably proceed along the following lines (at least):

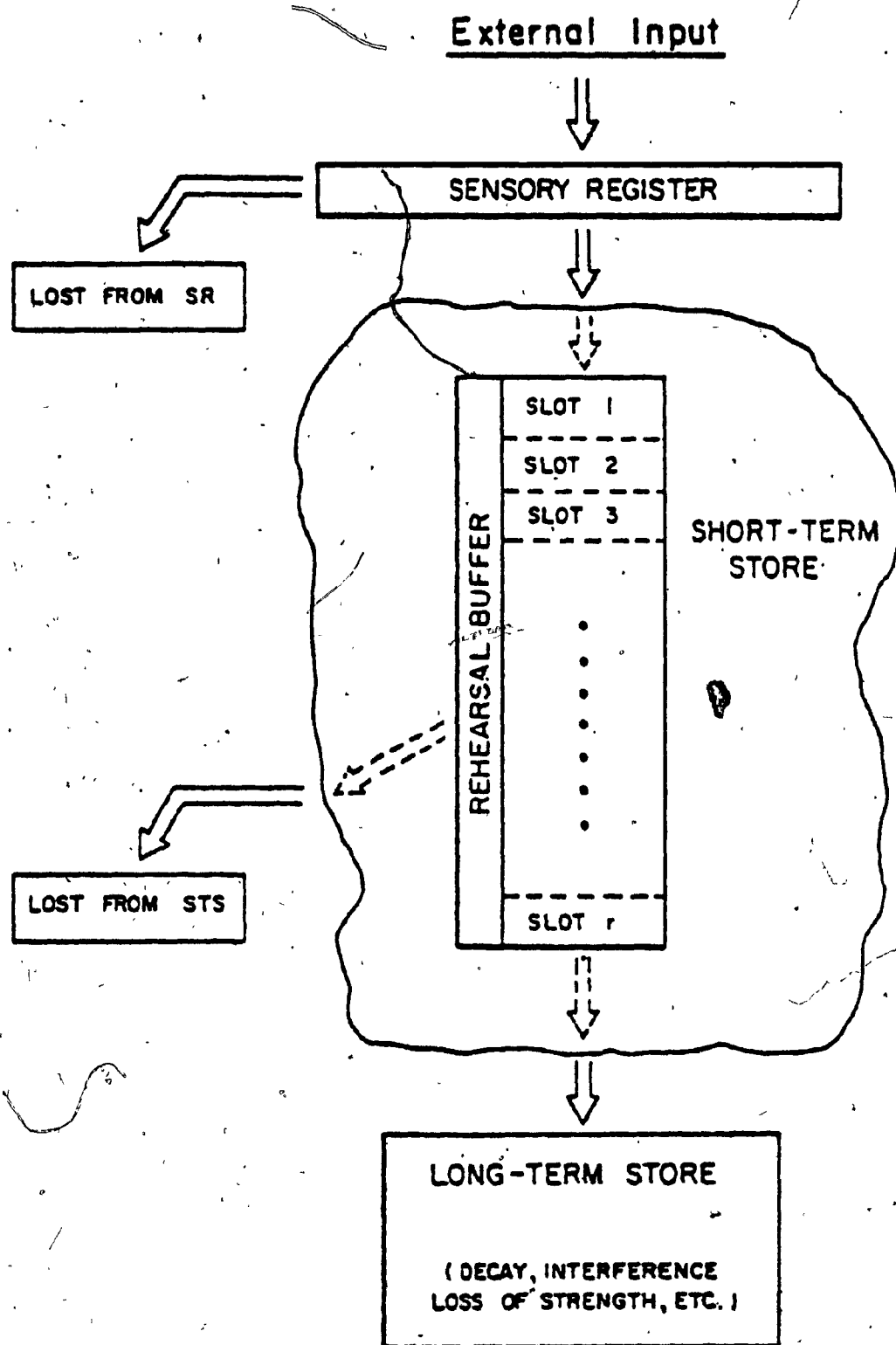


Figure 3. Detailed schematic of the short term memory with assumed processes therein. (From Atkinson and Shiffrin, 1968)

- (1) Extraction and apposition of meaning to each of the words making up the sentence. The meaning should be extracted in the LTS and transferred to the STS.
- (2) Attribution of meaning, via deep structures, to the words organized in the string "The sky is red". In other words the string is now transferred back from STS to LTS for apposition of meaning to the string as a whole.

Following this, one has to transfer to STS again. From the point of view of Atkinson and Shiffrin this process must be obviously automatic.

(iii) Control Processes in the Long Term Store

Let us now consider what happens to the LTS-side of the STS-LTS interface during the encoding process. To illustrate their argument Atkinson and Shiffrin use an experiment done by Clark and Bower. In this experiment Ss were required to learn lists of paired associate words. Two groups were tested. One of the groups was told that the best way of remembering the list was to form elaborate visual images connecting the paired associates. The control group was not thus instructed. Subsequently, the groups were told to pair-associate sequences of words. It was found that the experimental group outperformed the control group by 40%.

The effectiveness of the visual image methods has been known since antiquity. Such a method, for example, is the locus method, a fuller account of which may be found in Norman's book (1976). Through this method whole speeches could be remembered word for word. Locus techniques were in fact part of the curriculum of the study of Rhetoric. Norman citing a manual by Quintilian describes this method by saying

that artificial memory is established from places and images. (Constat igitur artificianca memoria ex locis et imaginibus). "A locus is a place easily grasped by the memory, such as a house, an intercolumnar space..... or the like. Images are forms, marks or simulacra (Formae, notae, simulacra) of what we wish to remember." If we wish to remember much material we must equip ourselves with a large number of loci. Within each locus we imbed the item we wish to remember. Further, if these loci are arranged in order their contents will then be retrieved in order. Atkinson and Shiffrin thus led to examine why some particular techniques such as the locus methods and others like them help in memory acquisition. They reach the conclusion that success of a given method depends on the fine structure of LTM.

According to them a likely reason for success is the possibility that the encoding process might be making use of strong pre-existing associations. In their first paper Atkinson and Shiffrin did not have much to say about long term search processes. These processes they studied in a later paper, (Shiffrin and Atkinson, 1969) to which we now turn since this part of their work bears directly on the results of our own work.

iii (1) Storage and Retrieval Processes in Long Term Memory

In order to speak about placement of information in LTS, Atkinson and Shiffrin make a further assumption concerning the nature of LTM: Given the vast amount of material in LTM, and since it seems unlikely that all information in this store is searched when a particular item is required one must assume that the long term memory is partly self-addressing. Self-addressing proceeds by determining an appropriate

and unique storage location based on the information make-up being processed. In their own words (1969) "An ensemble of information, presented to the memory system will define a number of memory areas in which the information is likely to be stored... just, as for example, a book on caulking methods used for the 12th century Egyptian riverboats, will be placed in a specific library location. (The Egyptian Room). If a user desires this book it may be located by following the shelving plan used to store it in the first place. The LTS is assumed to be partially self-addressing in that the degree to which storage locations are specified by the ensemble will vary from one ensemble to another and it will also depend on the task in which the particular ensemble is imbedded. On the basis of the above one would not expect an exhaustive search in a recognition process. For example, an ensemble of information presented to a subject for recognition would determine a search process whose spread would be proportional to the number of sharp features the ensemble could be broken into. The 1969 article does not specify the microstructure in any further detail. For example, the term "location" in memory is defined in relation to the organizational schema of long term memory; a "location" in the LTS is basically defined by the place in the organizational structure occupied by an information ensemble. Two locations are supposed to be "near" or "close" to each other if they tend to be retrieved together. A "code" or an "image" is an ensemble of information that is closely related and likely to be retrieved together with the stimulus ensemble. It should be born in mind that a given code might comprise several codes for other inputs, depending on the tasks at hand.

(2) Storage and Retrieval in fine detail

The requirement of a partly self-addressing memory makes storage and retrieval two very closely related processes. One reflects the other so to speak. For purposes of logical clarity each of these processes is subdivided into three subprocesses. Storage is assumed to consist of three basic components or mechanisms: Transfer, placement and image production. The transfer mechanism includes those control processes whereby one decides what, when and how to store in LTS. The placement mechanism determines the location in which the information vector (or ensemble) is to be stored in the LTS. Obviously, the primary mechanism that determines storage location and placement is the actual organizational network in place. Self-addressing can work provided placement depends on prior, already established, LTS organization; each item is sent to locations that depend upon a pre-existing organizational framework. The image production mechanism is supposed to determine the proportion of the currently available information in STS that gets stored in LTM. On the basis of what we said about rehearsal, image production is supposed to be linearly dependent on the amount of rehearsal. Atkinson and Shiffrin do not involve any other mechanisms determining transfer from STS to LTS. It should be obvious that there must exist others. For if STS is to be construed as some sort of working memory then a mechanism must exist which decides which parts of our current stream of consciousness are relevant to a given task and which are not. In essence we are talking about the mechanism that controls attention and set.

Retrieval is assumed to consist of three mechanisms: Search, recovery and response generation. The search process of which we shall

speak more on, is basically a recursive loop during which different memory locations are pulled up for examination. The recovery process determines what proportion of information must be selected per recursion and inputted into STS. The response generator examines the contents of STS after each recovery and decides whether the process must be terminated or another loop of the recursion process must be executed. The retrieval process begins with the presentation of some sort of query such as "what did you have for lunch two days ago?". This type of query delimits the contours of the search process and the type of answer that is considered satisfactory. On the basis of these S starts the recursive loop by looking at one likely memory one after another. The process of recovery of information from a particular chosen location is actually called recovery. The recovered information is put into STS which must also contain, at that time, other pieces of information necessary to correlate to the recovered information. Depending on the suitability of the recovered information the total current contents of STS are "processed" and the next stage of recursion is set if necessary. Basically, the STS is considered as some sort of "window" to the sequential processing of information in the LTS. The decisions regarding the suitability of current contents of STS (from the point of view of the task at hand) and what to do with them come under the title of the response generator-process. If a decision is made to continue the search then a new location is selected on the basis of information just recovered, or in accord to some external search strategy. The total cyclic and recursive process that continues until termination is called the search process. Shiffrin in his 1970 article paid special attention on the problems relating to the topology, so to speak, of the search process. By this we mean the

the space on which the search strategy operates. His ideas, as he himself admits, were based on ideas of William James and James Mills. The search process according to them is akin to the activities of somebody who lost something and is trying to find it by going and looking at various likely locations. To sharpen his analytical tools he first modifies the definition of memory trace. Instead of considering codes and images as they are elicited by given stimuli, he studies the content of codes directly. Accordingly he breaks up the trace into the information units (or I-units as he calls them) that make up its code. He assumes that these units of information are interassociated and, as a rule, tend to be recalled together. The search process is based on these elements rather than the images. (Images are supposed to be composites of informational units). The LTM is viewed as a space of I-units on a subspace of which the search process sequentially happens. This subset is known as the search-set. Schematically, the topology of the LTS is shown in figure 4, due to Shiffrin (1970). The large irregularly shaped area enclosing all others is supposed to represent the LTS. Crosses stand for the units of information. The search set is the large solidly outlined areas within the LTS. The search occurs in this area. The dotted circles with their crosses separated from each other to emphasize the unitization property of images. An image is supposed, in other words, to be a set of distinct objects $O_i, i = 1, 2, \dots, n$. The selection phase of the sequentially occurring search process consists - according to Shiffrin's schema - in randomly choosing a unit information from within the elements of the search set. If the search set has say n units in it, whereas a particular image is composed of n_i units then the probability of 'hitting' a particular image is n_i/n .

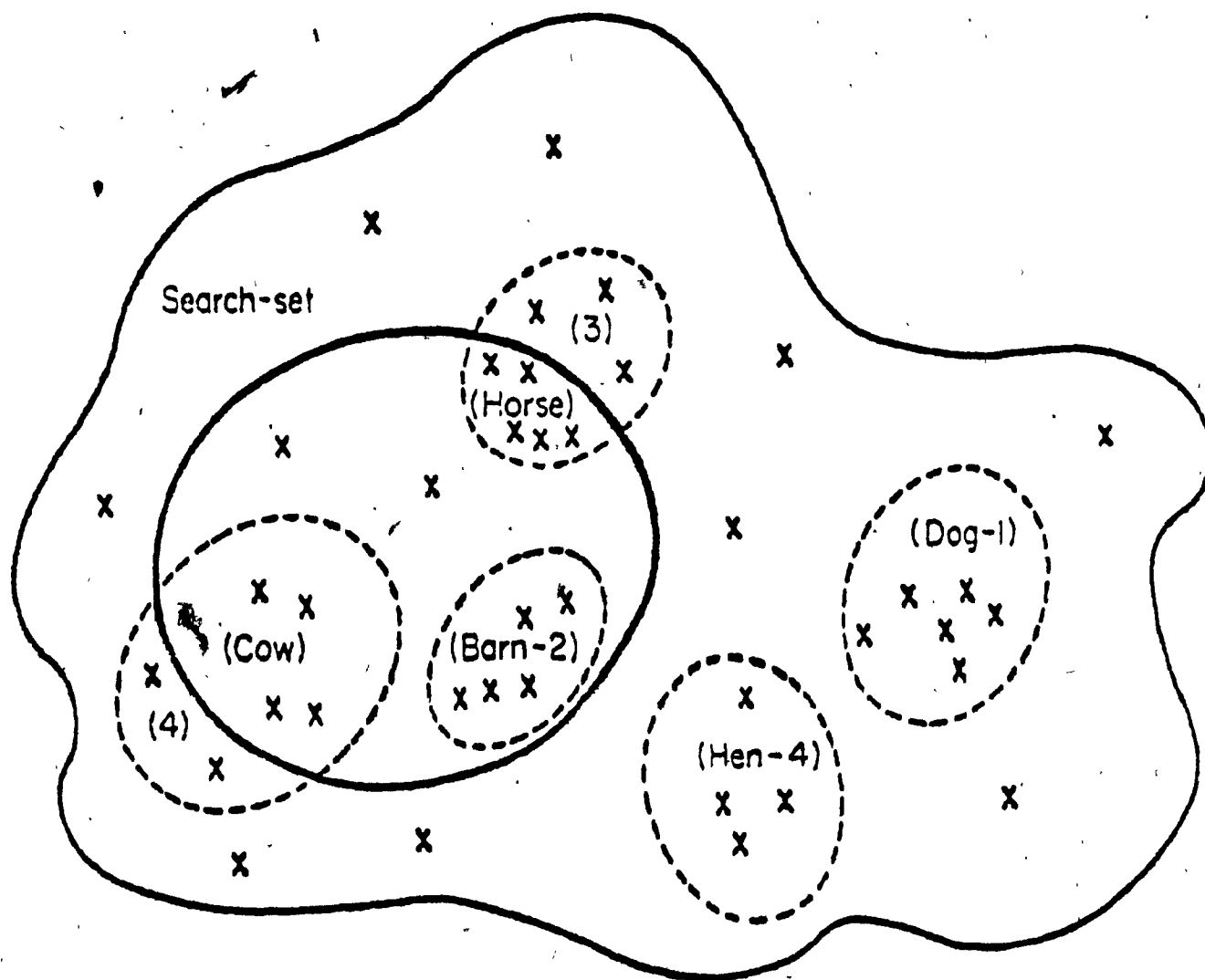


Figure 4. Assumed topological structure of long term memory according to Shiffrin (1970)

Anticipating our own discussion we must state here that it is unlikely that the image space is made up of mutually exclusive image sets S_i . For example in a paired-associate task words which were learned together will tend to have images S_i, S_j so that $S_i \cap S_j \neq \emptyset$. This would be a result of the context within which a learning task is initiated. Triggering of S_i implies triggering S_j because processing of S_i finally leads onto I-units of the $S_i \cap S_j$ set which is subset of S_j .

In the Shiffrin schema the unitization property of a given trace is further exhibited by his claim that if one I-unit belonging to S_i gets triggered the whole of S_i gets triggered as well.

The global aspects of the search process is shown in figure 5. It is basically the same as that of the Shiffrin and Atkinson article (1969) only now Shiffrin uses search sets and I-units to describe the process. In the Shiffrin recursion a search-set is selected at the start of each of the loops of the search process. This is the crucial aspect of the whole operation and of major interest to us. Of course one must somehow describe the conditions, factors and constraints that determine a particular choice. Among the constraints Shiffrin mentions such things as the task set for the subject, response required, clues or stimulus information given in the test. An extremely important factor governing selection is based upon temporal clues. Search sets may change after a cycle is completed. This would be due to information uncovered during the search of the LTS. In problem solving searches this is exactly what we do. For example, if one is queried as to what he ate for lunch five days ago one is obliged to embark on a reconstructive search tracing one's actions backward in stages - each stage

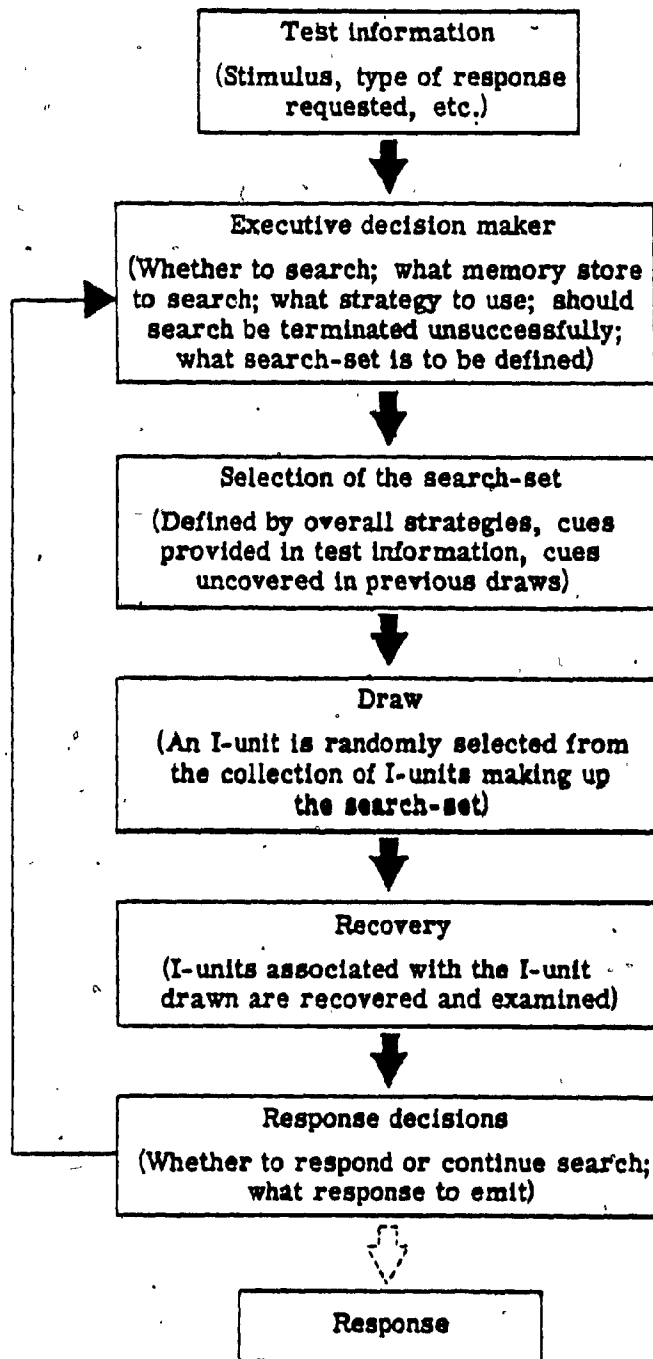


Figure 5. The flow chart of the search process according to Shiffrin (1970)

defining a new search-set with a different time tag. The information uncovered after each loop helps to define the new search-set. To further understand the process aspects of memory one must have a more intimate acquaintance of the microstructure of memory. Basically, what we are talking about is the representation of our world in our own LTS. Figure 6 shows a global schematic of the memory system from the point of view of Atkinson and Shiffrin.

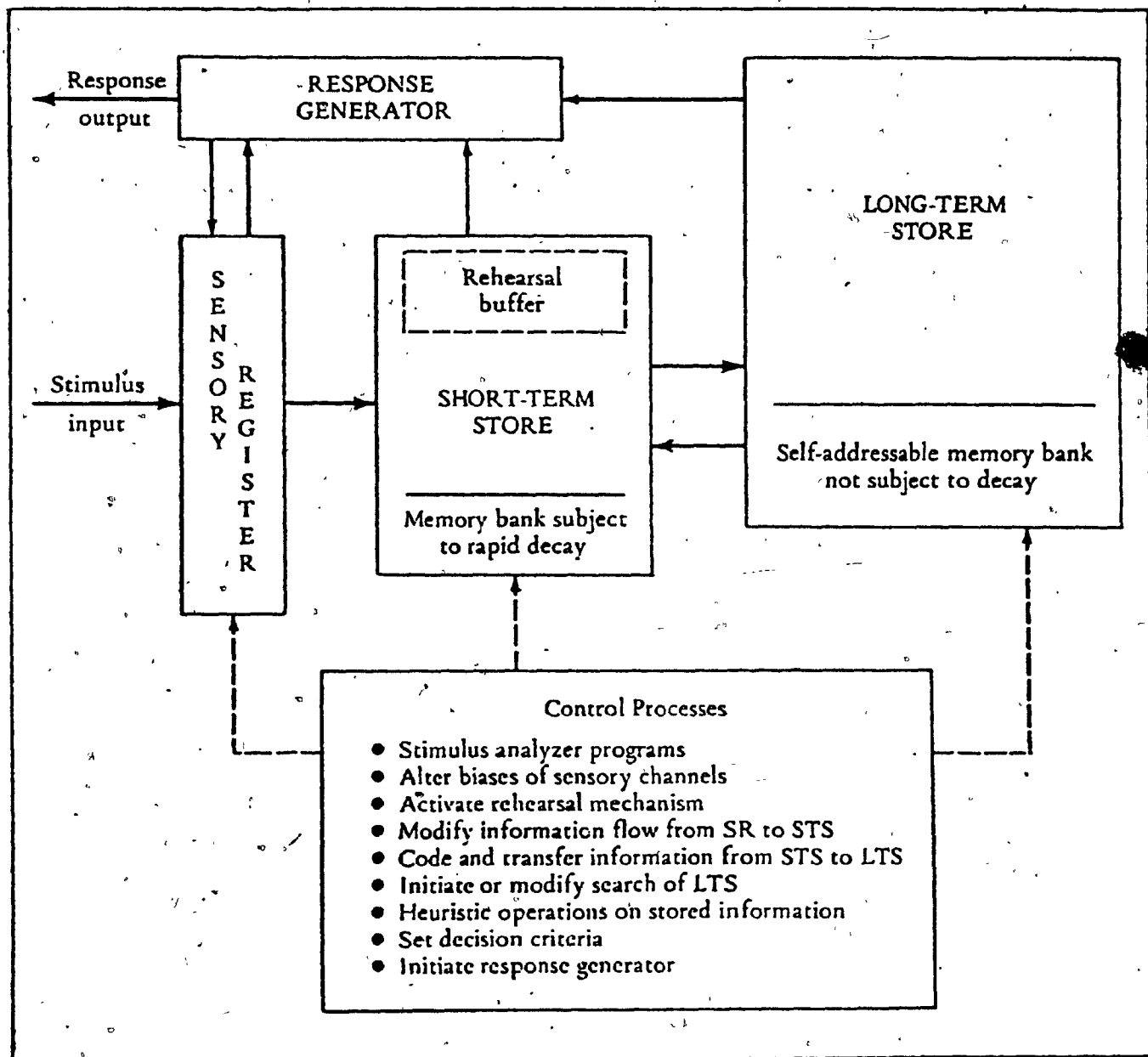


Figure 6. Global schematic of the memory system and its control processes. (From Shiffrin and Atkinson, 1969)

SOME ASPECTS OF PRESENT DAY MEMORY RESEARCH

PART III: PROCESS ASPECTS OF MEMORY CREATION AND THE REPRESENTATION OF INFORMATION IN LTSA. Introduction

The representation of knowledge is one of the fundamental and probably the most difficult subject in present day memory research (Norman, 1976). Answers to some of the questions its problematique poses will have serious and far reaching consequences in such fields as educational strategies for children to treating neurotic and psychotic conditions. Answers to most of the questions are, at this stage, speculative with very little experimental data to go on. Most of the evidence is indirect. A typical question in this kind of research would be the following: "In remembering a book, for example, do we somehow remember the physical attributes of the book such as visual appearance of words and sounds of these words, or do we have in mind some abstraction of its meaning? Do we keep some mental image of the events in the book?"

Before we embark on a more detailed presentation of some of these issues it is perhaps useful to define some of the vocabulary used in these studies.

When mental representations of the external world reflect our perceptual experiences directly we say that we have analogical representations of our world. Another closely studied representation is the so-called propositional representation. This approach takes its inspiration from the work of linguists such as Chomsky and others and it has adopted a great number of the analytical categories of these

disciplines to the requirements of memory research. (Such for example, is the attempt to analyze incoming verbal messages into surface and deep structures). A propositional system encodes information in terms of interpreted and/or abstracted statements concerning the external world. This is justified by the fact that the meaning of a given word is often related to the way we have categorized it. The simplest propositional representation in linguistics is the declarative sentence of the type "A is B". Psychologists such as Lindsay, Norman and Rumelhart, known as the LNR research group have constructed models for the structure of LTM and have made models claiming that all knowledge of the world is of a propositional nature.

B. The Propositional Representation of Rumelhart, Lindsay and Norman (1972)

This representation unfolds along two basic dimensions of the memory system: On the one hand one must take into account, the permanent structural components of LTM namely what they call the data base. On the other hand the nature of the processes can be considered to be part of the data base and can be activated to process data from the data base. An external interpretive process operates on the data base, retrieving information when necessary or activating library subroutines when it is necessary. This process is responsible for example, for answering questions, storing and evaluating new information etc. The LNR group takes the attitude that memory works basically like a problem solving process, tackling serially parts of the memory structure until final solution is obtained.

B1. The Structural Aspects

The basic structural element of the data base is a set of nodes interconnected by a relation R , and symbolized by $R(C_1, \dots, C_n)$. Nodes represent any cluster of information in memory. A relation is, in other words, an association among a set of nodes. Two important properties of a relation are its labeling and its direction. Depending on the direction of a relation the relation assumes its meaning. Reversing the direction of a relation changes its meaning. Our memory system is supposed to be using these relations (or associations) to encode semantic and logical relations among nodes.

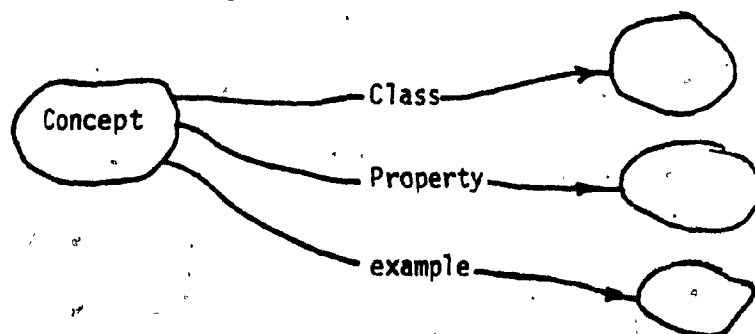
The graphic representation in the simplest possible case i.e. that of a binary relation can be symbolized by

$$A \xrightarrow{R} B$$

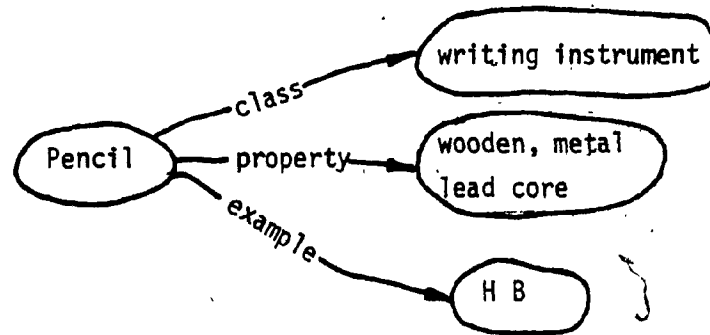
The rationale for the above abstract construction becomes clear if one examines the lexical definition of various concepts in a typical dictionary. One finds that typically definitions include one or all of the following:

- (i) the class of concepts to which the given concept belongs
- (ii) the properties or attributes that particularize or make the concept unique.
- (iii) examples of the concept.

Schematically, then, a concept may be defined according to the schema below. (Lindsay and Norman, 1972).



For example, the word 'pencil' could be defined according to:



In general, and to economize space, Lindsay and Norman (1972) adopt the following conventions in verbalizing the above schema.

class \longleftrightarrow isa . As in 'John is a man'

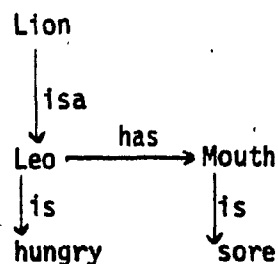
Property \longleftrightarrow has . As in 'John has feet' (properties are objects)

property \longleftrightarrow is . As in 'John is tall' (properties are qualities).

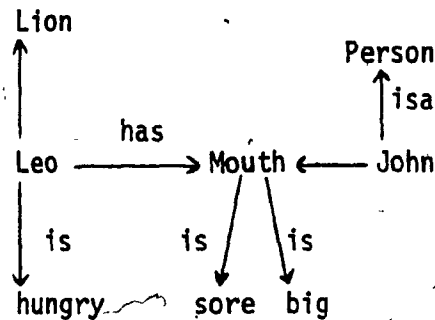
Figure 7, from Lindsay and Norman (1972) would represent the definition of the word tavern as it is imbedded in the semantic memory net. The meaning of the concept is extracted from the interrelationship it has to other proximate nodes. Further, every piece of information is stored as a node plus its relations. The node is the only addressable unit in the data base.

B2. Primary and Secondary Nodes

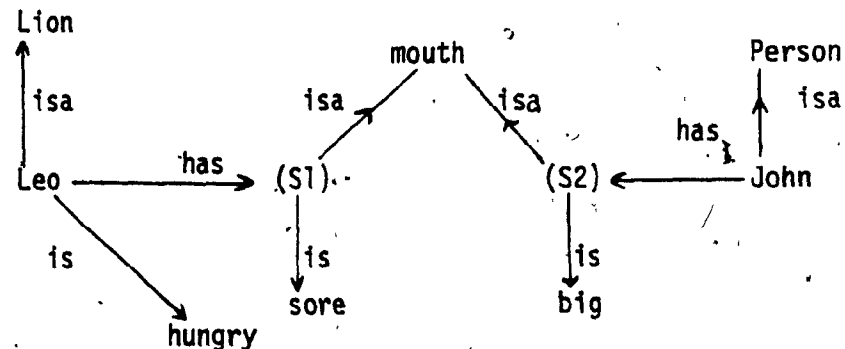
Consider now the process of coding the sentence "Leo the hungry lion has a sore mouth" (Lindsay and Norman, 1972). The semantic net fragment appropriate to this sentence could be something like the schema shown below.



One requirement, however, of our constructs must be their ability to interface with new incoming information without destruction to the rest of the semantic net. How would we then add to the above fragment the additional piece of information that "John is a person who has a big mouth"? If we just grafted this piece onto the old piece we would get:



This obviously won't do since the lion's mouth in addition to being sore now becomes sore whereas it is only big. The way out of our difficulty is to realize that the concept (mouth) can become the name of a class with particular instances as they occur in specific situations. We only need to have one definition of the word mouth known as the primary definition. Specific daily instances of the definition are called secondary definitions. The usual way of denoting secondary definitions is with brackets. Consequently, our new correlated graph would now look like the schema shown below



Here, when we are retrieving information about the mouth of the lion Leo we will only read the secondary node (S1) which truly describes the soreness of Leo's mouth.

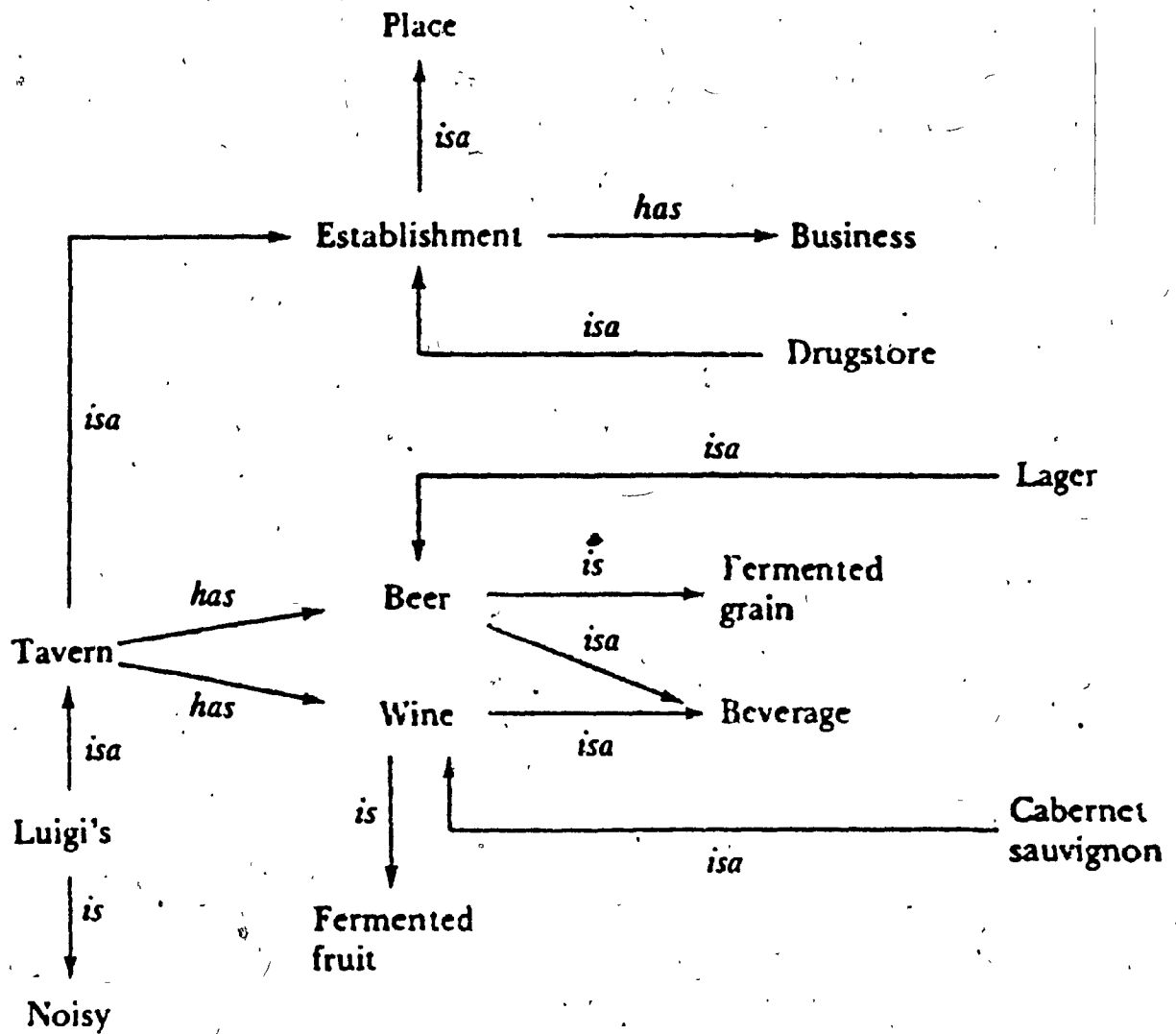


Figure 7. Nodal structure in long term store: the definition of the word 'tavern' according to Lindsay and Norman (1972)

Rumelhart, Lindsay and Norman in addition to concepts, consider such situations as events and episodes. An event represents some sort of scenario with actors and actions, and an episode is a connected series of events. This, in other words, represents the internalized aspects of human activity. Since this is of no consequence in this work we shall make no further use of it beyond the final statement that events and scenarios can also be represented as nodes. As a final example of the kind of net one is assumed to have stored in his LTS the reader can consider figure 8, due to LNR (1972). This figure illustrates a section of the memory space which encodes the proposition that "In 1865, an avalanche destroyed a tiny village at the bottom of a mountain in Switzerland." The figure gives the entries for avalanche, mountain, fall, move, tiny, high and mass, as well as the proposition itself.

The definitions of relations and concepts is called the meaning-store and is presumed to be derived from experience with repeated appearances of the lexical item in specific uses or by some other mechanism. The exact physiological process which is the substrate for meaning acquisition is the subject of a great number of hypotheses the most interesting of which is the Hebbian Theory and its various modifications.

The picture that emerges from the LNR construct is that of memory system consisting of sets of decision points (nodes) and directed routes. In principle, however, it is possible, starting with any given node to reach any other by making the right kind of decisions as to directions. The important conclusion from this is the interconnectivity property of all information. This is quite important whenever one comes to study such aspects of behaviour as free associations and their physiological substrates.

C. Process Aspects of the Model

We must now consider the kind of cognitive processes that operate on the data base. We will then be able to see the relation between this model and the search procedure of Shiffrin. Since the memory system must be able to accept information from the environment we must see how new facts are added to the base. Second, we must examine the reverse process, namely, the output in a coherent form of information already imbedded in the memory system.

C1. Addition of new information

This process has already been illustrated in the previous paragraph. If one is informed for example that (Lindsay and Norman, 1972): "Mary is a fat, pimply person". Then one decomposes this statement into three propositions:

Mary is a person

Mary is fat

Mary has pimples

The ensuing three-branched node <Mary> is grafted on the appropriate position of the data base.

C2. Making Responses and Searching the Data Base

Assume we have a data base shown on figure 9 (due to Lindsay and Norman, 1972). If the owner of this data base is queried as to the meaning of the word 'teapot' he could produce the following string of statements: "A teapot is a pot. It is ceramic. It has a handle, a spout and a liquid in it that is hot. Further, a pot is a container that....." The procedure implied is the following. A control procedure is involved that first accesses the node "teapot". The

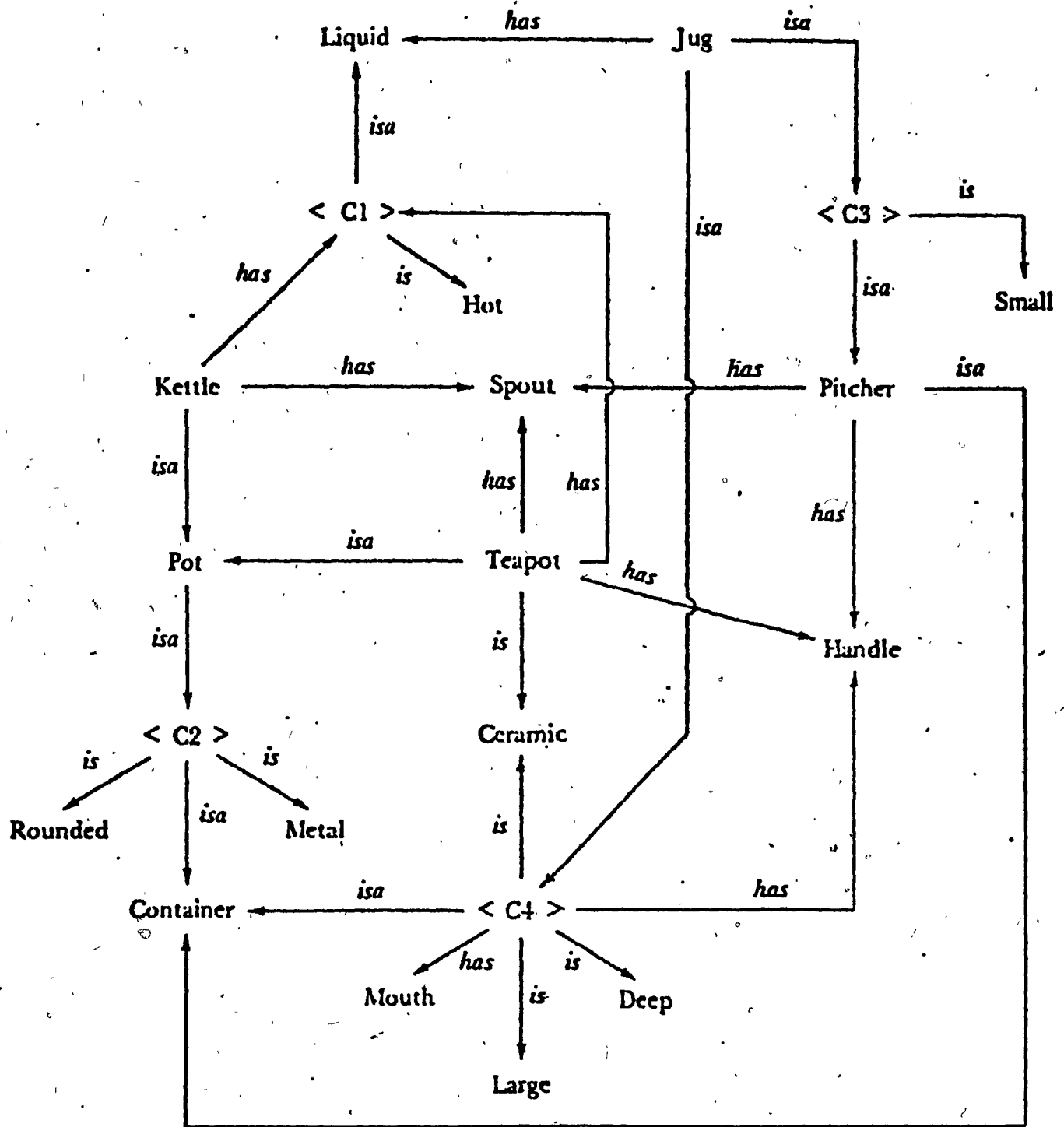


Figure 9. Primary and secondary nodes on a fragment of memory space as they are understood in the system of Lindsay and Norman (1972)

attempt is then made to follow the various paths emanating from this node. The strategy for recall that this process seems to imply is sequential. First we describe the information immediately surrounding the node. This always brings us to some other node. We now describe this node and so on.

Lindsay and Norman (1972) show that their system can, not only scan the data base but also readjust it to a certain degree in order to make corrections due to new empirical findings. A typical adjustment occurs when the system generalizes and produces abstract concepts. This it does by examining all instances where a given property or quality is held in common by several concepts. Whenever the same information is found to be stored in all concepts one generalizes the common information. In other words one "moves upwards" or abstracts the common attributes.

We should here point out some common ground between Shiffrin (1970) and Lindsay and Norman (1972). A node representing a concept is a connected unitized set of informational elements x_i such that a concept is defined by the assembly of its information units. According to Lindsay and Norman, however, a node corresponding to a given concept must be analyzed by other nodes. Recall the definition of the word "pot" above. We believe that Lindsay's and Norman approach is more adequate. It has greater analytical power. It implies among other things the sharing and interpenetration of nodes. This becomes obvious whenever one considers the fact that many attributes of various nodes are held in common. Consequently, if c_1, c_2 are two nodes that are defined according to attributes a_i so that:

$$c_1 = \{a_1, a_2 \dots a_i \dots a_n\}$$

$$c_2 = \{a'_1, a'_2 \dots a'_i \dots a'_n\}$$

one has many times the possibility that $a_i = a_j$. In other words, if C_1 and C_2 have attributes in common $C_1 \cap C_2 \neq \emptyset$.

In general, both Shiffrin's and Lindsay and Norman's schemas may be abstracted so that the search and recognition process for a given concept C may be represented as an attempt by the CNS to produce memory unit "vectors" M_i so that for any given concept C :

$$C = \bigcup_{i=1}^N (C \cap M_i). \quad \text{Equation 1.}$$

The above representation based on set theoretical grounds implies, among other things, the following:

1. A given input signal C may have different meanings for a given individual during different times of his life and among different individuals sharing different experiences.
2. The learning process defined as the process whereby the different M_i are laid down on a CNS by societal institutions such as schools, etc. have the effect that a great number of M_i 's are held in common.
3. Discourse becomes possible because of 2.
4. If one takes the informational point of view one can say that C is decoded into the set of M_i 's or C 'means' the set of M_i 's. If no M_i 's occur for a given input signal C no meaning exists or C cannot be decoded.
5. The structure of the memory system is used to analyze the external world of experience in a dynamic fashion. Our perceptions of this new process will be direct results of our memory base.

6. For any given civilization and historical epoch the maximum set of M_i 's is limited and slow to change.

C3. Possible Constraints on the Scanning Process

It is obvious that all of the data base cannot be handled at once. The scanning mechanisms of our memory system is assumed to be working in a sequential fashion interpreting a fixed number of chunks of information each time. Lindsay and Norman (1972) view the interpretive process as the analogue of a beam of light shone on the data base. Depending on the distance the apex of the cone of light has from the data base - viewed as a flat surface -- one covers more or less area.

Speculating, Lindsay and Norman postulate that there exists a maximum number of distinct chunks one can view. This maximum number is the span of STS. Since they take the chunk as the node they claim that the interpreter keeps 7 ± 2 nodes under its illumination each time. The number 7 ± 2 is the well known span of STS.

C4. Memory as an Information Processing System

Throughout their exposition Lindsay and Norman go to great pains to distance themselves from any view that presents the memory system as a passive system. In their model no account is thus given to either classical or operant conditioning. Their main emphasis is on the process, integrative and control aspects of memory systems. In this writer's view, however, it is very difficult to explain first the "seeding" and secondly the growth of the semantic net without association processes. Simon Newel (1971) in fact makes the flat statement that a memory structure cannot be built without the ability of our CNS to form associations. We believe that the ability of our memory system to form associations is a necessary but not sufficient condition for constructing a dynamic memory

system. For obviously control processes are there as well. In fact one has not only conscious control processes. Automatic processes that become unconscious as Schneider and Shiffrin (1977) have shown also exist. Other non-conscious control processes must operate as well. Let us recall Lashley's (1958) paper on Cerebral Organization and behaviour. In this paper Lashley cites a great number of non-conscious control processes with memory components. One such cited example is the control process implicated in the grammatically correct enunciation of sentences.

Recapitulating their view of how their data base grows one says that there is a continuous grafting of new information (via new nodes) onto the old structure. At any given time the world can only be perceived with the aggregate structure that exists. In a profound sense the statement that one perceives the world with the 'tinted' glasses of one's structure is obviously true. As the system matures reorganization becomes more difficult. The child's system, however, is quite labile. Further, a message encoded in a child's brain does not get the same elaboration it gets in a grown up brain. For children each new concept encountered must be built from scratch; by rote according to Lindsay and Norman. In our view this rote learning occurs via all sorts of conditioning and association building. Further, the ontogenesis of memory must proceed via a preponderance of conditioning in its early stages whereas control processes play a major role in a mature memory. The last remark follows from the fact that the basis of the semantic net is already in place. New things are learned primarily by analogy to already existing structures. Thus in the abstraction process associations are now being set up with selected groups of clusters of nodes rather than with individual nodes. As, however, the external world affects us both in a common and in an idiosyncratic fashion the idiosyncratic aspects of

an individual are also being built up.

The process view of memory due to Lindsay and Norman, as we noted above, made it easy to understand claims that the STS and LTS are not two distinct structures. We saw them speak of STM as the activation of a finite number nodes from the data base. Atkinson and Shiffrin characterized STM as a mode of excitation of the LTS.

The last paragraph brings us to the work of Craik and Lockhart which is a conscious attempt to explain the hitherto known results of memory research from the point of view of the process aspects of memory systems.

PART IV: PROCESS ASPECTS OF MEMORY OPERATION AND THE "LEVELS OF PROCESSING APPROACH"

In 1972, F. Craik and R. Lockhart of the University of Toronto published a paper of considerable importance. It was an attempt to explain the whole spectrum of memory phenomena directly from the process aspects of the memory system. Considering the set of data then available, (see figure 10), they showed that these same data could be interpreted not only from the storage model point of view but from another point of view as well. From their point of view, in other words, the criteria listed in the table, which incidentally, are more or less the same as those of Simon (1972), did not provide adequate grounds for distinguishing between separate stores. The actual theoretical structure of Craik and Lockhart is of interest to us and hence, we shall go into it in some detail.

A. Capacity

One of the salient differences between STM and LTM is their

Feature	Sensory Registers	Short-term Store	Long-term Store
Entry of information	Preattentive	Requires attention	Rehearsal
Maintenance of information	Not possible	Continued attention	Repetition
Format of information	Literal copy of input	Phonemic Probably visual Possibly semantic	Largely semantic Some auditory and visual
Capacity	Large	Small	No known limit
Information loss	Decay	Displacement Possibly decay	Possibly no loss Loss of accessibility or discriminability by interference
Trace duration	.25-2 sec	Up to 30 sec	Minutes to years
Retrieval	Readout	Probably automatic Items in consciousness Temporal/phonemic cues	Retrieval cues Possibly search process

Figure 10. Some commonly accepted differences among the three stages of verbal memory, according to a structural model. (From Craik and Lockhart, 1972)

hypothesized difference with regard to capacity. STM has a limited, whereas LTM has unlimited capacity for information. Craik and Lockhart, however, shrewdly pointed out that it is indeed unclear as to how this limitation in capacity is to be interpreted. One could, for example, consider it as a limitation in processing capacity of the memory processor rather than a limitation on storage capacity of STM. In other words they asked themselves whether the STM bottleneck is in the amount that can be stored or the rate at which information can be transferred to, or copied into LTM. Attempts to measure STS capacity usually imply a measurement of storage capacity. The problem, however, with all these measurements as Craik and Lockhart point out is the variety of results one obtains. The range is, from about three, (Badeley, 1970) to about twenty words in a sentence. (Craik and Massani, 1969). Craik and Lockhart take the disparities of these results as a phenomenon to be explained by any valid memory theory. The fact that chunking has been proposed as a mechanism by various storage model theorists is unacceptable since as they point out chunking is an elaboration process. Hence, they conclude, this capacity limitation must be taken as a limitation of processing; storage capacity being a derivative of processing capacity which is the more fundamental of the two.

B. Coding

We saw that according to Atkinson and Shiffrin information is coded in STS in the acoustic form (a-v-1), whereas when it passes to the LTS it is given a semantic mantle. Craik and Lockhart pointed out however, that even this distinction between the two stores can be blurred. Citing evidence from various workers they indicate that STM can be visual as well as semantic in addition to its being acoustic. On the other hand the LTS could be shown to have a visual component as well.

C. Levels of Processing

Taking their cue from Selfridge and Neisser (1960) and others, Craik and Lockhart make the claim that perception involves the rapid analysis of stimuli at a number of levels or stages. One recalls the pandemonium model of feature extraction for which a hierarchy of feature extractions progressively filter and finally define a given stimulus. (Lindsay and Norman, 1972). Preliminary stages are concerned, for example, with the filtering of different angles, lines, brightness etc. whereas later stages "template" match against stored abstractions of these stimuli from past learning. In other words the last stages of analysis are supposed to attach a semantic label to the input pattern. In the Craik and Lockhart treatment, this series, or hierarchy of processing stages is called depth of processing. The word depth is to be taken to indicate the number of cognitive acts required for pattern or input recognition and elaboration. Greater depth is to imply a greater degree of semantic or cognitive analysis. After the stimulus has been recognized further processing is possible via elaboration and/or association with other relevant structures. A typical example is a recognized word triggering its attendant associations, images etc. from the S's idiosyncratic memory. "In general, analysis is to be perceived as proceeding through a series of sensory stages to levels associated with matching or pattern recognition and finally to semantic-associative stages of stimulus enrichment". (Craik and Lockhart, 1972). One of the results of perceived analysis is the memory trace the strength and coding characteristics of which are a function of depth of processing. Deeper levels of analysis are associated with more elaborate, longer-lasting and stronger traces. Preliminary analyses are usually not stored since we are usually concerned only with the extraction of meaning from a given signal.

What are called the SR and STS are basically the sensory part and the intermediate part respectively of this analysis.

Further, highly familiar stimuli such as sentences and any other well known information can be retained more easily because these structures are highly compatible with previously existing cognitive structures. Consequently, elaboration for these will be greater and hence their trace will be stronger. This implies that if a message is imbedded in familiar cues the ability of the system to retain it will be greater. This has been verified by a great number of workers including Craik and Tulving (1975). These authors presented various words to Ss and by asking them various questions during the time of presentation influenced the degree of elaboration the given stimulus would have. In fact the stimulus now should be understood to be word + environment of word (in question). Shallow encodings were produced by rhyme imbedding of the word whereas deep imbeddings were provided implicating the given word in some sort of semantic or sentence structure. Results are as shown in figure 11.

In another experiment by Craik and Tulving as related by Reynolds and Flagg (1977) the degree of encoding was varied by imbedding the word in sentences of varied complexity. Steady increase in accuracy of recall was obtained with increase of semantic enrichment. On the basis of these experiments Craik and Tulving speak of spread or elaboration of the coding process rather than depth. Presumably, since sentences were used the same depth was attained for each one of them.

In summary, Craik and Lockhart view memory as tied up to perceptual processing which can be viewed as a continuum of processes. Memory, too, is viewed as a continuum of processes. On one end of the spectrum one finds the transient products of sensory register analysis. At the other

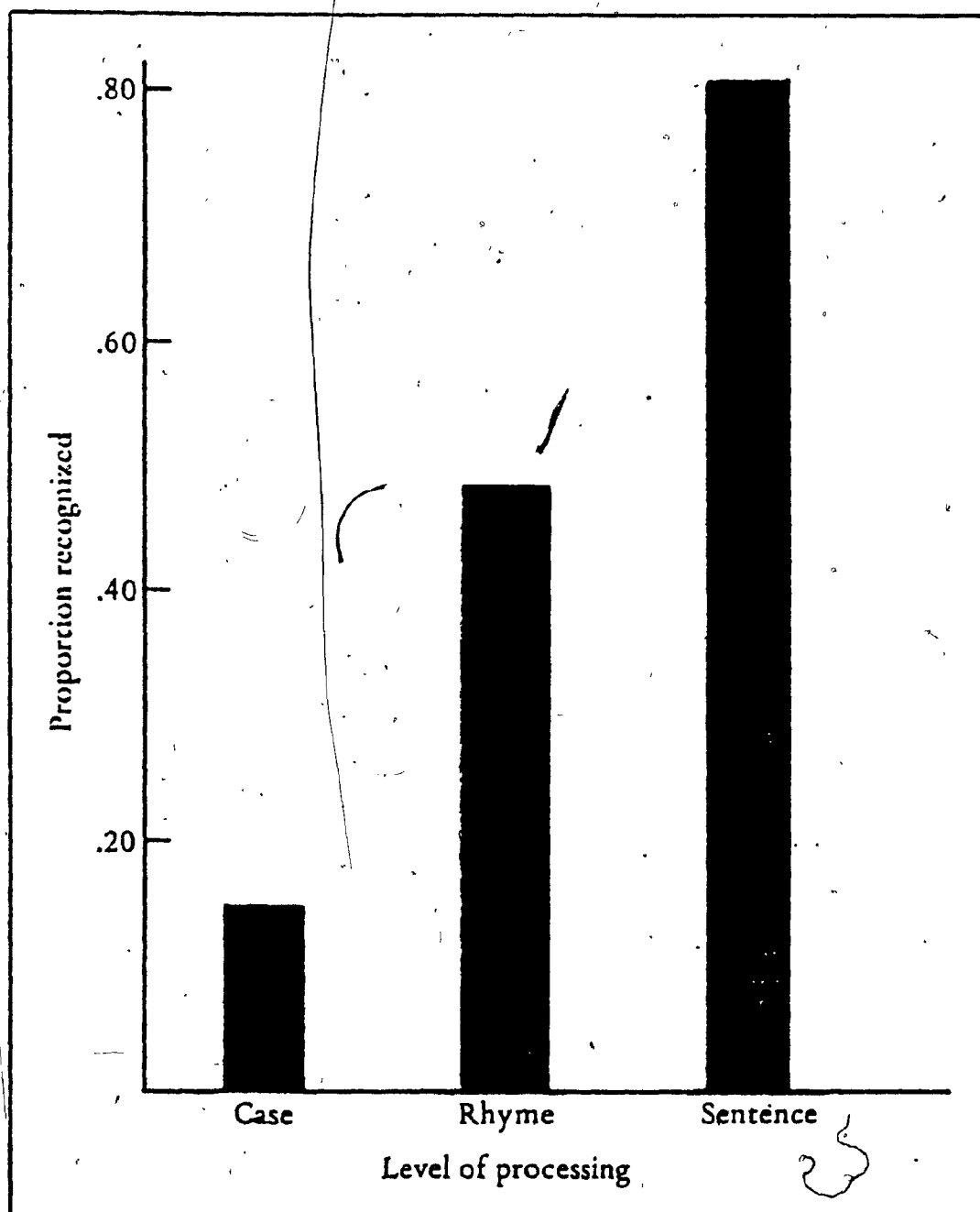


Figure 11. Effect of elaboration on recognition. (From Craik and Tulving, 1975)

end we will find the permanent products of semantic-associative operations.

Superimposed, however to this continuum of memory processing, they consider the existence of yet another memory mechanism. This mechanism is able to retain the current results of analysis at a given level via the process of recirculation. By this they mean the same thing as such statements as "keeping items in primary memory" etc. This process they call primary memory (PM) because they attach capacity limitations to it. They view brain function as being similar to the processes of a limited capacity central processor of a computer centre that at any given time can do a finite number of tasks. For example, if the processor is forced to keep the results of analysis at a fixed level of perceptual analysis then the observed phenomena of STM will appear. Another characteristic of this limited capacity is that the amount of information held at any one level would be a function of the "height" of the level at which recirculation occurs. At the deepest level already existing structures make for greater capacity. The basic characteristic of PM is that material here is continuously attended to. If attention is shifted elsewhere memory for material just attended to will decay. Information, however will be lost at a rate proportional to the level of analysis. The deeper the analysis, the longer the trace. Recirculation at a given level is called Type I processing. Shift to a deeper level is termed Type II processing. Obviously one can rehearse with Type II processing as well. This kind of processing students do when they prepare for an examination. According to Craik and Lockhart this is the kind of process that leads to better and longer lasting memory.

As Reynolds and Flagg (1977) have pointed out the essential components of the "levels of processing" alternative to the duplex theory of memory are:

1. The memory trace of a given stimulus is proportional to the depth at which the stimulus has been analyzed. Since more familiar stimuli get elaborated more they can be more easily remembered (if we followed Craik and Tulving we should talk, perhaps in terms of spread of coding rather than depth).

2. Memory persistence is best seen by analogy to the continuum of perceptual processing.

3. Recirculation or rehearsal in PM is not sufficient to improve memory because, if attention must be diverted from information in PM to elsewhere, one would have loss of information at a rate which is proportional to the height of the level the process occurs.

It is obvious from the above that their approach to memory processes could fit quite easily with the abstract characterization of the operation of the memory system we presented in the concluding remarks of page 54. Thus equation 1 can be adopted to the Craik and Lockhart schema as follows: If S_1 and S_2 are the memory traces of two comparable stimuli and if (say) S_2 is better remembered than S_1 this would imply the following:

$$S_2 \supset S_1 \quad (\text{comparability condition}) \quad (2)$$

$$S_2 = \bigcup_{j=1}^N (S_2 \cap M_j) \quad (3)$$

and

$$S_1 = \bigcup_{l=1}^M (S_1 \cap M_l) \quad (4)$$

with $N > M$.

Using the language of topology we would say that:

The space S_2 triggers, (or spreads) and contains the space that S_1 triggers. This is the equivalent to saying that:

$$\bigcup_{j=1}^N M_i \supset \bigcup_{i=1}^M M_i$$

which is obviously true since $N > M$.

Anticipating later chapters we can see from, even here, that recall for S_2 would be easier than S_1 since access to S_2 may be obtained with a greater number of combinations of the M_i 's than is the case with S_1 .

Using the language of Craik and Lockhart S_2 is better remembered because its spread (3) is greater than the spread of S_1 (4).

Our reference to the model aspects of memorial processes ends here. The next two chapters being a review of relevant neurophysiological material are necessary if one is to understand theories - such as those of Hebb, (1949) John (1967) and Lashley (1958) that attempt a neuropsychological synthesis.

CHAPTER III

SOME NEUROPHYSIOLOGICAL PREREQUISITES

As we said earlier on, our basic task, in this work is to relate functions of the nervous system, including higher processes such as memory to structural aspects of this same system. To do this we start by presenting aspects of neurophysiology relevant to this work. Our main focus and interest are the signal processing properties of the CNS.

A. Neurons

The most fundamental structural concept, known as the neuron doctrine, is that the nervous system is built up with elementary units known as neurons capable of transmitting electrical signals. These basic units are assembled via their connections known as synapses into functional aggregates capable of carrying out the workings of our brain. Although neurons come in all sorts of shapes and sizes one can abstract and make the following statements concerning the so-called typical neuron. See figure 12.

Being a cell neuron must have a nucleus (or soma or cell body) within which the metabolic processes occur. Emanating from this nucleus one has multiple branching fibers known as dendrites and the axon with its axon collaterals. Dendrites are supposed to input information from other neurons (afferents). Axons are presumed to transmit information from the cell body to other neurons (efferents). Information within dendrites, soma and axon is conducted via electrical signals. While a given wave is passively carried within the dendrites

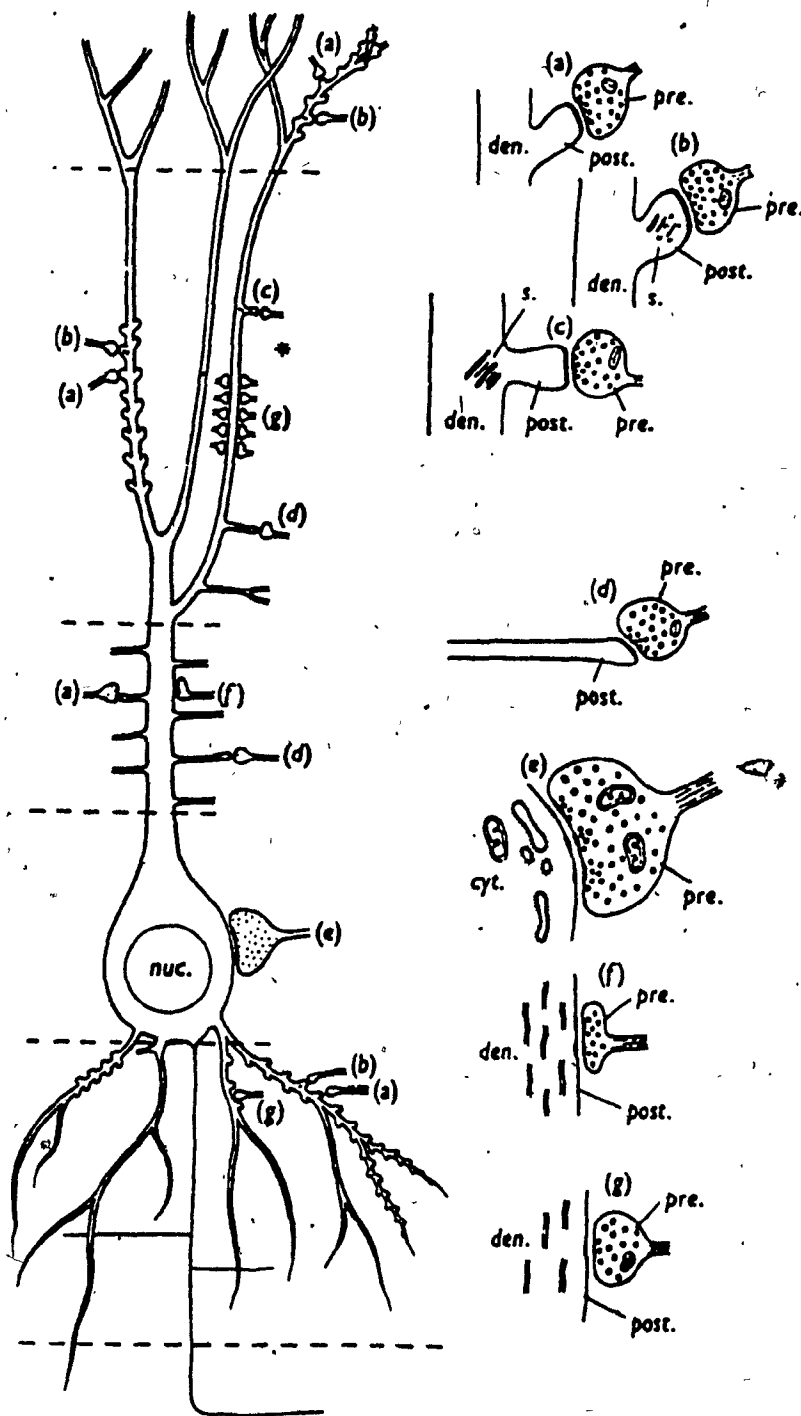


Figure 12. A neuron schematic according to Eccles (1973). Synaptic contacts are shown on the right in magnification.

and soma, a signal upon entering an axon, and assuming it has the requisite amplitude, will be converted into trains of very brief pulses, called nerve impulses, action potentials or spikes. These are propagated without attenuation along the axon and its collaterals to various target cells (e.g. to other neurons, muscles, etc.). Typically ongoing nervous activity involves the continuous spike bombardment of a given neuron via its dendrites and consequent production and sending out of other spike trains due to this neuron, via its efferent axon, to a great number of other neurons.

Surrounding the neurons are the glial cells with graded potentials of their own. The functional significance of the glial system is not known. It has been surmised (Stevens, 1966) that the glial system is the support system on which the neuron system rests. Others claim that the glial system plays some role in memory storage. The total system of neurons and glia is bathed in a good electrolyte through which both chemical and electrical signals may be sent out. These signals are presumed to be averages of the electrical activity within cell bodies.

We distinguish several types of recordings of electrical activity. On the one hand one could implant large diameter electrodes (macroelectrodes) either on the scalp or within brain tissue. He would then obtain the electroencephalogram (EEG). Microelectrodes of the order of a micrometer can be used on the other hand to pierce individual cells and axons and thus gain access to the ongoing electrical spike activity of individual neurons. The external fields of neurons can be probed by putting electrodes outside these same neurons. (MacGregor and Lewis, 1977).

B. Synapses and some of their functions in Neural Aggregations

Axons generally branch repeatedly and end with their terminals called butons on the dendrites or cell bodies of other neurons. The area of contact between one neuron and the next is known as the synapse. The architectonics of synaptic contact present a great variety as well. Of interest is the fact that in many cases the axon makes contact not with the dendrite directly but with a stumpy branch of it known as the dendritic spine. (Shown as diagram (a), (c), (d) in figure 12. Typically a neuron makes contact with thousands of other neurons (up to 10,000 in the CNS) and receives messages from a comparable number of other neurons.

As we said above many neurons are arranged into functional units with specific functions. Such aggregates may be determined by the specificity of channelling their produce within the brain whenever an electrical excitation blankets the aggregate. Another way of possible definition of a functional assembly is the fashion in which the synapses of a given cell can emit one of several chemical transmitters to various other target cells with the ensuing result that selected pathways become facilitated. The latter remark makes it conceivable that more than one assembly may be formed from a given population of neurons.

C. Information processing between Neurons

As we said before potential waves reach the soma and dendrites of a given neuron via a time series of spikes coming down an axon cable. Synaptic connections between two neurons are either inhibitory or excitatory. If the existing potential of the soma becomes momentarily depressed due to the arrival of a signal then we say that we have an inhibitory post synaptic potential (IPS). If the potential of the soma

is elevated from its base value one speaks of an excitatory post synaptic potential, (EPSP). With respect to the outside fluid in which the cell is bathed the soma has a potential of about -70 mV when it is not active. Differences in concentration of ions such as Na^+ , Cl^- , K^+ , Ca^{++} within and without the cell, maintained by a mechanism known as the sodium pump are mainly responsible for this potential difference. When signals arrive and depart from the cell the conductivity of the cell membrane with respect to penetration of various ions changes producing or "carrying" the well known potential waves. For the purposes of this paper the electrochemistry of signal generation and transmission in the neuron is of no importance and it will not be presented. (For further details see chapter 1 of Rugh and Patton, 1965).

Many neurons make contact and hence "talk" to a given neuron at a given time. The neuron must therefore relay this information along its axon. The place where various signals are "summed" is known as the axon hillock, i.e. a place on the soma near the initial segment of the axon.

If the cell has been inactive recently the incoming post synaptic potentials (PSP) waves will sum algebraically. (The summation process is known as temporal summation when all PSPs are initiated at the same synapse. If they are initiated from different synapses we use the term spatial summation.) If the summation process exceeds a certain critical value known as the threshold for firing for this particular neuron the system will "fire" and send out a potential spike. Normally more than one spike is produced. The summation process in this case is modified by the fact that after production of the first spike the threshold potential is elevated to a very high value so that for a period known as the absolute refractory period (ARP), no additional

excitation may elicit a new spike. After this time interval the threshold potential declines exponentially to its normal value.

As we said before the unattenuated signal moves along the axon much like a lit fuse of an explosive device until it reaches the synaptic end of its axon. It is there translated into a chemical message via various transmitters that migrate through the synaptic cleft between sending and receiving neuron and it is translated back into an electrical signal upon entering the reception site. According to Dale's principle (Eccles, 1964) a given cell produces a unique synaptic transmitter. (The transmitter being the chemical substratum of the translation process through the approximately 200 Å synaptic cleft). In general, the effect of this transmitter on the receiving cell and its receptors depends on the biochemical state of the receptors at that particular moment. A given transmitter may have an excitatory effect on one receptor and an inhibitory effect on another. In general, however, this is not true; a given transmitter is either largely inhibitory or largely excitatory. This is one reason why we denote particular neurons as either excitatory or inhibitory.

D. Neuronal pools

The present understanding of individual neuron functioning, however, does not lead to the understanding of brain function. This is because aggregations of neurons with definite architectonics have properties beyond those of individual neurons. (A parts and whole problem). Functional aggregates of neurons are given the generic name of neuronal pools. For example, the entire cerebral cortex by itself, of surface area 47 cm², could be considered such a pool. It contains between 10⁹ and 10¹⁰ neurons. Each pool can also be considered as made

up of different pools with more specific functions. Thus in the cortex, one first and obvious subdivision is its halving into two hemispheres. (Each hemisphere is supposed to be subdivided into 4 lobes named according to their position on the skull. This subdivision into lobes however is purely anatomic and tells us nothing about function. (Schmidt, 1978). The classical subdivision of the cortex on the basis of function is first derived from the afferent and efferent connection of its various areas. These are shown in figure 13, due to Eccles (1978). Thus the cortical areas to which sensory afferents terminate (after various relays) form the sensory cortex. Within the sensory cortex we have further subdivision into motor, auditory, somatosensory, olfactory and visual. The rest of the cortex is the nonspecific or association cortex where it is presumed "pure" thinking occurs. The association cortex in man takes up considerably more space than the sensory cortex. (Schmidt, 1978). Another classification, due to Broadman is shown in figure 14. This classification provides the basis for discussion of the functional areas of the brain. Broadman's original classification was based on his histological studies and the forty or so areas involved were thus specified on the basis of the types of neurons found in them.

Studies in human beings during operations have revealed specific functions localized in various areas of the brain. These have been mapped as well. See figures 14 and 15. In figure 14 due to Penfield and Rasmussen, the lightly shaded areas are the so-called primary sensory areas while the darkly shaded area is the primary motor area from which muscular movements may be elicited by passage of a weak electric current.

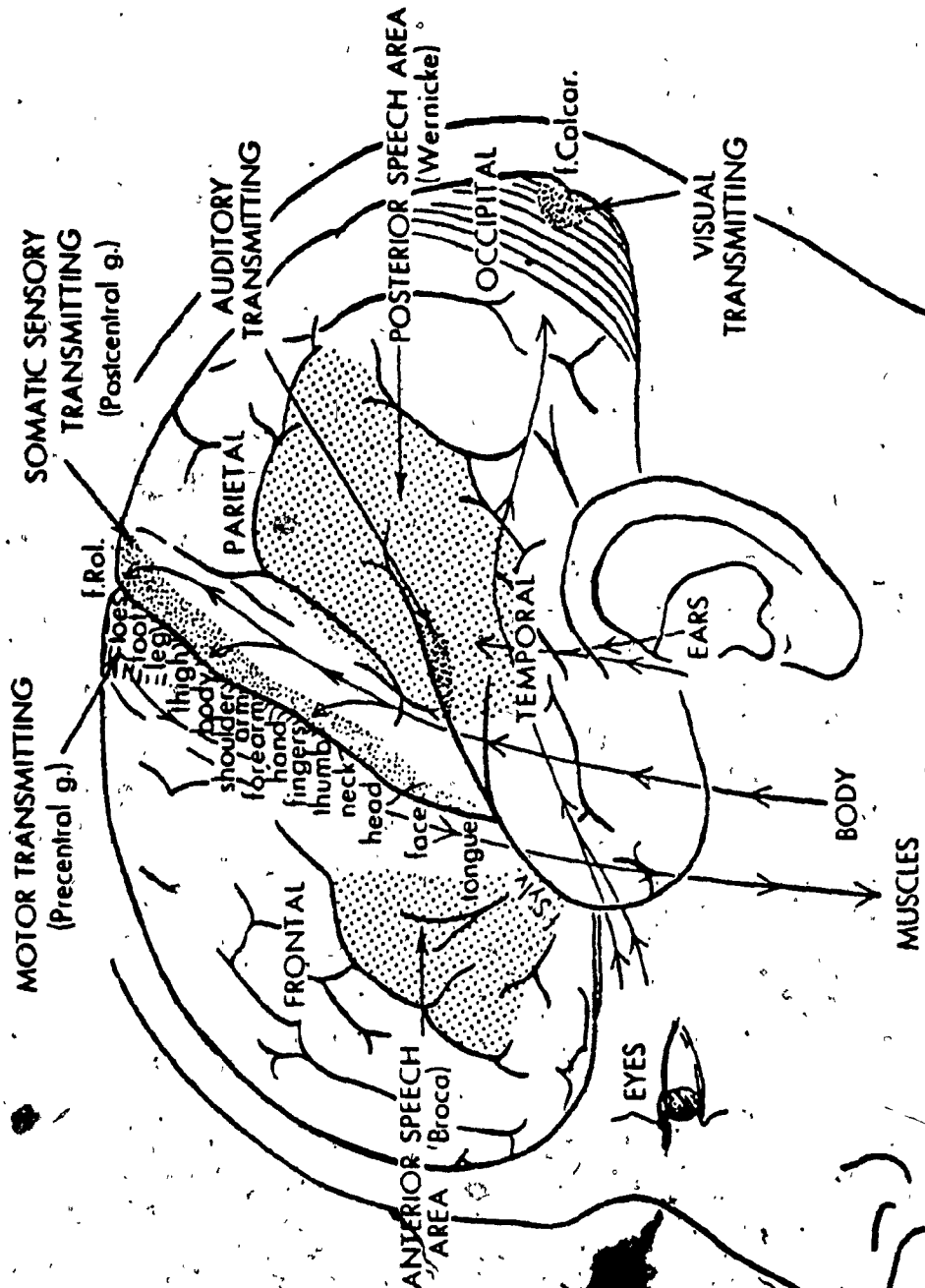


Figure 13. The motor and sensory transmitting areas of the cerebral cortex. A schematic due to Eccles, (1978)

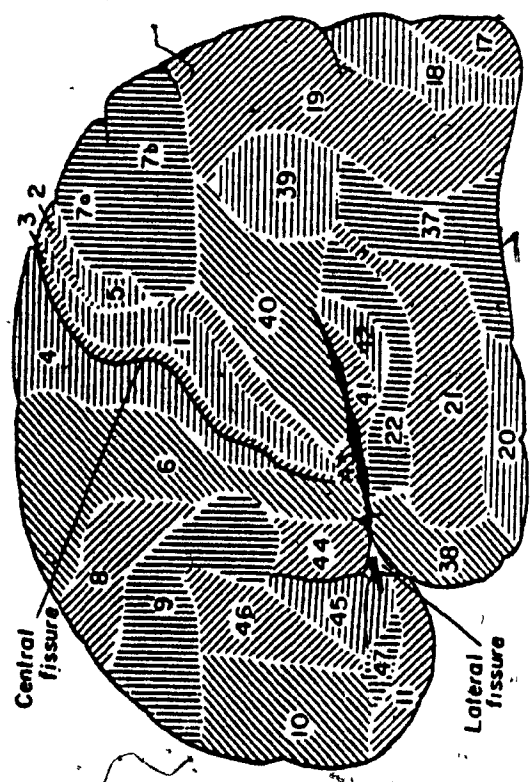
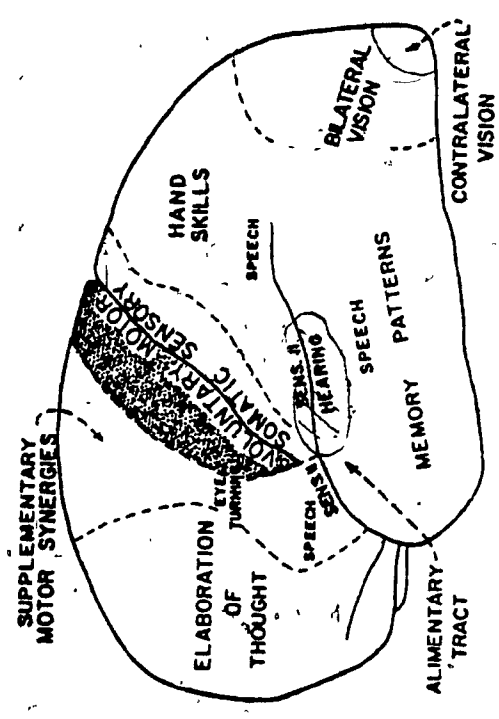


Figure 14. The Broadman classification of the cerebral cortex: (a) numbering and boundaries of various areas, (b) putative functions assigned to some areas according to Penfield and Rasmussen. (From Schmidt, 1978)

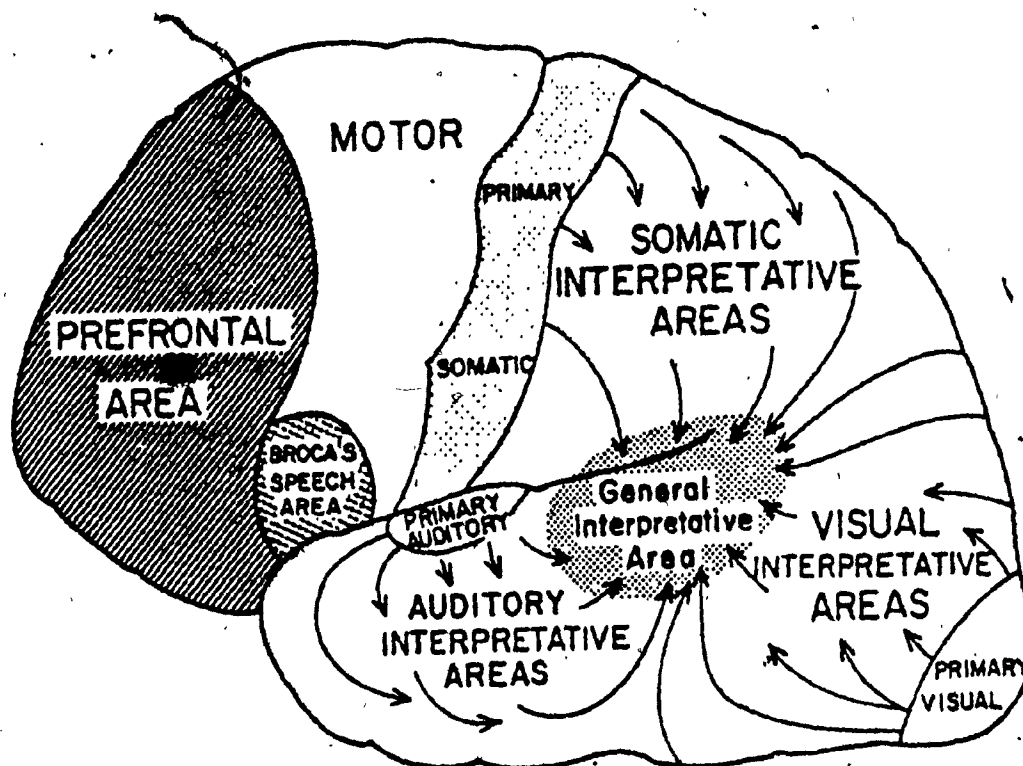


Figure 15. Gross functional division of the cortex according to Guyton, (1976). Note information flow toward the general interpretative area.

Figure 15 due to Guyton, (1976), shows the rough organization of the various association areas as determined by various methods. It should be pointed out, however, that various gross structures and nuclei of the brain interconnect and pass information from one structure to another. Figure 16, which is a schematic due to Schmidt, (1978) is a flow chart showing the basic features of information flow and control in the CNS.

Let us now, however, come back to a more topological view of neuronal pool functioning. This will make it easier to consider the transmission and processing of information in neuronal circuits.

E. Transmission and Processing of Signals in Neuronal Pools

E1. Organization of Neurons in the Neuronal Pool

Figure 17 due to Guyton shows the basic organization of a neuronal pool. To the left we have the input fibers (1) and (2) (made up of axon bundles) that transmit information to neurons (a), (b), (c) and (d). The area into which the endings of a given fiber arborize is known as the stimulatory field of the given nerve fiber.

E2. Threshold and Subthreshold Stimuli

Since only one synapse rarely fires a given neuron one must assume that a number of synapses must be simultaneously active to fire a neuron. In the Guyton diagram if one assumes that 6 or more knobs must be active simultaneously or in rapid succession, then one sees that neurons (a), (b) may be fired by fibers (1) and (2) individually. Neurons (b), (c) if excited by (1) or (2) in isolation do not fire. However their membrane potential changes (depolarizes)

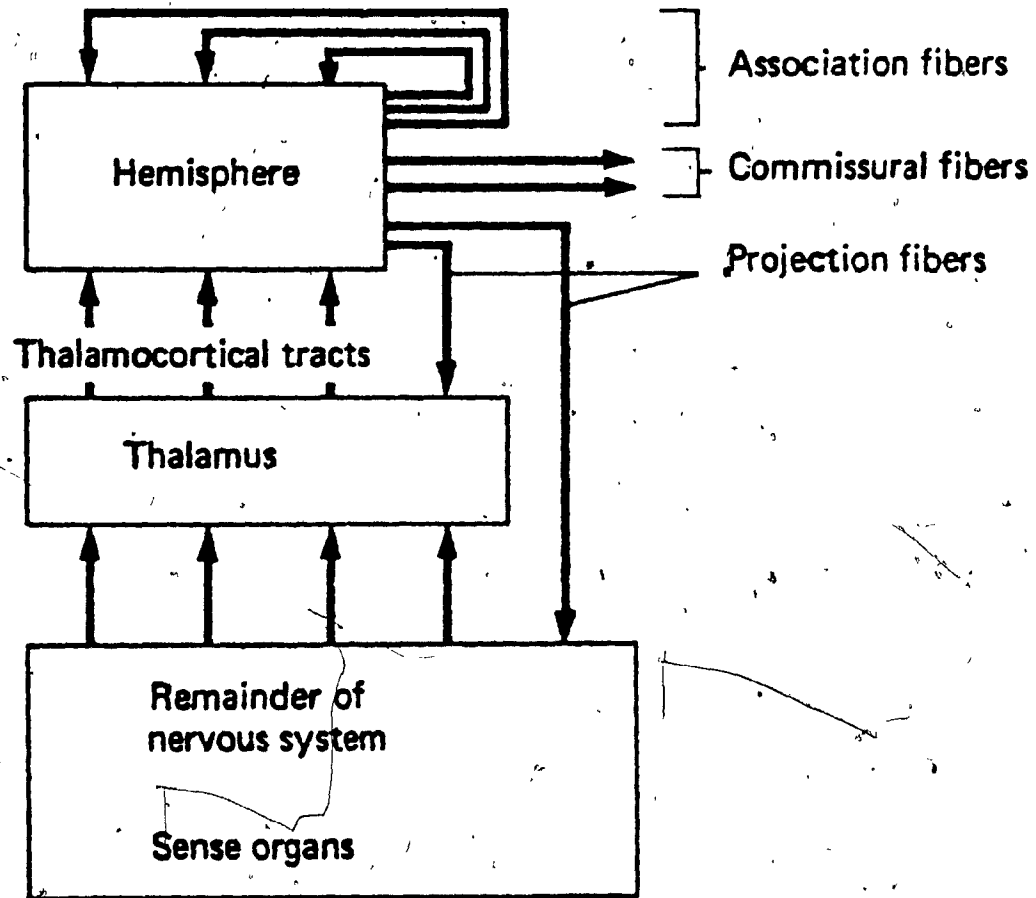


Figure 16. Rough schematic of information flow in the brain.
(From Schmidt, 1978)

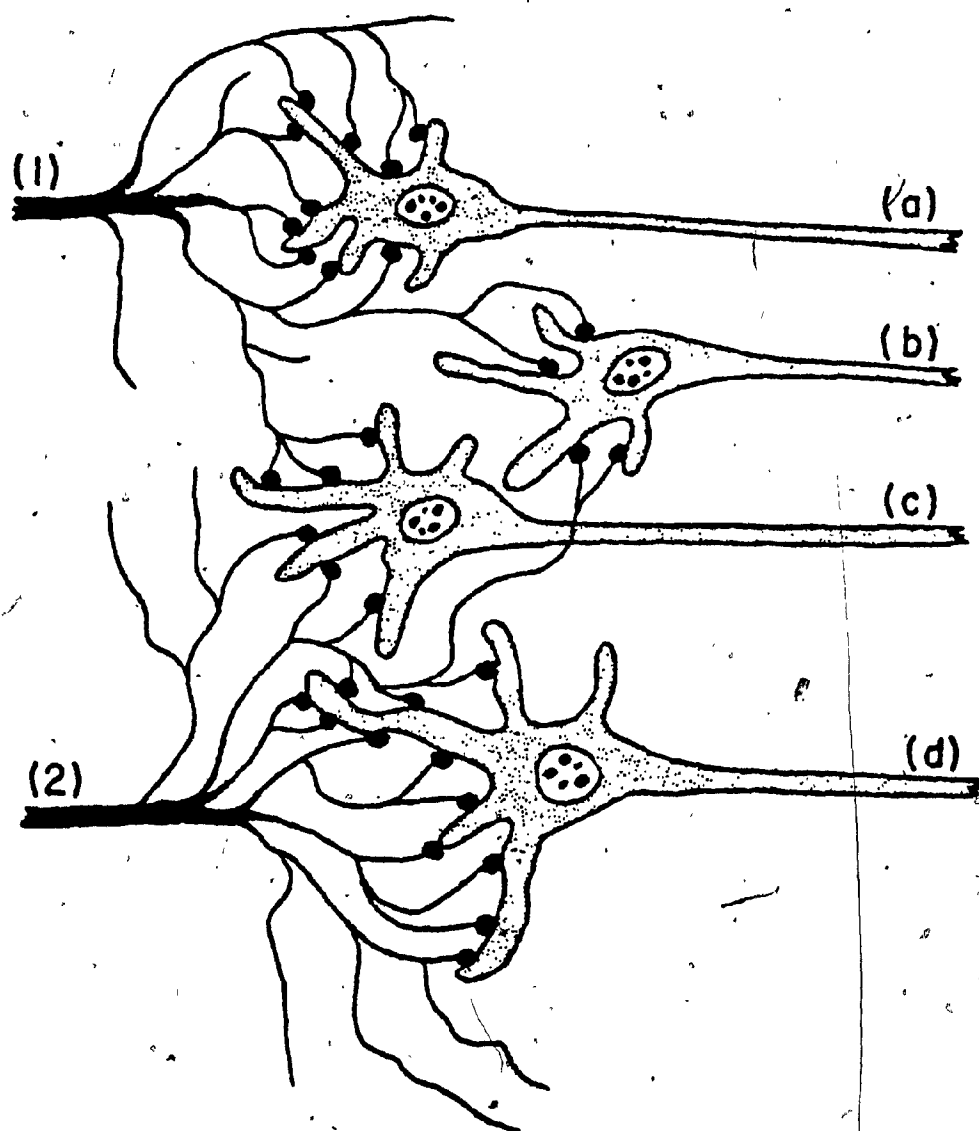


Figure 17. Schematic of a neuronal pool according to Guyton, (1976)

and if another input comes in time the neuron can fire. We describe the two interaction modes as excitatory and subthreshold stimuli respectively. The neuron that has been given a subthreshold excitation is said to be facilitated.

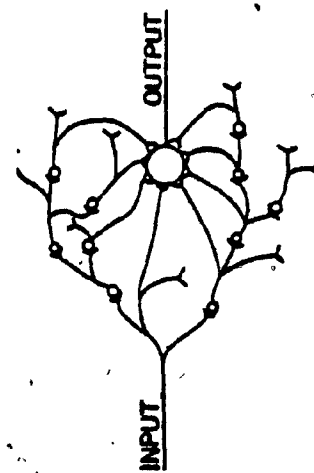
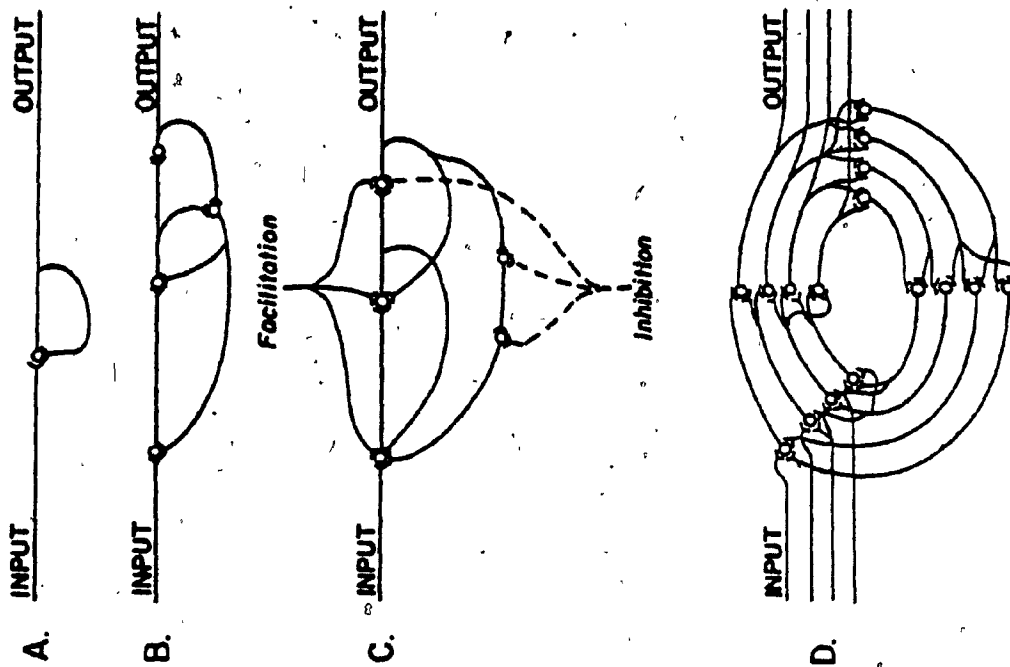
Usually each nerve fiber makes contact with about 10^2 neurons (Guyton, 1976). In general, mainly the set of neurons 'in front' of a fiber can be made to fire. This set is known as the field of the fiber. The area in the pool where all neurons can be made to discharge is called the discharge, or liminal zone. Peripherally to it one finds the subthreshold, subliminal or facilitated zone.

Most information is transmitted from one part of the nervous system to another through several successive neuronal pools.

E3. Prolongation of a Signal by a Neuronal Pool - "after discharge"

Even after a signal has entered a pool one can have prolonged output discharge called after discharge (Guyton, 1976). Of the possible mechanisms involved we shall be interested in the so-called reverberating (oscillatory) circuits. Schematic realization of reverberating pools are shown in figure 18 due to Guyton. Theoretically the simplest type of oscillatory circuit is due to A. It consists of one neuron with a feedback mechanism that if properly adjusted can keep itself firing for a long time. It is of no interest in this study. Schemes B and C are generalizations of Scheme A, with and without inhibition. Set ups of the type D (of the parallel variety) are the more common forms of reverberatory circuits. The reverberatory pathways are constituted of many parallel axonal pathways. At each cell station one sees that the terminal fibrils diffuse widely.

Figure 18. Various types of reverberating pools. (From Guyton, 1976)



E4. Characteristics of after discharge from a reverberating circuit

In figure 19 one sees what output pulse rate one can have from a reverberating neuronal pool. The input signal need last for only 1 msec. Output can last from a few seconds to even minutes. Note that both early stages of output signal as well as its last stages have rather steep slopes. This could mean on the basis of the fact that the intermediate zone is almost a plateau that the pool "codes" one signal per input stimulus. This is especially easy to see in the facilitated case. The three different curves show the obvious effects or influence of other parts of the brain on a given neuronal pool.

E5. More on the functional histology of the Cerebral Cortex

On the average, the cerebral cortex is 3mm thick. Within the cortex layers predominantly made up of cell bodies alternate with layers made up of axons. Typically we have six layers. (See figure 20 due to Schmidt, 1978). Histologic and electrophysiological experiments (some of which will be mentioned) have shown that a prominent feature of the functional organization of the sensory and motor cortex is its structuring according to the columnar arrangement. Cells having similar properties with respect to response to peripheral stimuli are "stacked" in columns perpendicular to the cortical surface. Figure 20 shows a typical layout of the cross section of levels I-VI. The superficial layers I-IV presumably serve to receive and process afferent signals. The neurons sending information from the cortex have their cell bodies in levels V and VI which are called the cites of origin of the cortical efferents (Schmidt, 1978). By stimulating and recording from very small regions of the motor cortex, it has been

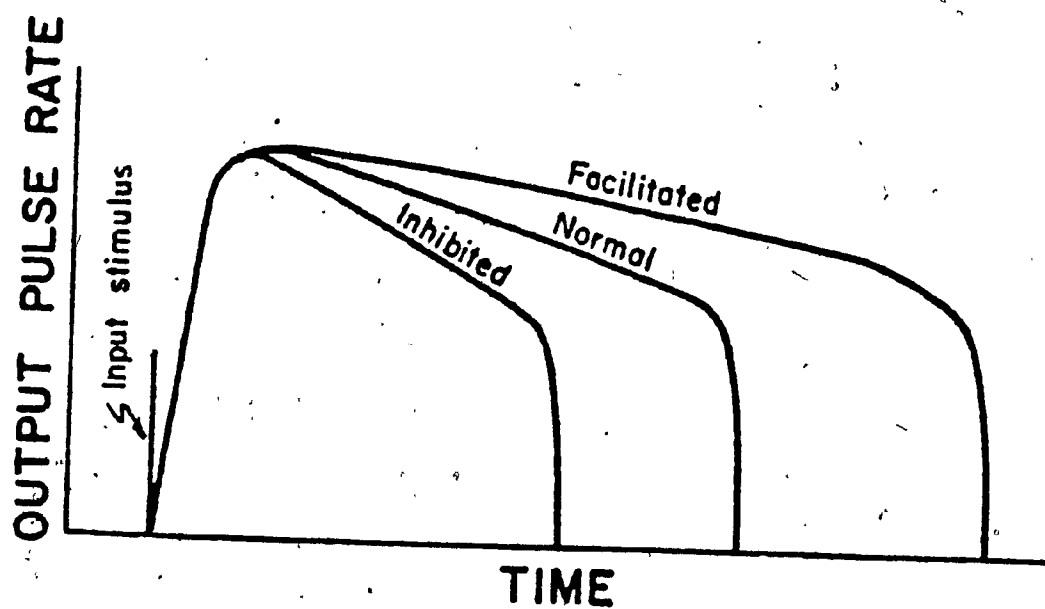


Figure 19. Pulse rates from normal, inhibited and facilitated pools. (From Guyton, 1976)

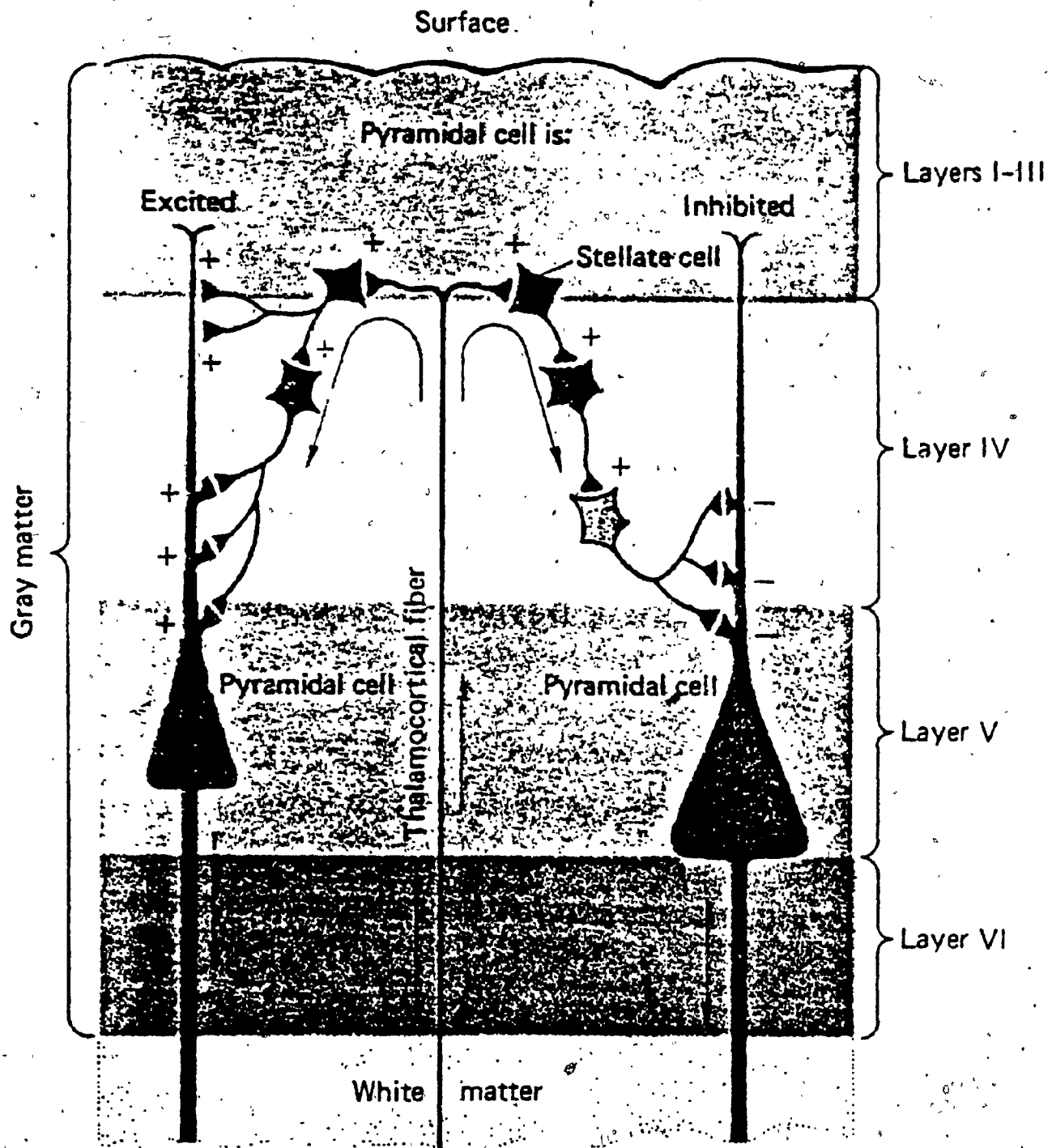


Figure 20. A schematic of layering just below the pial surface.
(From Schmidt, 1978)

shown that the functional cortical columns there have diameters of about 1mm or less.

Further, in a series of experiments over the last years Hubel and Wiesel in their attempt to investigate information processing in the visual cortex of cats and monkeys produced similar findings. They found that the neurons in the visual cortex respond to complex cognitive elements such as directed lines, edges, corners and even moving edges (Hubel and Wiesel, 1962). Further, they established that these functions, and their corresponding neural organization were present at birth rather than being the result of learning (Hubel and Wiesel, 1968).

On the basis of the Hubel and Wiesel findings, Colonnier (1966) suggested that the efferent sensory fibers may go only to some of the cells in the column, specifically to the stellate cells of layers III and IV. The uniform response of all neurons in a column would then be the result of local spreading of neuronal activity. Such radial spreading was observed directly in the work of Stoney et al (1968) by intracortical microstimulation; it was shown further that the effect of such stimulation fell off sharply with distance in a plane parallel to the pial surface.

The columnar arrangement idea has been further refined by Szentagothai. Eccles (1977) citing work by Szentagothai (1969, 1972, 1973, 1974, 1975) divides the whole of the cortex into approximately 2×10^6 functional units called modules each consisting of 10^4 neurons. It is proposed that each of these modules has in it both excitatory pyramidal cells which have the function of building power within these

units and outputting information to other areas of the cortex. Inhibitory neurons within the module tend to strongly suppress the activity of modules which are immediately adjacent. Szentagothai (1972) has generalized from the specific sensory areas, somesthetic and visual to the neocortex in general.

E6. Modules in information processing interactions

Eccles (1973) identifies these modules as substrates of conscious perception. He does not elaborate. Hebb (1949) is more specific (See below).

A schematic of the interaction process generated by triggering two initial clusters (modules) A, B is shown in figure 21 due to Eccles. Nerve cells are represented as circles solid or open, aggregated in clusters corresponding to modules.

The modules are supposed to be power units owing to the existence within them of self exciting (reverberating) chains. They can also influence adjacent columns because they also connect to them via inhibitory neurons. An excited cluster projects on to other non adjacent modules excitation and thus one can have sequential excitation from module to module. A particular sensory modality (A) is presumed to sequentially activate clusters identified by solid circles. Sensory modality B represented by open circles triggers chain B of open circles. Finally, convergence of the transmission of the A and B modalities gives activation of clusters of cells with modality reference corresponding to both A and B and so are shown with a symbolism (AB) (Eccles, 1973). The confluence AB would correspond to the recognition phase of an object requiring for example two sensory

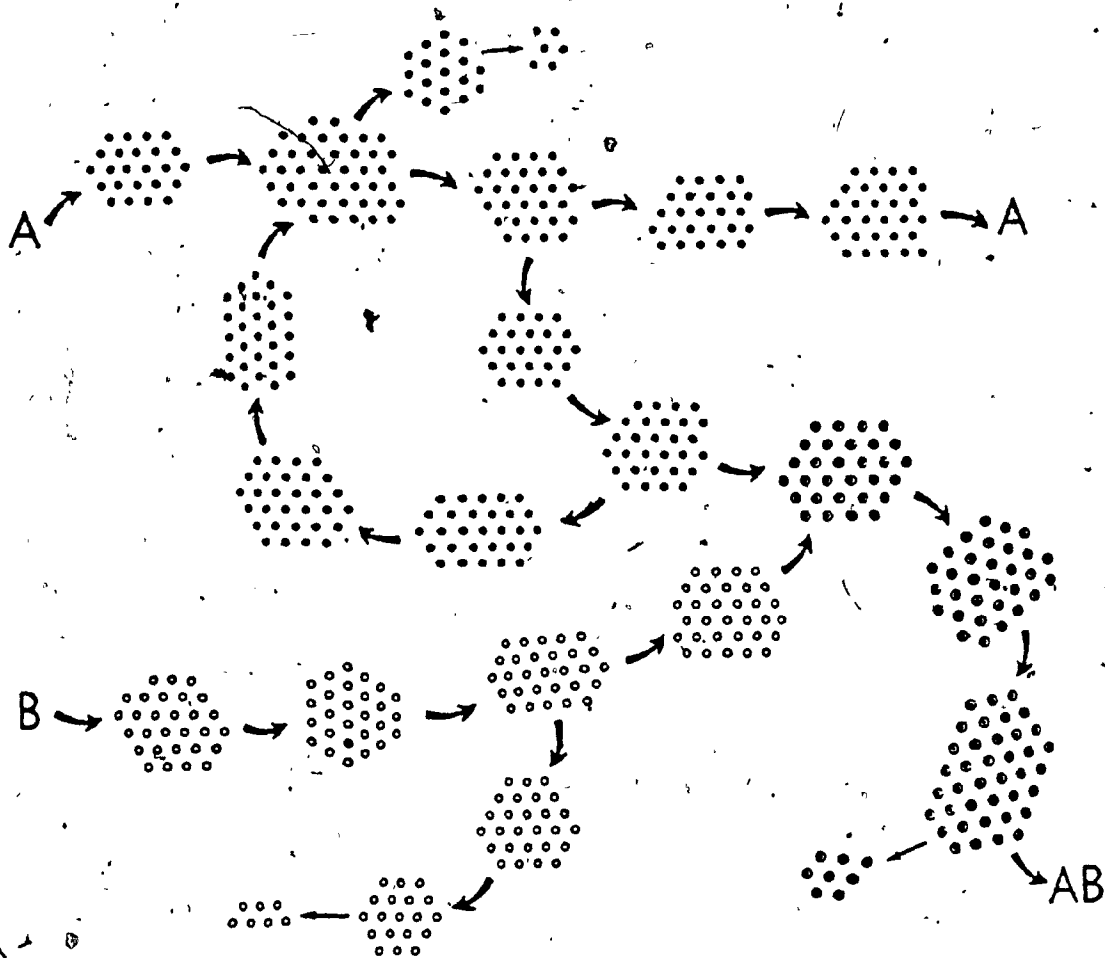


Figure 21. Schema of the cerebral cortex looked from above. Large pyramidal cells are represented by circles solid or open. Clusters represent modules. The large arrows symbolize impulse discharges along hundreds of lines in parallel which are the mode of communication from column to column. Two inputs, A and B, and two outputs, A and AB, are shown. (From Eccles, 1973)

inputs in combination for identification. Eccles (1973) is very cautious; he states that the diagram is only a schematic and that no identifiably specific area of association cortex is responsible for any specific awareness of some arbitrary event.

This last remark brings us to a very controversial area. For obviously whenever one is aware that he is looking at an apple, say, some set of neurons must be the substrate of the memory process involved in recognizing the apple. The question is whether this set becomes excited whenever another apple comes to view later on.

John (1967) supposes that the coding of information internalizing a given event corresponds to activation of a large number of diverse assemblies. For any given assembly what matters is not which particular neurons have been excited but rather how many i.e. what counts is the activity level associated with the event. John further elaborates his theory based on his findings (1972) that diverse structures and nuclei in the brain seem to produce similar wave forms of activity whenever the brain is occupied by some particular task (given stimulus context). He notes that although individual cells display a variable firing pattern in response to conditioned stimuli ensembles of neurons display essentially invariant temporal patterns. (microelectrode recordings vary, macroelectrode recordings do not vary per task.)

However it must be pointed out that the statistical versus connectionistic dilemma can be circumvented. In our theory where nets (standing for modules) are the units of organization one can have determination at one level and randomness at another.

If one says: event e corresponds to a vector $\langle \alpha_1, \alpha_2, \dots, \alpha_n \rangle$ where α_i are activities in determinate nets (modules) one has determinism. On the other hand different permutations of neurons within a net can combine to produce a given activity, thus generating macrocanonical degeneracy and probabilism.

CHAPTER IV

ASPECTS OF THE NEURAL BASIS OF MEMORYA. Introduction

Having described overt aspects of memory behaviour as well as neurophysiological facts of interest the ground is now ready for a consideration of the kind of relationship that could exist between the two. It must be stated at the outset that so far, one has various theories concerning this relationship. The arguments, however, we will be making in this work will be of a consistency type rather than of a full analytical one. Let us therefore take a look at some of these theories.

B. Neural Circuits of Memory

All present studies on the neurophysiology of memory seem to agree that without chemical or structural change in parts of the neural mass no retention of information is possible. The details of the ontogenesis of neuronal memory circuits is usually where divergences appear. Further, in addition, most workers would state that conscious experience and short term memory are mediated through electrical activity. This is not to be taken to mean, however, that when one is not conscious, as when one is sleeping, no electrical activity goes on. One has electrical activity as evidenced by the electroencephalogram. However, the profile of the EEG is different for different states of CNS activity.

One of the central problems of present day research in the neurophysiology of memory is the interface between structural and/or chemical changes in neuronal masses leading to neuronal circuits able to sustain specific electrical activity (waves). Since specific signals are presumably the carriers of information one is further faced with the problem of constructing neuronal circuits capable of (a) carrying and (b) reproducing i.e. having a memory capability for a given signal.

The problem is akin to that of a communication engineer who is faced with similar problems as the construction of communication systems capable of transmitting, coding and decoding information. The analogy would even be closer if one could consider the communication process between two computers or between a production system and its cybernetic control. Without memory systems whose substrates are the hardware of the system neither communication nor control would be possible. Just as a computer engineer is faced with problems of memory storage (and hence program complexity) requiring specific characteristics and memory sizes in the same way neuropsychologists attempt to construct neuronal circuits capable of memory. These constructs however cannot be arbitrary. Due regard must be given to the electrochemical properties of the individual building blocks of these systems i.e. the neurons.

Most workers, however, consider the problem from the point of view of neuronal pools rather than from the point of view of individual neurons. The most basic requirement of a memory circuit is that the effects of an input persist after stimulus offset since this is the crux of any definition of memory for an event.

Lindsay and Norman (1972) list three approaches to the solution of this problem.

B1. The Unique Cell Theory

This theory assumes that there exists a unique cell that corresponds to each object we remember. This cell is 'stamped' with the image of the object. In a recognition test, for example, the mechanism that comes into play is some sort of template process analogous to that used by computers in pattern recognition schemes. The basic neurological inspiration for this hypothesis seems to have originated with the work of Hubel and Wiesel who showed that a hierarchy of feature extractors exist in the visual cortex. The vast complexity of memory structure makes this 'stamping' theory quite inadequate. For these reasons and the fact that this theory is too much of an ad hoc nature, very few people give credence to it.

B2. The Unique Pattern Theory

The unique pattern theory maintains that an object is remembered whenever a determinate set of neurons become active. A typical theory of this kind is that due to Hebb (1949) which we shall explore in some detail in the next chapter and a generalization of which will be the basis of this work. Without going into details, at this point, it must be stated that this theory has behind it a great deal of analytical work. Its assumptions, especially those relating to the ontogenesis of the memory circuits have been the inspiration for a great deal of both theoretical and experimental work. It is the theory accepted by most brain modelers.

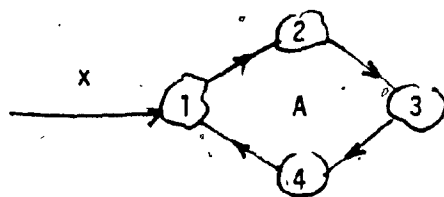
B3. The Unique Code Theory

The unique code theory assumes that each stimulus is represented by a unique signal the electrical code of which in the CNS is also unique. It assumes chemical and histological permanence in the brain,

and no changes as a result of learning and experience. The neural mass can remain, in other words, undifferentiated as a result of experience. Further, the same set of neurons can carry more than one memory since a great number of codes could ride on the electrical activity of a given set. In addition, memory could reside anywhere in the CNS since what counts is the signal and not the neural substrate that carries this signal. This is a non local theory Pribram (1966, 1969) is one of its major exponents.

C. Reverberating Circuits

Pursuing our line of thought outlined in the previous paragraphs one might now ask what kind of structural neural arrangement could produce the activity and function that is necessary in the act of remembering. Neuropsychologists such as Hebb (1949, 1972) and psychologists such as Lindsay and Norman (1972) usually invoke the reverberating circuit idea. Ramon y Cajal had, as long ago as 1911, noticed their existence in the cerebral cortex. Consider now how a reverberating circuit might remember: (See schematic below).



Let X be the route through which some signal S arrives. We say that the group of neurons A1, A2, A3, A4 reverberates provided once A1 gets triggered the cycle A1, A2, A3, A4, A1 follows. Thus reverberation in loop A is said to represent event S. More accurately

reverberation in A represents an electrical memory of the fact stimulus S came along route X. As we saw in the previous chapter there are several mechanisms whereby this reverberation might cease. The simplest is obviously inhibitory input to any part of this chain.

D. Consolidation

If whenever signal S coming along X, invariably excites the neural chain A we say that permanent memory for the event S has been created. Most workers attribute the creation of chain A to the possibility that repeated stimulation of the chain leads to chemical and/or structural changes at the synapses joining the various member neurons of the chain. The process of thus altering the chain is known as consolidation. Consolidation of specific circuits occurs over a period of time. Consolidation implies a local theory of memory since specific memory circuits must carry specific memories. Electrical activity through stabilized chains represents a temporary activation of these circuits. This temporary activity is what we know as short term memory. (Lindsay and Norman, 1972). Long term memory is attributed to the stabilized chain itself. In other words Lindsay and Norman (1972) consider short and long term memory as attributes or properties of a given, determinate neural mass. Short term memory is a functional-dynamic aspect of an electrical nature whereas long term memory is a structural property of this same mass.

D1. Synaptic changes due to Learning

The last paragraph brings us to the question as to the type of mechanism that operates in order to produce aggregations of neurons that are stable enough to carry permanent memory. The natural place

to situate the change is the synapse since it is there that electrical transmission may either pass or not pass to the next member of the chain i.e. the neuron next in line. The two possibilities usually considered are: chemical encoding or growth of new synaptic junctions. Further, we must not forget our original hypothesis, namely, that these changes occur as a result and during the electrical activity of a chain of neurons.

The most abstract statement one can make about these changes is due to Hebb (1972, p. 64): "When an axon of a neuron x is near enough to help fire neuron y and does so, a change takes place such that x becomes more effective at exciting y ". Put in a simpler language and since more than one input is required to fire a given cell we can rephrase the above to read as follows: If neurons x , y are simultaneously active and they are in contact with each other than repeated activation of the pair x , y would tend to make the pair adhere. It is obvious that this mechanism could be invoked until a complete chain is formed.

The specifics of synaptic plasticity especially in the adult human brain are at this stage not completely elucidated.

Hebb (1972) hypothesized that the change would be chemical ---- more specifically an increased ability of the presynaptic part of the synapse to produce transmitter substances such as ACh. He also speculated on a possible increase in contact area in the synapse.

Today, we have at hand evidence sustaining Hebb's hypothesis. The most interesting work to this writer's knowledge is the work of Bourdois, McCandless and MacIntosh (1975). In a series of experiments in the cat superior cervical ganglia conditioned by prolonged

preganglionic stimulation at high enough frequency to cause initial depletion of the store of acetylcholine (ACh), they found "rebound" increase of stored ACh. They found that a conditioned ganglion can release, as well as store, more ACh than an unconditioned one.

"During and after conditioning, changes in the level of ganglionic ACh occur in the following sequence: (1) a fall by as much as 30-40% over the first few minutes of stimulation; (2) a gradual return to the initial level and beyond it, while stimulation is continued; (3) a further "rebound" rise up to 50-70% above the initial level, within minutes after stimulation is stopped; and finally, (4) a slow return to the initial level, which is still incomplete after 4 hours of rest. Phases 2 and 3 of the sequenceare both suppressed by hexamethonium, and presumably they reflect the same process, namely, the creation of new storage capacity for ACh. The new storage sites must be presynaptic." They conclude that "Perhaps such a 'trophic' effect can only occur when presynaptic and postsynaptic activity are greatly intensified at the same time."

In an article concerning possible synaptic mechanisms subserving learning, Eccles (1972) mentions Ramon y Cajal (1911), Loennies (1949), Young (1956), Eccles (1953, 1961, 1966) and Kandel and Spencer (1968) as people who accept changes in the efficiency of synapses as the foundation of permanence of knowledge (learning). "There is further, the complementary concept that diminished usage leads to regression in effectiveness and thus an explanation of forgetting is hand". A great number of workers is engaged in synaptic plasticity research today.

Two extensive review volumes and a textbook on the subject are of interest in this regard. Horn and Hinde (1970) and G. Gottlieb

(1976) published the review volumes. Lund (1979) is the author of the textbook. Kandel (1970) divides the theories concerning plasticity into two basic groups. He first poses the problem in its paradoxical aspect:

"If the development of connections between most neurons in the nervous system is rigidly determined, how then is behaviour modified? How does one reconcile the known malleability of behaviour with preprogrammed and rigidly 'wired' nervous system?... Do memory and learning require some further additions to the wiring diagram?

The two (solutions) that have proved most interesting experimentally are based on notions of the plastic and dynamic capability of neurons. The plasticity hypothesis was first put forward by Cajal and several other neuroanatomists and then in more modern form by two psychologists, Jerzy Kornorski in Poland and Donald O. Hebb in Canada. A current version of this hypothesis states that even though the anatomical connections between neurons may develop according to a rigid plan, the strength or effectiveness of the connections is not entirely predetermined and the effectiveness of synapses and other properties of neurons can be altered by experience. This hypothesis predicted that neurons, and in particular their synapses, should be able to change their functional properties as a result of altered activity.

The dynamic hypothesis... of Lorente de No derives from his observation that neurons often form closed chains. Neural activity could therefore be sustained by the circulation or reverberation of impulses within a closed chain of interconnected self-exciting neurons. This again would not require anatomical change. In fact, the hypothesis does not even require functional change in the properties of neurons."

The second possibility (in its pure form) as Kandel points out is untenable. If, for example, a number of inhibitory neurons find themselves in the chain, activity would cease. Since memory survives cooling of the brain, and since its activity is not then possible, one must assume that its memory should reappear upon warming up again.

In 1947 Larabee and Bronk of John Hopkin's showed that if a pre-synaptic neuron leading to the stellate ganglion in the autonomic nervous system of the cat is stimulated at high frequency (tetanized), then the postsynaptic neuron would show an increase in its postsynaptic potential even if it is normally excited 90 seconds after cessation of tetanus. They called this phenomenon posttetanic potentiation. Kandel mentions the work Wigdor, Beswick and Conroy who have even found that posttetanic potentiation can last for as long as two hours after a period of tetanization lasting from 15 to 30 minutes.

Of interest to subsequent discussion is Kandel's heterosynaptic facilitation, i.e. the EPSP's produced by stimulation of one pathway would be greatly facilitated by activity in another pathway.

Eccles (1972), in the paper mentioned earlier, commenting on various frequency and posttetanic potentiation experiments claims that although this ability of potentiation does not directly prove that memory has been stabilized it does suggest that some residual potency would remain. This potency could then be the basis for memory acquisition of the given circuit.

Of special interest to Eccles (1972) however, is the work of Lomo with the entorhinal pathway into granule cells in which potentiation is also inhibited. They are in a part of the central nervous system that is believed to be importantly concerned with memory.

The synapses involved in these experiments of Lomo are the spine synapses on the dendrites of granule cells. "It is now generally recognized that the spine synapses are almost uniformly excitatory in character...and furthermore, there is now much evidence

to show that these synapses are plastic in that they exhibit regression and development to a quite remarkable degree under appropriate experimental conditions. A pyramidal cell is estimated to have at least 10,000 such spines and these provide powerful paths for synaptic stimulation (Eccles, 1972, p. 51).

In Lomo's experiment disuse led to a striking depletion of dendritic spines. Eccles (1972) making the converse argument then argues that use would indeed lead to an increase in spine population. He then brings in the molecular theory of memory by assuming that this growth must come about as a result of the operation of the appropriate metabolic machinery. He cites evidence due to Morell (1961) and Hyden (1964) who showed that there is an increase in the RNA content of neurons which are subjected to excess stimulation. Presumably the synaptic growth theory of learning is consistent with RNA increase in the neuron since this would be responsible for the manufacture of the necessary protein that would make for an enlarged structure.

This writer agrees and would perhaps put the matter in a slightly different fashion:

.. Growth occurs in space-time. Consequently, if more ACh is to be released from different sites as the Bourdois et al paper would imply, that is to say if more sites must be created one must occupy more space for them, i.e. the "bouton" must "bulge" more. Similarly, if more macromolecules are to be found residing (or constituting) on "learned" synapses the assumption must be that the packaging process would produce a bigger geometrical aggregation with more rather than with less micromolecules.

To put the matter in another fashion: with a given thickness of wire a transatlantic cable can carry so many messages. If you want to carry more messages you have to add more wires and thus make the cable thicker. Just as you cannot identify a given wire with the message it carried so you cannot identify memory with specific macromolecules.

It is, further, unlikely that permanent memory can be explained by potentiations. The time span of potentiation is very small (many hours) compared with the permanence of some memories that go on for years.

All that these studies show is that synapses are modifiable. Speculating, this writer would say, that perhaps the process of consolidation comes about because stimulation produces in the neighbourhood of the synapse the appropriate electromagnetic environment in which chemical reactions can occur of the kind that lead to the aggregation of macromolecules and thus to more stable structures. The model I have in mind is what happens as an electromagnetic wave passes through an organic solution assuming that crystallization centers can exist on which aggregation can begin. In any case no change in the nervous system is not compatible with permanent memory, in this writer's view.

That some such mechanism seems to be at work is suggested by two recent pieces of research. Browning et al at the University of California (1979) repetitively stimulated (100 pulses per second for 1 second), the Schaffer collateral commissure system of the rat hippocampus. This induced long term potentiation of synaptic strength and produced significant changes in the subsequent endogenous phosphorylation of 40,000-dalton protein from synaptic plasma membranes. They suggest that these findings provide evidence that repetitive

synaptic activation can alter the phosphorylation machinery of the synaptic region revealing thus the biochemical substrate of synaptic plasticity with its putative relation to learning and memory.

Indirect support, however, for the above is also implied by experiments due to Rodan et al (1978) who examined the effects of external oscillating fields (1166 volts per centimeter, 5 herz) on cartilage cells isolated from the proliferative zone of tibia epiphyses from 16-day thick embryos. It was found that such treatment enhanced the incorporation of (^3H) Thymidine into the DNA of chondrocytes. In other words these findings would support the hypothesis that Na^+ and Ca^{2+} fluxes generated by the electrical perturbation trigger DNA synthesis in these cells.

Although the second piece of research is not related to the CNS it is quite suggestive. The last remarks end our brief review of the vast subject of the neural basis of memory. The basis purpose of this, rather short set of remarks, is to prepare the ground for Hebb's ideas to which we now turn.

CHAPTER V

HEBB'S THEORY OF CELL ASSEMBLESA. Introduction

Hebb's work to which we now turn in considerable detail is an attempt to solve in neurophysiological terms the problem of representation of knowledge as posed in chapter II. The original work dates at least as far back as 1949. Aspects of the Hebbian schema are accepted by the majority of the workers in neural modeling. MacGregor and Lewis (1978) find Hebb's work quite fruitful and urge a methodology of neural modeling suggested by Hebb's ideas.

According to Hebb, (1959, p. 632) learning should be the crucial question in psychology theory. Hence, the focus of his theories are three basic psychological categories: (i) perceptual learning and generalization, (ii) memory and, (iii) attention. His basic method of attack was to search through the available physiological and neurological data for possible substrates that could carry the psychological process. The physiological possibility of reverberatory circuits had been established by Lorente de No (1938a, 1938b). He then asked himself how these nets might be set up. In other words, how and why would two cells in the cortex become effectively associated. A paper by Arvanitaki (1942) demonstrated that contiguity alone will permit excitation aroused in one cell to be transmitted to another. Now since it usually takes at least two inputs to fire a given cell, Hebb makes the following statement (1949, p. 65) "In the intact nervous system, an axon that passes close to the dendrites or body of a second cell would be capable of helping to fire it, when the second

cell is also exposed to other stimulation at the same point". In his psychology textbook, (1972, p. 64) he repeats his basic idea in the following abstract form. "When an axon of a neuron x is near enough to help fire a neuron y and does so, a change takes place such that x becomes more effective at exciting y." Hebb called the reverberatory circuit that can be set through the operation of the above principle the cell assembly. This according to Hebb (1959, p. 628) is the brain process which corresponds to particular sensory events, or a common aspect of sensory events. Since the assembly is a closed system, activity can reverberate in it even after the event which started it has ceased. He thus produces a memory circuit.

Further, this assembly can not only be triggered by an external sensory stimulus, but it can also be switched on via another assembly. The activity of an assembly when active via the second mechanism is (called) the image or the idea of the original stimulus. The formation of connections between assemblies is the mechanism of association.

A cell assembly might consist of from 25 - 100 neurons and its build up would require many repetitions of the original stimulus. (Hebb 1959, p. 628).

B. Mechanisms of Learning and Development

B1. The Growth of Cell Assemblies and the Infant's First Learning: Introduction

Hebb distinguished two basic stages in mental growth. The first stage involves the ontogenesis of assemblies and it corresponds to the infant's first learning. This he terms perceptual learning because it

enables the baby to perceive better and to develop a capacity for images and ideas. To this stage also belongs the period of creation of many associations which at a later stage become crucial for the process of thinking. The second stage corresponds to the dynamics of the stabilized system of cell assemblies as they function in an adult brain. This section devotes itself to the first stage whereas the next sections illustrate aspects of function of a developed cell assembly systems.

B2. Perceptual learning in infants: General Considerations

According to Hebb (1972, p. 29) perceptual learning can be defined as a lasting change in the perception of an object or situation resulting from repeated exposure to the same or similar events. Primarily the change is toward a clearer perception of some aspect of the given event.

Hebb, (1972, p. 66) to illustrate the above, considers some fairly common events in the life of an infant. Events in other words that occur, repeatedly such as sight of mother's and father's face, milk-bottle, taste of milk, etc. Each of these events according to Hebb gives rise to an excitation of sets of neurons in the CNS. But the neurons that get excited whenever mother appears, for example, are not always the same. However, there is always a common core of neurons that does get excited when the aforementioned event occurs. In the language of set theory, if $A(t_i)$ is the set of neurons excited upon presentation of a given stable stimulus S in time t_i ,

then

$$A = \bigcap_{i=1}^N A(t_i) \neq \emptyset.$$

The more S appears the more this core becomes intertwined owing to Hebb's postulate. Further, since infants can also be trained via association A may also be triggered as a result of presentation of another stimulus S'. This increase in organization and stability of A making it capable of reverberating electrical activity (and hence memory) even after stimulus offset, Hebb called perceptual learning. Theoretically, it leads to increased clarity and distinctiveness in perception. Anticipating further discussion we can in addition claim that perceptual learning can also be classified as conceptual learning since a stable assembly when active can become the source of activation of another assembly to which it is connected. While a cell assembly is excited by its own sensory stimulation the activity of the cell assembly is perception; if it is active after stimulus offset or if it is excited by another cell assembly and not sensorily, the activity is imagery or ideation or a mediating process. (Hebb, 1972, p. 66).

B3. A Closer Look at the Semiprobabilistic Mechanism of Single Assemblies

The effects of repeated stimulation as stated above are obviously derivable from his postulate that if two neurons or groups of neurons are near each other and if they are simultaneously active they tend to become connected. The setting up of cell assemblies is semi-probabilistic and the cell assemblies constructed have a high degree of redundancy, a fact which provides an answer to Lashley's objections concerning localization. The argument is essentially like this:

The external world is given and stable to a high degree, hence the stimulus delivered is constant. (A moving person or object retains

its identity during motion.) Secondly, a part of the sensory apparatus up to (for example) the sensory cortex is completely deterministic. It is only in the association cortex that divergent conduction occurs. Here is where the probabilistic considerations come into play. Let us recall his line of thought concerning the fate of a visual signal as it leaves the visual projection area 17. (Pertinent areas of the brain have been mapped and numbered by Broadman. Besides the visual projection area we have: the parastriate area, area 18, which surrounds area 17 in a concentric band, and to which area diffuse connections are made from 17; area 19, the peristriate area, which is a second, more anterior concentric zone, and to which connections run from 18, and area 20 in the lower part of the temporal lobe, which may also have connections from area 18.) As Hebb remarks (1949, p. 65-70) "all topographical organization in the visual process seems to have disappeared." And yet a 3 year old can recognize a square! What happens? His answer: The pattern of excitation of the square falls repeatedly on area 17. Repetition of this projection leads to repeated use of certain paths leading from 17 to 18, 19 and 20. He does not say that this happens all the time. If we think probabilistically, we can see the argument:

Suppose, a, b, c.....z are neurons that belong to 17, 18 19 and 20. Further, suppose that these happen to be excited upon presentation of the stimulus. Suppose a is the specific neuron in the visual projection area from where the path begins. Upon presentation for a second time of the same stimulus at a, the chances that it will connect with b (let us say b is in 18) are higher than chances that it will connect with some other similar neuron. (If we assume his postulate.) One can therefore operate the induction

machine and the law of probabilities and construct a closed path. (A closed path will be formed if the chain has N neurons that are simultaneously active. They have by our postulate, a higher probability to make connections among themselves than with other neurons.)

When certain determinate pathways in areas 18, 19 and 20 are traversed, something particular is "seen". Seeing had not occurred before because either (1) the firings associated with a particular path were too weak, or (2) the paths associated with a particular object had not formed. With regard to (1) Hebb's notion of recruitment is significant. When a pathway is used repeatedly, its electrical activity causes synaptic growth in the neighbouring pathways. The overall effect is to make "thicker" pathways.

In this context, the ideas of Kandel (1970) regarding hetero-synaptic facilitators are of interest. Since a given neighbour of the pathway could be "primed" by activity in the assembly, additional activity from some other source could fire this neuron and thus switch on either activity into the original assembly that originally primed it, or switch activity into another assembly.

B4. Establishing the Network of Associations

Not only common excitation of neurons can lead to the formation of a cell assembly but going one level of structure higher Hebb (1972, p.66) says that cell assemblies that become active at the same time tend to become interconnected. To continue with the example we used in the previous paragraph:

• "When the baby hears footsteps, let us say, an assembly is excited; while this is still active he sees a face and feels hand picking him up, which excites other assemblies - so

the "footsteps assembly" becomes connected with the "face assembly" and the "being picked up assembly". After this has happened, when the baby hears footsteps only, all three assemblies are excited; the baby has something like a perception of the mother's face and the contact of her hands before she has come in sight - but since the sensory stimulations have not yet taken place, this is ideation or imagery, not perception. According to the work of the great Swiss psychologist Jean Piaget, ideation can be detected about the age of four to five months. At this point, the baby is ready for the development of a wide network of associations - he is no longer limited, in his brain activities, to ones that are excited by here-and-now sensory stimulations - which (again theoretically) must help in the formation of new associations".

So, as far as Hebb is concerned, adult is distinguished from infant thinking by the degree of ideation an adult does. This ideation is possible because (a) the cell assemblies have become stabilized and (b) one cell assembly can trigger another and thus a train of thought occurs.

C. Structure and Ontogenesis of Perceptual Learning

C1. Introduction

According to Hebb our perception of the external world is a learned process. Perception itself is a complex of unitary processes whose neural substrates are cell assemblies that through the course of time became fused owing to the mechanism of association. Although early stages of the ontogenesis of cell assemblies built-up consisted in producing 'islands' of cell assemblies here and there, later stages involved more intimate interconnections between different cell assemblies. This process of perceptual learning does not only occur in young children but throughout one's life since according to Hebb there is no limit as to the number of cell assemblies that can be joined together. It is, in addition, obvious that higher order cell

assemblies (in the sense that they operate on higher and higher integrative levels) can be formed. This would correspond to the ability of the CNS to abstract from the specific to the general. Thus Hebb describes the process of perceptual learning of a square as follows: Hebb (1949, p. 32)

"The course of perceptual learning in man is gradual, proceeding from a dominance of color, through a period of separate attention to each part of the figure, to a gradually arrived at identification of the whole as a whole: an apparently simultaneous instead of a serial apprehension.the perception of square or circle is slowly learned and depends originally on multiple visual fixations."

Hebb was lead to the conclusion that perception of the external world is a learned process by reading the observations of Von Senden regarding visual perception in recently operated blind people. These observations were recorded in the late thirties. More recent results corroborating Von Senden's are reported by London (1960). The basic import of their observations is that when congenitally blind people who have regained sight are confronted with a shape like a triangle or a square they cannot distinguish between them. The only thing they are capable of doing is to distinguish between figure and ground. Very laboriously, over a period of months, they can finally tell the difference among various geometrical figures, as long as conditions of illumination are constant. To make Hebb's argument more palpable we now turn to a more detailed description of the process of perceptual learning. We must show how this structure comes about and what factors influence its genesis. Before doing this, however, it is important that we become clear as to Hebb's terminology.

C2. Some Terminological Preliminaries

One of the appealing qualities of Hebb's theory of learning is the way time appears in his theory. Experience and hence learning, unfolds in time. It is sequential. Cell assemblies, once established, get connected in sequences. Such organizations Hebb termed phase sequences. They underlie his theory of perceptual learning, and give the actual mechanism of our stream of thought. As Hebb writes (1959, p. 629):

"The phase sequence is a temporarily integrated series of assembly activities; it amounts to one current in the stream of thought. Each assembly activity in the series might be aroused (1) sensorily; (2) by excitation from other assemblies; or (3) in both ways. It is assumed that the last, (3), is what usually happens in an organized flow of behaviour. Each assembly must establish connections with a number of other assemblies, at different times; which of the others it will arouse on any specific occasion will depend on what other activity, and especially what sensory activity, is going on at that moment. At each point in time, behaviour would be steered both sensorily and centrally, jointly controlled by the present sensory input and the immediately prior central activity".

A centrally excited phase sequence was later called (Hebb, 1972) a mediating process. From what we said above perception is a type of phase sequence.

C3. Perception of a Complex and the Phase Sequence

(i) Preliminaries: Lines and Angles in Perception

Before embarking on this discussion the reader should bear in mind that as of 1968 (Hebb, 1968) angles are not considered perceptual elements. Points of fixation are considered perceptual elements. So whenever the reader comes to the words "angle A" in Hebb's older

writings he should remember this caution.

According to Hebb more complex figures such as triangles and squares are gradually assembled from what he calls perceptual elements. By this term he means processes that are partly innate and partly learned. Lines and angles are considered as the elements from which more complex perceptions develop (Hebb, 1949, p. 48). Perceptual complexes such as triangles and squares are learned after the idea of a line becomes clear in perception. This means that an isolated cell assembly for a straight line is what is set up first when somebody is learning to perceive a rectangle.

(ii) Movements during Perceptual Integration

Let us now see how we can put the above to work in order that the perception of a triangle, say, can come about. (See figure 22).

Suppose I fixate at A (angle) with the purpose of looking for the first time at the triangle ABC and suppose I consider A, B, and C as being the perceptual elements.

If I fixate at A, then B, C become peripheral stimulus points that have equal probability of being fixated next. In fact the eye moves either to B or to C. According to Hebb, (1949, p. 82):

"When one point of a straight line is fixated, every point to one side of the fixation tends to arouse exactly the same direction of eye movement, and every point in the line on the side exactly the opposite direction. At times the two vectors may balance, but often they will not. It follows that there is a strong tendency for the eye to make a sweep along the line, in one direction or the other". (Hebb 1949, p. 82)

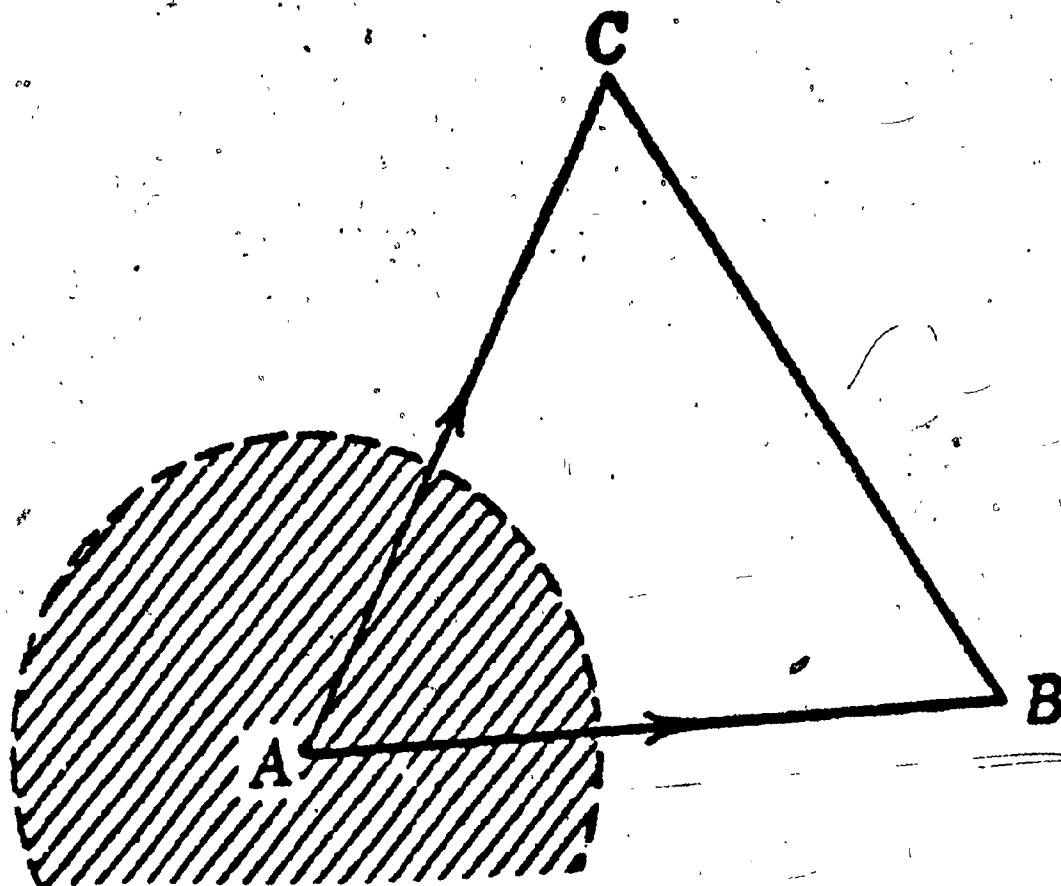


Figure 22. The triangle ABC is seen with fixation on the point A. The macular field is represented by the hatched circle, so that points B and C fall in peripheral vision. Arrows represent the direction and strength of eye movement tendencies aroused by stimulation from B and C. (From Hebb, 1949)

In other words, fixation on A would have two effects. One would be primarily motor with consequent motion toward either B or C. The other would be central. In other words, a set of neurons would be excited in the occipital lobe and beyond. (area 19 and beyond of the visual cortical area). If we call this set of centrally excited cells a then a disproportionate number of these would be triggered from the area around A although the other parts of the triangle would also contribute. Call b and c the areas of the visual cortical area corresponding to B, C as point of fixation. The aforementioned motor components have the effect of joining, sequentially, fixation points A, B, C. Repeated fixation on A, B, C would result in the creation of three diffuse assemblies each of which would be capable reverberatory activity. While a particular assembly reverberates there would also be excitation of a motor assembly tending to move the eye either toward one or the other fixation point. In other words, fixation A assembly would facilitate a motor assembly which in turn would facilitate a motor assembly which in turn would facilitate B assembly for example, and so on. In a real life situation facing the stimulus triangle ABC occurs in variable conditions of illumination, distance, orientation, etc. Under these conditions:

"When the growth (of cell assemblies for A, B, C) occurs in a variable background of visual stimulation, those cells that are affected by the variability would not be consistently active at each repetition of the specific stimulus which excites the assembly and therefore would not be part of it. In other words, as growth proceeds in a, b, and c, fractionation would eliminate elements that are active only when a certain fixation has preceded, and those which are inactive following any particular fixation. Activity in a can

therefore occur independently of b or c, once the integration of these assemblies is complete". (Hebb 1949, p. 89)

(iii) The Development of Superordinate Perceptions

Once assemblies a, b, c have stabilized one can turn to the question of integrating the parts of a figure into a whole. Taking our triangle again, and as Hebb continues (1949, p. 95) one can look at fixation points A, B, C in any order. If the duration of this process is small enough then reverberation in any two of the assemblies can be going on while assembly c say is sensorily excited. In other words one would frequently have activity which is simultaneous in a, b, c. Since the tissue from which a, b, c are made up is organized in such a fashion that tissue of the a assembly is contiguous to tissue of the b assembly and so on one would have a gradual integration of these three systems. What, in other words, we end up with is a trio of assemblies each one of which facilitates the other two with the consequent strengthening of synaptic connections. The resulting complex which Hebb calls a superordinate system is the assembly corresponding to the triangle. This is essentially a new system with a structure of its own.

(iv) The Phase sequence in Perception

We recall that during the development of assemblies a, b, c arousal of any of these was accompanied by motor activities only one of which led to change in fixation to one of the other two fixation points. Hebb (1949, p. 98) schematizes this possible sequence by

a-b-c-b-a-c.....

To this 'ideational' sequence together with its motor components Hebb gave the name phase sequence. However, and this is important, when t becomes organized, the psychological evidence suggests that its activity (t 's activity) intervenes between the activities of the subordinate assemblies a , b , and a and does not supersede them. The sequence, in other words, becomes something like this:

a-b-t-a-c-t-c-t-b-

As evidence for the above, Hebb (1949, p. 99) suggests cases of congenitally blind people who after an operation at first see figure as an amorphous mass, but may be able with effort to count its corners; their perception alternates, consciously, between perception of a whole and of its parts.

Recall here the phase sequence of triangle integration. Let us put it in a slightly different notation.

$a \rightarrow m_{ab} \rightarrow b \rightarrow m_{bc} \rightarrow c \rightarrow t$

where m_{ij} is the motor assembly aroused in going from the i th assembly to the j th. Symbolizing the phase sequence thusly is a more accurate description of what actually happens.

D. Divergent vs Parallel Conduction: Attention

Very young children have notoriously short attention spans. One may ask why this is so and what the role of assemblies is in producing longer attention spans. Hebb (1972, p. 69-70) reasons as follows:

(i) In cases where the sensors (or their continuations) are bunched together so as to form parallel conduction bundles ending at the narrow areas of the somesthetic cortex one has a high degree of reliability. (Such is the case with the skin sensors).

(ii) The excitation when carried from the sensory cortex to the association cortex follows a divergent mode. Neurons lead to different directions and subsequently transmission must be less reliable. Hebb considers this arrangement a boon rather than a curse. To him what happens is that a given excitation can "proceed" through the association cortex provided that along the path of the excited axons assemblies are encountered which are excited due to the task at hand of the CNS. As Hebb simply puts it (1972, p. 73): "The region of divergent conduction will act like a screen, allowing through only the "messages" that can obtain support from activities that are already going on".

The change from parallel to divergent mode of neural transmission is shown in figure 23 due to Hebb (1972). Here S represents the external stimulus as it impinges on the sensory transducers. In parallel conduction branching at each synaptic stage - as is shown in the insert - allows for both time and spatial summation from several neurons. Where divergent conduction occurs this is not in general the case. Thus S which fires group A will certainly fire group B and the single cells C, D, E. At the level, however, of F, G and H transmission would become unreliable unless facilitation occurred due to active assemblies there. Schematically the effects of signals reaching the association cortex are also shown in general in figure 24. Here X, Y and Z are three signals from the sensory cortex which could be coming from three different modalities. When they arrive at the association cortex assembly A is already active. Hence facilitation from A and X can make B fire whereas an active B with the aid of Y would trigger C. B and C now would fire D which by feeding back on B helps to keep it active. A temporal series of activities

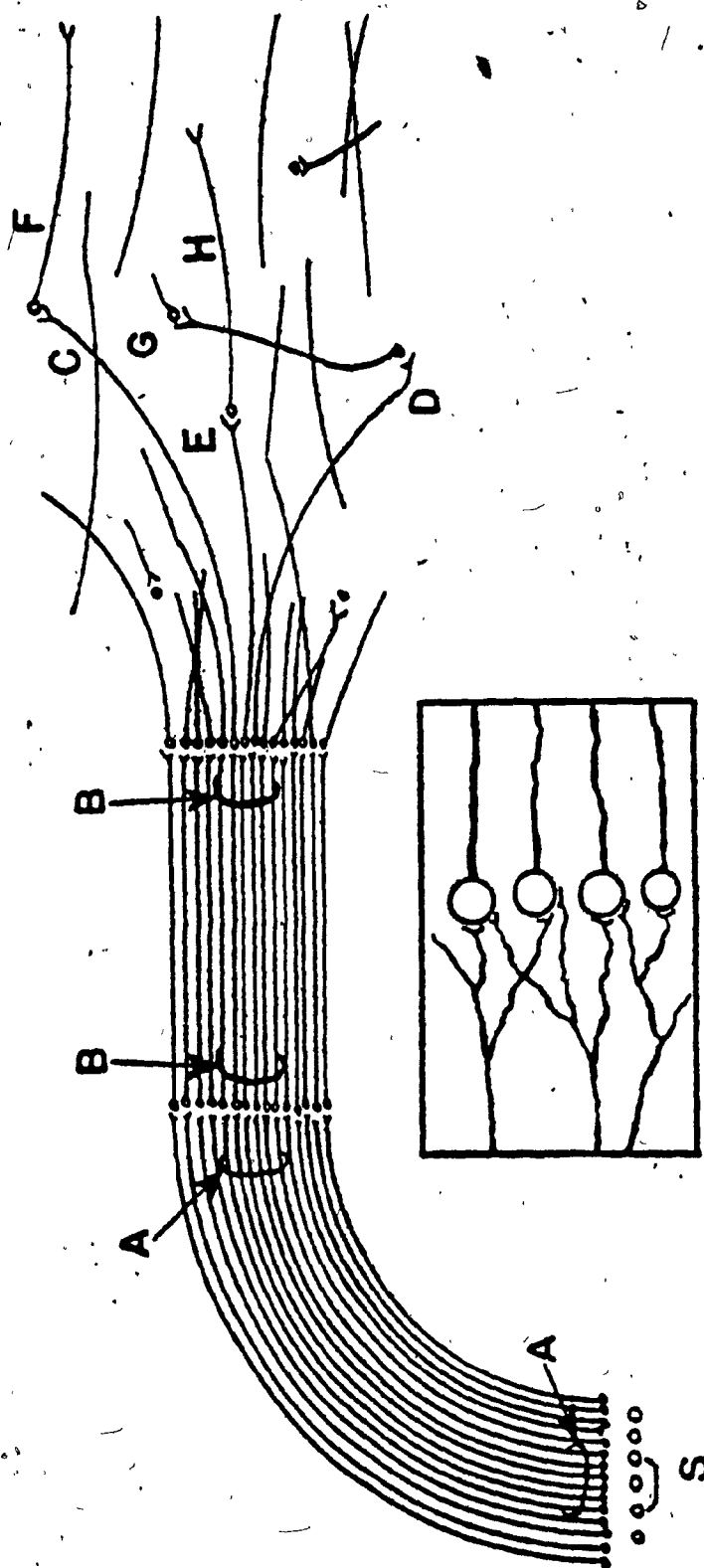


Figure 23. Schematic of neural conduction: parallel on the left, divergent on the right. S, stimulation on part of a receptor surface, excites a group of neurons A which converge at the next synaptic level and provide summation in a postsynaptic group B, which therefore fires reliably. At the next synapse there is divergence; B produces summation in C, D and E and so fires them reliably. There is no summation for F, G and H. Inset: the convergence in greater detail, showing overlap of branching fibers resulting in summation. (From Hebb, 1972)

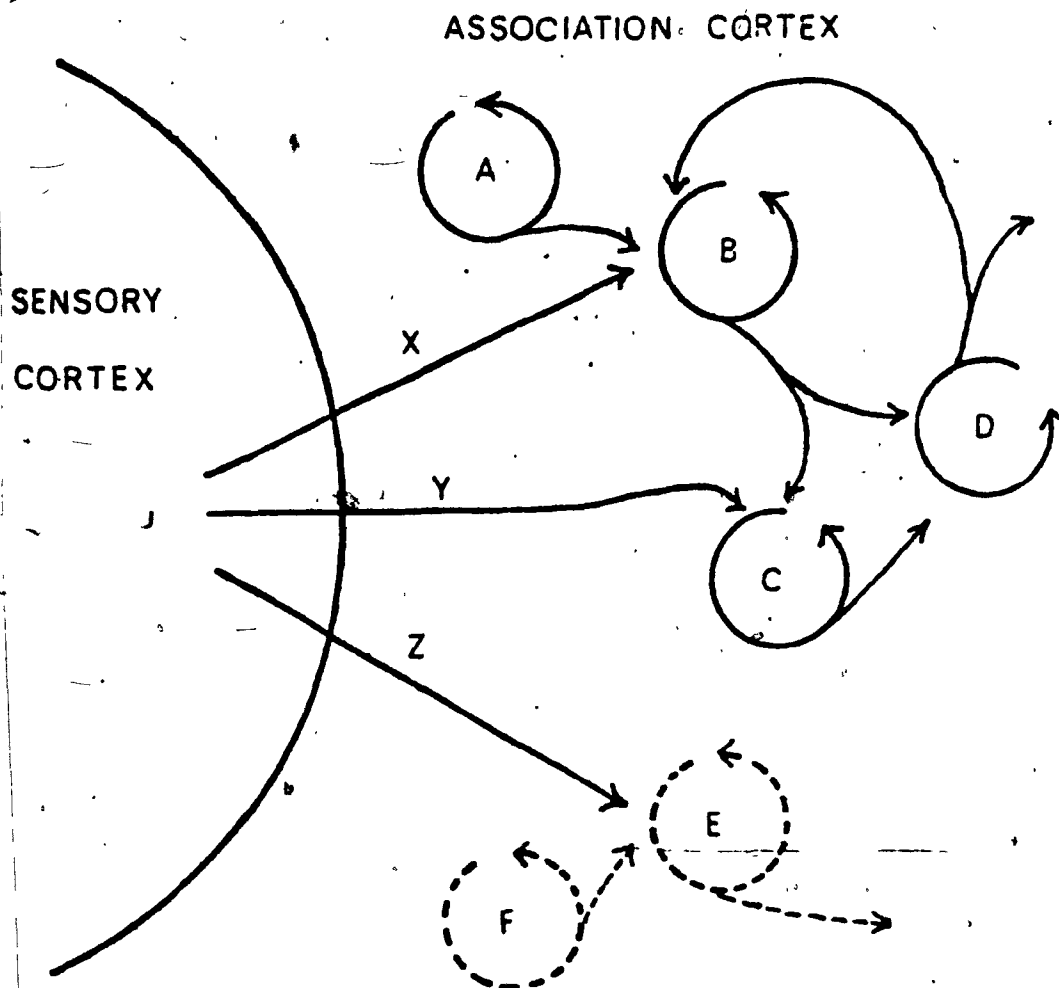


Figure 24. The structuro-dynamic aspect of the attention mechanism according to Hebb. Here it is seen as it filters information from the sensory cortex. X, Y and Z are three different sensory messages. A, B, C and D are reverberating cell assemblies. A is active before whereas C, B and D become active after sensory stimulation. E and F are inactive assemblies. First stimulus X and afterwards message Y would be attended. Z would have no effect. (From Hebb, 1972)

A-B-C-D ensues centrally. Since F is not active, input from Z is filtered out. If we suppose that XYZ are the attributes of a sensory event and if F were active whereas on the other hand A were not, then X Y would have no effect and Z would trigger E. This means that for a given sensory event different central processes would ensue depending on the current status of activity in the central areas. This is the theoretical basis of the distinction that Hebb draws between sensation and perception.

This selectivity in response according to Hebb is what constitutes attention. Attention in other words, is both a structural and a dynamic quality of a duly constituted system. It is structural because unless the assemblies exist no conduction can happen and it is dynamic because it also depends on the current activity of the aggregate of assemblies.

E. Achievements and Limitations of the Hebbian Schema

E1. Recapitulation

Let us now see in retrospect what our schema can do. We saw that:

(1) Perception depends on learning first to see the parts of an object clearly, a process involving a series of visual fixations, and presumably from seeing, at first, an amorphous mass containing several foci (the corners), to seeing a distinctive figure at a glance. Even at this final stage we know that perception of the whole is dependent on eye movements for maximal clarity. According to the schema, perception is constituted by a temporal sequence of activity in supra-sensory (or association-area) structures, which owe their organization to changes at the synapse.

(2) Set-attention-expectancy are basically the central facilitation of a perceptual activity e.g. activity in b may be aroused in two ways: sensorily, by looking at B , or centrally, by the "association of ideas." When B is looked at just after A , activity in b is aroused both ways, and the central facilitation from a , is an instance of attention; a central reinforcement of a particular sensory event.

(3) Generalization: Given perception of the various parts or properties of a pattern on the one hand separately: and the possibility of an association of these perceptions with a perception of the whole: we have given at once the possibility of a single response to two patterns in their total properties but which have a property or properties in common.

Specifically, suppose t is the perception of a given triangle which goes according to

$$a+b+c \rightarrow t_1 - a+b \rightarrow t_1$$

and where t_1 corresponds to an assembly with complex cells. And further suppose that we have any other triangle with fixation points a' , b' , c' leading to t_j . The Hebbian schema now that leads to the conception of triangularity in general is as follows:

$$t \rightarrow t_j \rightarrow t_j \rightarrow \dots t$$

where t is a superordinate assembly composed of hypercomplex cells.

The reason, briefly, why hypercomplex cells are required for the general notion of triangularity was forced on Hebb by Hubel and Wiesel's findings. It may be recalled that in Hubel and Wiesel's

schema simple cells are those in which a number of retinal cells converge, complex those on which simple cells converge, and hyper-complex those on which the complex cells converge. Each of these levels increases the process of abstraction by one degree.

Here is a beautiful example due to Hebb (1968) of how this mechanism might be used. "Let us say an infant has already developed assemblies for lines of different slope in his visual field. He is now exposed visually to a triangular object fastened to the side of his crib, so he sees it over and over again from a particular angle. Looking at it excites three primary assemblies, corresponding to the three sides. As these are excited together, a secondary assembly gradually develops, whose activity is perception of the object as a whole - but in that orientation only. If now he has a triangular block to play with, and sees it again and again from various angles he will develop several secondary assemblies, for the perception of the triangle in its different orientations. Finally, taking this to its logical conclusion, when these various secondary assemblies are active together or in close sequence, a tertiary assembly is developed, whose activity is perception of the triangle as a triangle, regardless of its orientation".

E2. Some Recent Evidence Corroborating Organization in Perception

We shall here review two pieces of research that strongly indicate that as Hebb claims perception is a complex of processes with an internal structure. Hebb (1972, p. 239) mentions others including the work of D.E. Broadbent which demonstrated holding in auditory perception.

(1) The Work of Noton and Stark

Let us first consider the work of Noton and Stark (1971): In this experiment subjects learned and recognized patterns which were marginally visible, requiring them to fixate directly each feature to which they wished to attend. Fixed "Scanpaths", specific to subject and pattern, appeared in their saccadic eye movements, both intermittently during learning and in initial eye movements during recognition. The assumption was, that "if patterns are presented to a subject under conditions of poor visibility, so that he is forced to look directly (foveally) at each feature to which he wishes to attend, then the position of his eyes will reveal the features processed and his saccadic eye movements from feature to feature will reveal the order of processing". The experiment had two phases, one a "learning" phase and the other a "recognition" phase. These, however, were external labelings of the experimental process. Subjects were not told in which phase they were participating.

"In the learning phase the subject viewed five patterns, which he had not previously seen, for 20 seconds each. In the recognition phase, which followed immediately, these patterns were intermixed with five other patterns, to make the recognition process less easy, and the set of ten patterns were presented to the subject three times, in random order which changes in time, each view of each pattern lasting 5 seconds".

Under these combined conditions of poor visibility the subject, who was fixating on one extremity of a pattern, could not see any significant details of the opposite extremity. His eye movements during learning and recognition were measured.

"Analysis of the records obtained showed, first, that when a subject was freely viewing a pattern during the learning phase, his eye usually scanned over it following intermittently but repeatedly, a fixed path characteristic of that subject viewing that pattern (quite different fixed paths being followed by the subject with different patterns and by different patterns and by different subject with the same pattern); and second, that when the same pattern was presented to the subject during the recognition phase, his first few eye movements usually followed the same path he had established for that pattern during the learning phase. For example, figure 25A shows a typical pattern viewed by a subject in these experiments; figure 25B is an idealized drawing of the fixed path observed repeatedly in his eye movements.... Such a fixed path, characteristic of a given subject viewing a given pattern, we have termed his "scanpath" for that pattern..

During the learning phase, 25 percent of the total viewing time, on the average, was occupied by occurrences of the scanpath, although this figure was as high as 40 percent with one subject. During the recognition phase, the appropriate scanpath appeared in the initial eye movements in 65 percent of the cases, on the average - that is, in two of three recognition-phase viewings of each pattern, on the average..... On the average, scanpaths in the learning phase contained 10.1 fixations and lasted 4.2 seconds, whereas in the recognition phase the corresponding averages were 7.3 fixations and 2.5 seconds." And here are the implications of their work according to them:

"What implications do these results have for theories of pattern recognition and perception? Although the particular scanpaths used by a subject for different patterns are potentially informative concerning his particular mental processes, we believe that the mere existence of scanpaths is a

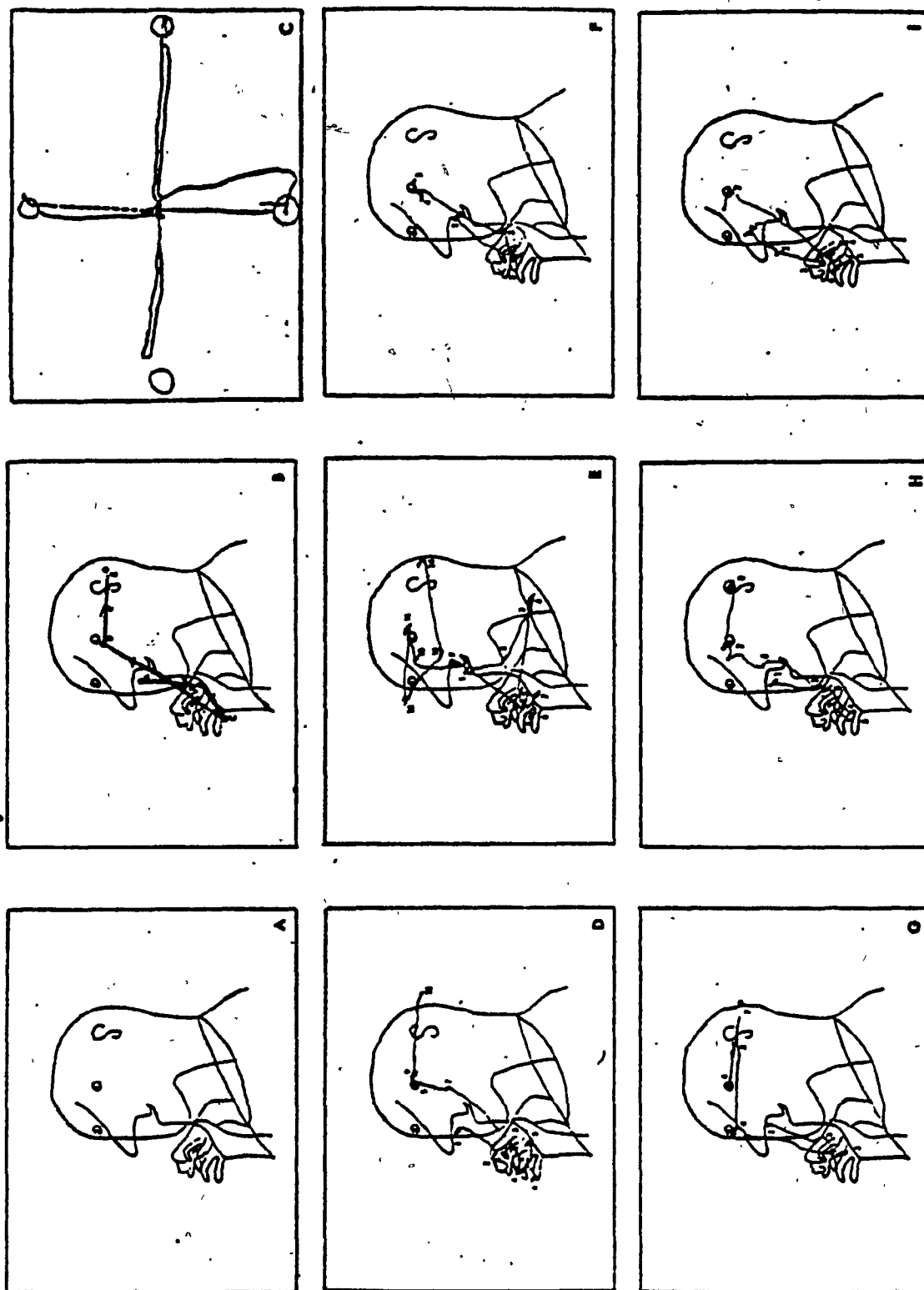
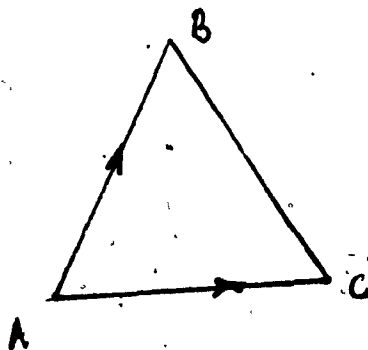


Figure 25. Examples of scanpaths in eye movements during learning and recognition phases. (From Noton and Stark, 1971)

more immediately significant fact, suggestive as it is of how patterns are remembered and recognized by all subjects. That the subject follows a fixed path from feature to feature suggest that the eye movement motor components involved in perception are not merely involved in moving the pattern over the retina, but are an integral part of the memories on which recognition is based. It suggests that the subject's internal representation or memory of the pattern is an alternating sequence of sensory and motor memory traces, recording alternately a feature of the pattern and the eye movement required to reach the next feature. During occurrences of the scanpath in the pattern-learning phase, these memory traces are being laid down, and, when the scanpath appears in the initial eye movements during recognition, the subject is matching the internal representation with the pattern, by reproducing the successive feature memories, a successful completion of the scanpath indicating recognition of the pattern".

Upon reading the above one is struck with the similarity between the phase sequence of Hebb with its concatenations of fixations - motor movement - fixation -abstraction and the serial visual processing as reported by Noton and Stark. This writer would like, however, to make the following remarks:

First, it will be recalled that in our treatment of the learning process of perceiving a given triangle if one fixated on say, A (see diagram) then one would have an equal chance of motor excitation toward B as it would have toward C. According to Hebb, this would be true only on the first occasion that one had seen a triangle.



After repeated practice, a preferred strategy would develop. Each one of us apparently responds to either B or C and keeps this choice of stimulus permanently when one

is confronted with the same figure.

Secondly, the recognition process seems to be an abstraction process because if a less number of fixations (i.e. 7 as against 10) is required to integrate the figure it means that either the super-ordinate assembly can be triggered with less first order assemblies or that the assemblies that are triggered directly during the recognition phase trigger (by association) the non-triggered ones leading to the second order assembly.

Similar results to those of Noton and Stark have been obtained by Yarbus (1967) and Jeannerod et al (1968).

(ii) The Stabilized Image Experiments

In normal vision the image of an object on the retina is constantly in motion. Three motions are superimposed. These three motions of the eyeball that keep the image moving play an important role in the sensory function of the eye. It is possible, however, to keep the image of the object fixated on the retina. This has been achieved by suitable mounting on a contact lens a contraption that moves in phase with the eyeball thus keeping the image on the same region of the fovea. It is found that after a short time these images begin to fade and subsequently disappear.

Work on stabilized images was done, originally by D. W. Ditchburn in 1952, and independently, by L. A. Riggs. Here we shall comment on the work of Pritchard, Heron and Hebb as reported by Pritchard (1971). The basic facts are as follows:

- (1) A simple visual object such as a straight line disappears in perhaps 10 to 20 seconds.

- (2) Secondly, an object such as a square or a triangle does not act as whole but breaks up or regenerates in parts in the manner shown in the figure 26.
- (3) A single line may be visible for only 10 percent of the aggregate viewing time; a more complex figure may be visible in whole or in part for more than 80 percent of the time, as Pritchard presents it (1971, p. 119):

"The contrasting manner in which complex images fade and regenerate lends support to the role of learning in perception. For example, the figure of the human profile invariably fades and regenerates in meaningful units. The front of the face, the top of the head, the eye and ear come and go as meaningful units.

In the cell assembly approach to a theory of perception these observations are explained in terms of "perceptual elements" as opposed to purely sensory elements. The "organized", meaningful or recognizable parts of the image correspond to perceptual elements previously learned or established by experience. The parts of the human profile would thus function as perceptual elements at the outset of behaviour of the stabilized image.

When entire words are presented, the partial fragmentation of letters can cause different words to be perceived e.g.

BEER - PEER - PEEP - BEE - BE

In experiments with simple straight line figures the cell assembly approach is supported by the observation that the line is the apparent unit of perception just as the line is the unit of structure in the figure."

The cell assembly explanation of the above results is easy:

The assembly functions as a system. In fact, complex figures correspond to superordinate assemblies composed of first order assemblies and sub-assemblies. Each first order assembly owing to:

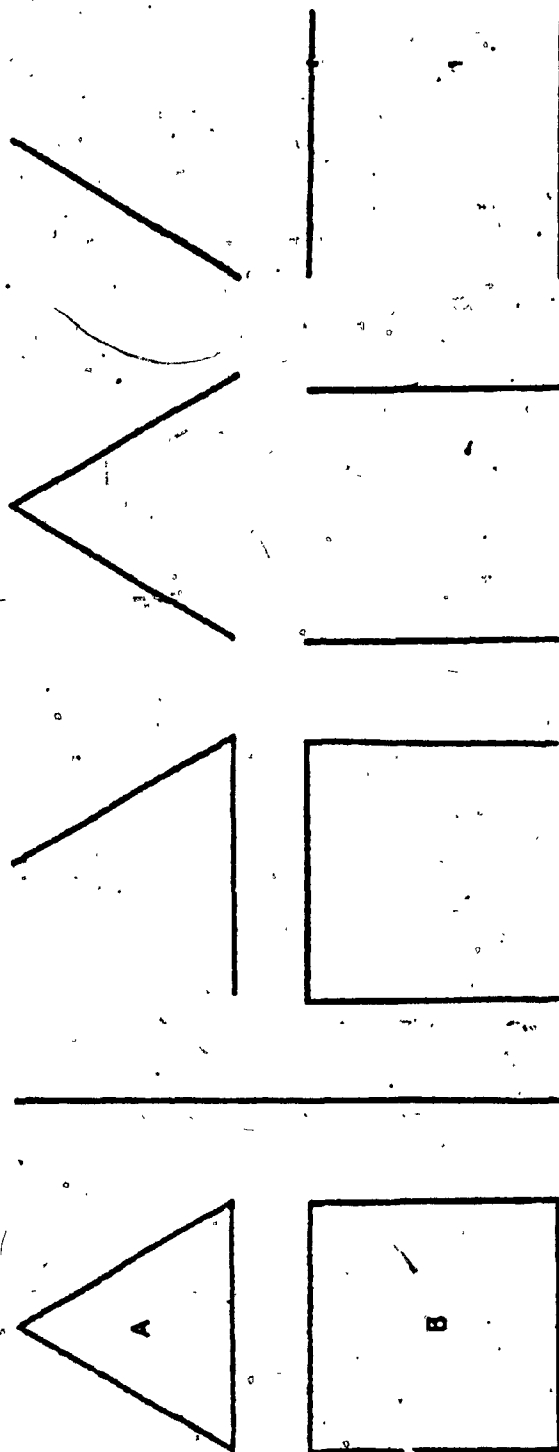


Figure 26. Disintegration of images in the stabilized image experiments.
(From Hebb, 1972)

the motion of the eye gets input from a strip of retinal cones. If the image is frozen on the cones a time will arrive that fatigue will set in as a result of which the particular groups of neurons making up the cell assembly will cease to function. So "chunks" of the target object disappear as far as the viewer is concerned. In regular vision because of tremor, for example, a given line gets shifted to another set of cones and hence no apparent disintegration occurs. (Hebb, 1963, p. 16-27).

F. Conclusion

Hebb's theory seems to be a very powerful construct. Hebb himself with others (1971) has given some indication as to how his general ideas could be applied in linguistics. A much more intriguing problem is the application of Hebb's ideas to the problem of self which according to Hebb (1960) is something arising partly out of a set of mediating-processes that come about by experience.

Hebb's basic appeal to this writer is the historicity property of cell assemblies. This I believe is its strongest point and link with real physical dynamics. Our next chapters follow another dimension of analysis. We shall attempt to construct an abstract mathematical structure (complemented by its computer analogue) which shall use facilitation as the mechanism of adhesion of formal elements known as neurons. We shall show cell assemblies can indeed be formed via this mechanism! Subsequently, an attempt will be made to show that this abstract mathematical model can be correlated to the theoretical conclusions reached in our chapter on memory. The cognitive aspects of our model will thus be revealed.

CHAPTER VI

THE PROBABILISTIC NEURAL NET: THE ANALYTICAL APPROACHA. Qualitative Aspects

The system to be described below, being a model, is an abstraction. Some of its component parts have names recalling biological entities which they are supposed to resemble in certain aspects.

A1. The Neuron

The brain is viewed as a set of identical interconnected elements known as neurons. Neurons can either be on or off, i.e. they can be active or inactive. The brain, in other words, is viewed as a vast neural net. The neuron, has the characteristics of a relay. The afferent area of the neuron consists of the soma and the dendrites. Information is supposed to enter the neuron via the afferent area. The efferent area, consisting of the axon and its collaterals, is the channel the neuron uses to send its information to other neurons. The neuron codes information in the form of electrical pulses. If the incoming excitation which is of the form of an electrical graded potential exceeds some critical value known as the threshold the neuron 'fires' and sends out an electrical signal of standard height known as the action potential. This action potential propagates along the axon and its collaterals unattenuated until the axon makes contact with the afferent area of the next neuron. When the neuron fires it becomes insensitive to any further excitation for a period of time

known as the absolute refractory period (ARP). Further, the firing process unfolds in time τ known as the synaptic delay which is the time between the arrival of a signal at neuron j and the firing of neuron i . The contact area is known as the synapse. At the synapse the standard size action potential (AP) will be modified in some manner and emerge in the afferent area of the postsynaptic neuron. This potential is known as the (postsynaptic potential), (PSP). All PSP's arriving at a neuron are summed instantly and, if they exceed the specified threshold, will cause the neuron to fire without delay. In this paper the ARP is to be taken as having a value greater than τ but smaller than 2τ . Temporal summation occurs, without decrement, for a period less than the synaptic delay. These assumptions have the effect that, if activity is initiated by a set of neurons at time t , then all subsequent activity will occur at $t + \tau$, $t + 2\tau$, etc.

In this model the action of the synapse is parameterized by a coupling coefficient k_{ij} which relates the AP originating on the j th neuron to the PSP generated on the i th neuron. Assuming that we can write the AP and the PSP as products of amplitudes and time functions we have (Csermely, 1968).

$$u_j(\text{AP}) = U_j(\text{AP}) f_1(t)$$

$$u_i(\text{PSP}) = U_i(\text{PSP}) f_2(t)$$

$$\text{by definition } k_{ij} = \frac{U_i(\text{PSP})}{U_j(\text{AP})}$$

Taking $U_j(\text{AP})$ as unity (since it is of standard size) we can formally write:

$$k_{ij} = \sum_i (\text{PSP})$$

Synapses are of two types. They either are excitatory in which case k_{ij} is positive, or they are inhibitory, and in this case k_{ij} is negative. The threshold θ of all neurons will be assumed to be identical. If the combined input $\sum_j k_{ij} a_j$ for the i th neuron exceeds θ it will go into the firing state $a_i = 1$.

A2. The Net

The topology of synaptic connections in the brain is not yet completely known. In some parts of the brain, as Sholl (1956), Colon (1968) and Smit (1968) have shown, synaptic connections with near neighbours seem to be preferred. However, following a suggestion by Szentagothai (1971), we shall assume in this work, that every neuron can be connected with any other neuron in the brain. This seems to require between five and ten synaptic stations between any two neurons in the brain.

A neural network in which the neurons have been numbered in an arbitrary fashion from 1 to A and for which all the coupling coefficients k_{ij} have been specified, can be pictured in various ways. For example, Harth and Edgar (1967) use what they term as the connectivity matrix. This matrix contains all the k_{ij} 's. If no connection exists between the j th and the i th neuron, then $k_{ij} = 0$. Since the net we have is random we assume that the seeding of the coefficients k_{ij} in the connectivity matrix is random.

If our neural net is considered to be made up of sensory, association and motor neurons, our matrix would look like that in figure 27. A cortical state means a complete specification of the firing state of A neurons. This may be represented by a column vector $\vec{a}(n)$

$$\vec{a}_n = \begin{bmatrix} 1 \\ a_n^1 \\ a_n^2 \\ \vdots \\ a_n^A \\ a_n^a \end{bmatrix}$$

where $a_n^i = 0$ if the i th neuron is in a resting state and $a_n^i = 1$ if it is active.

We can similarly define 'sensory vectors' $\vec{s}(n)$ (also known as sensory states) which give us the firing pattern of the sensory cortex in a similar fashion.

$$\vec{s}_n = \begin{bmatrix} 1 \\ s_n^1 \\ s_n^2 \\ \vdots \\ s_n^S \\ s_n^s \end{bmatrix}$$

B. Quantitative Aspects

Harth and Edgar (1967), Anninos et al (1970) have constructed a "statistical mechanics" of these nets as follows:

B1. Parameters of Neural Net Model

τ Synaptic delay.

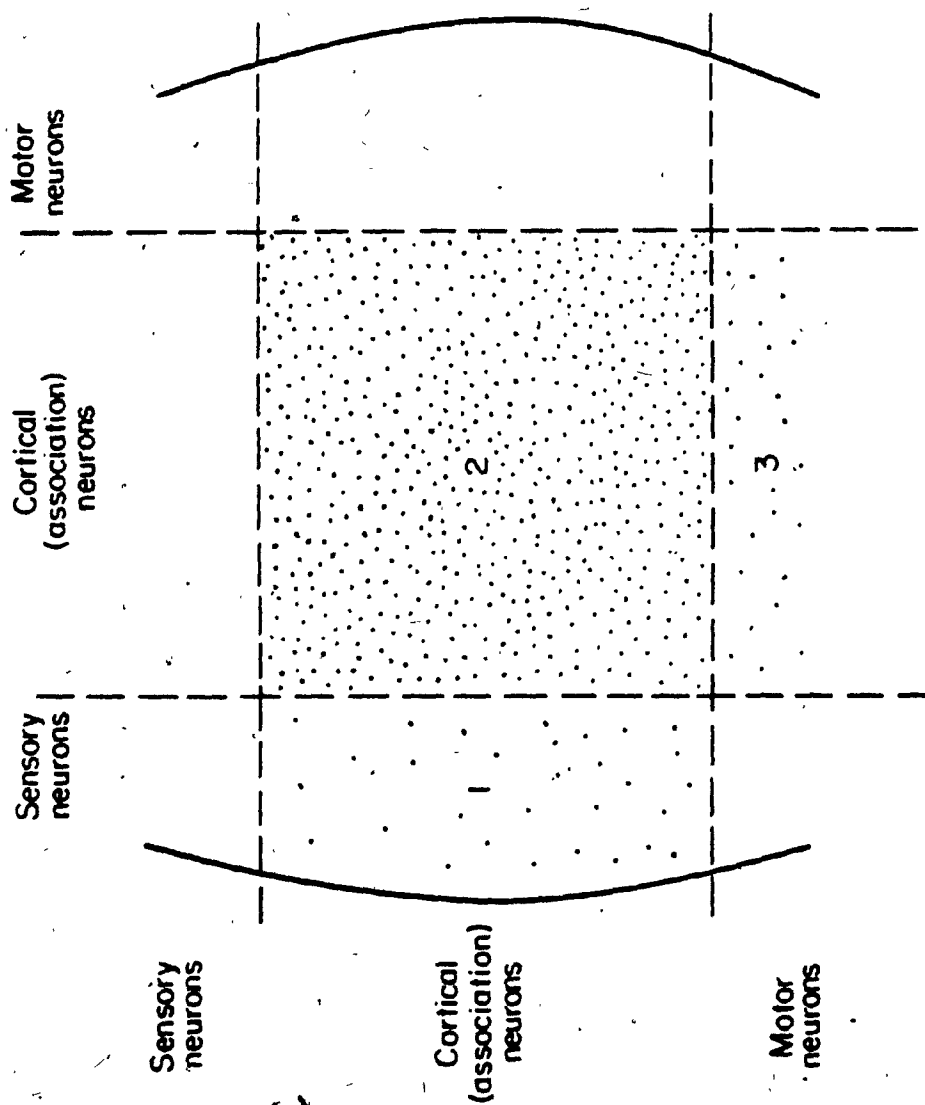


Figure 27. The connectivity matrix as constructed by Harth and Edgar (1967)

- A Total number of neurons in the netlet.
- h Fraction of inhibitory neurons in the netlet.
- $\mu+$ The average number of axon branches emanating from an excitatory neuron.
- $\mu-$ The average EPSP produced by an excitatory neuron in arbitrary units.
- K- The average IPSP produced by an inhibitory neuron in arbitrary units.
- θ Firing threshold of the neurons in the netlet.
- η The minimum number of EPSP's necessary to trigger a neuron in the absence of inhibitory inputs.
- η' The minimum number of EPSP's necessary to trigger a neuron in the presence of inhibitory inputs.
- α_n The activity; i.e., the fraction of active neurons in the netlet at $t = n\tau$.

B2. Expectation Value of the Activity

The expectation value $\langle \alpha_{n+1} \rangle$ of the activity at $t = (n+1)\tau$, i.e., the average value of α_{n+1} generated at the above time is given by:

$$\langle \alpha_{n+1} \rangle = (1 - \alpha_n) P(\alpha_n) \quad (1)$$

where $P(\alpha_n)$ is the probability that a neuron receives PSP's exceeding its threshold at $t = (n+1)\tau$ (Anninos et al, 1970). This quantity is a function of the activity at $t = n\tau$ and is given by

$$P(\alpha_n) = \sum_{m=0}^{M_{\max}} \sum_{l \geq \eta}^{M_{\max}} P_l Q_m \quad (2)$$

where P_l and Q_m are the a priori probabilities that a neuron receives l and m excitatory and inhibitory inputs, respectively i.e.

$$P_l = e^{-\alpha_n(1-h)\mu+} \{\alpha_n(1-h)\mu+\}^l / l! \quad (3)$$

$$Q_m = \exp(-\alpha_n h \mu-) (\alpha_n h \mu-)^m / m!$$

and

$$L_{\max} = A\alpha_n(1-h)\mu^+$$

$$M_{\max} = A\alpha_n h\mu^- \quad (4)$$

Further, $\eta' = u\{(\theta - mk^-)/K\}$. It is defined on the basis of the function $u(x)$ which is defined as the smallest integer which is equal or greater than x .

From equation (4) we see that L_{\max} will always be very much larger than the exponent in equation (3), hence, we can write:

$$\frac{L_{\max}}{\sum_{L=\eta'}^{\infty} P_L} \approx \frac{\sum_{L=\eta'}^{\infty} P_L}{1 - \exp\{-\alpha_n(1-h)\mu^+\}} \frac{\eta'^{-1} \{\alpha_n(1-h)\mu^+\}^{L/L}}{\sum_{L=0}^{\infty} \eta'^{-1} \{\alpha_n(1-h)\mu^+\}^{L/L}} \quad (5)$$

Substitution of (2) and (5) into (1) yields

$$\begin{aligned} \langle \alpha_{n+1} \rangle &\approx (1-\alpha_n) \exp(-\alpha_n h\mu^-) \sum_{m=0}^M (\alpha_n h\mu^-)^m / m! \\ &\times \{1 - \exp(-\alpha_n(1-h)\mu^+) \frac{\eta'^{-1} \{\alpha_n(1-h)\mu^+\}^{L/L}}{\sum_{L=0}^{\infty} \eta'^{-1} \{\alpha_n(1-h)\mu^+\}^{L/L}}\} \quad (6) \end{aligned}$$

In equation (6) we have replaced the upper limit M_{\max} by M , an integer sufficiently large so that the addition of higher terms may be neglected.

Curves of $\langle \alpha_{n+1} \rangle$ vs. α_n are shown in figures 28 to 31 (Anninos et al, 1970) for different combinations of netlet parameters. A convenient parameter is the minimum number of EPSP's necessary to trigger a neuron in the absence of inhibitory inputs. This quantity η is defined by

$$\eta = u(\theta/k^+)$$

C. Classification of Netlets with regard to the presence of sustained activity

(1) A convenient way of looking at the question of sustained

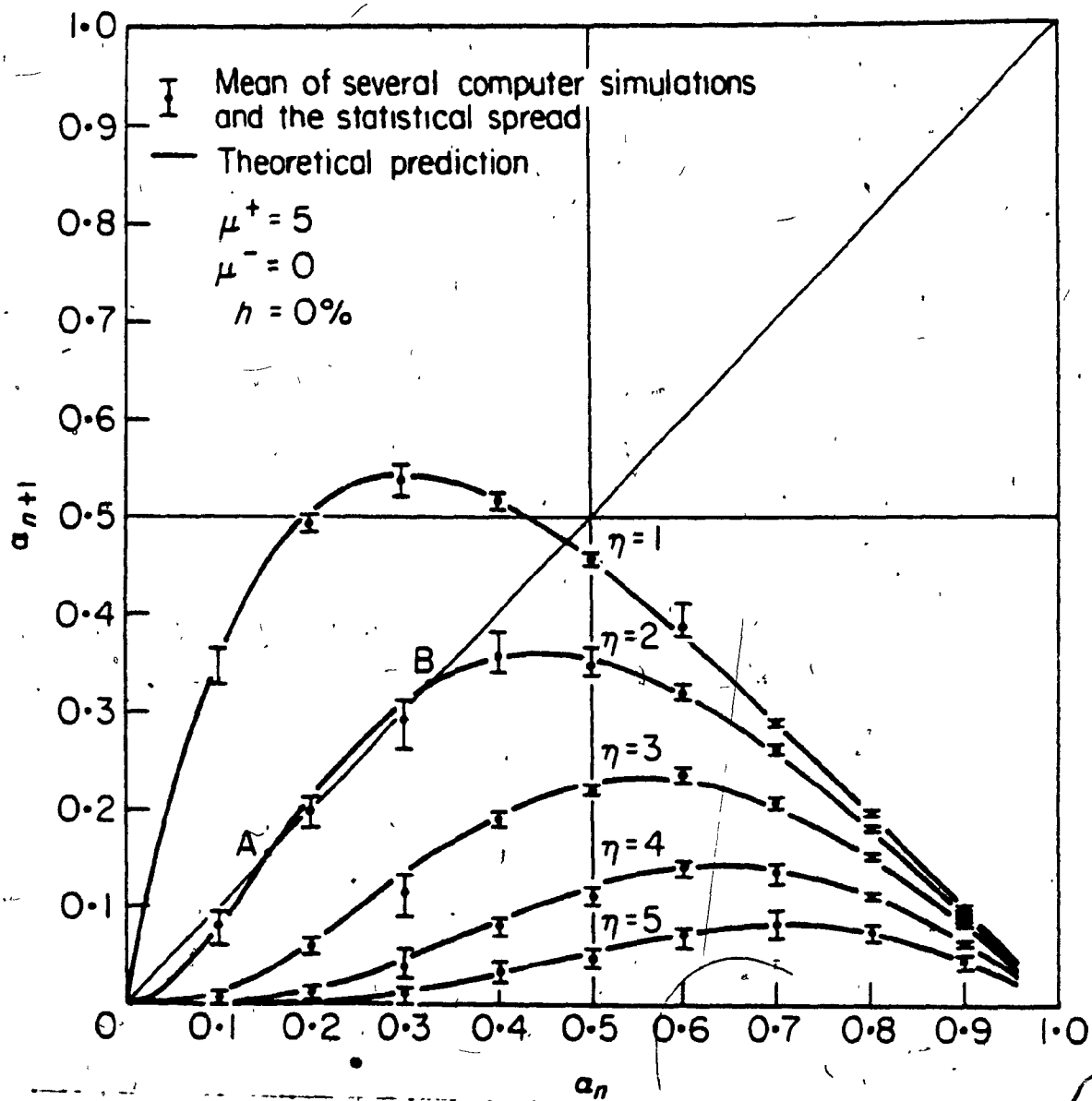


Figure 28. Expectation value of neural activity $\langle \alpha_{n+1} \rangle$ vs. preceding activity α_n . The netlet is characterized by the parameters $h=0$, $\mu^+=5$.

Curves are theoretical values from equation (6). Points are obtained by computer simulation of 1000 neurons.

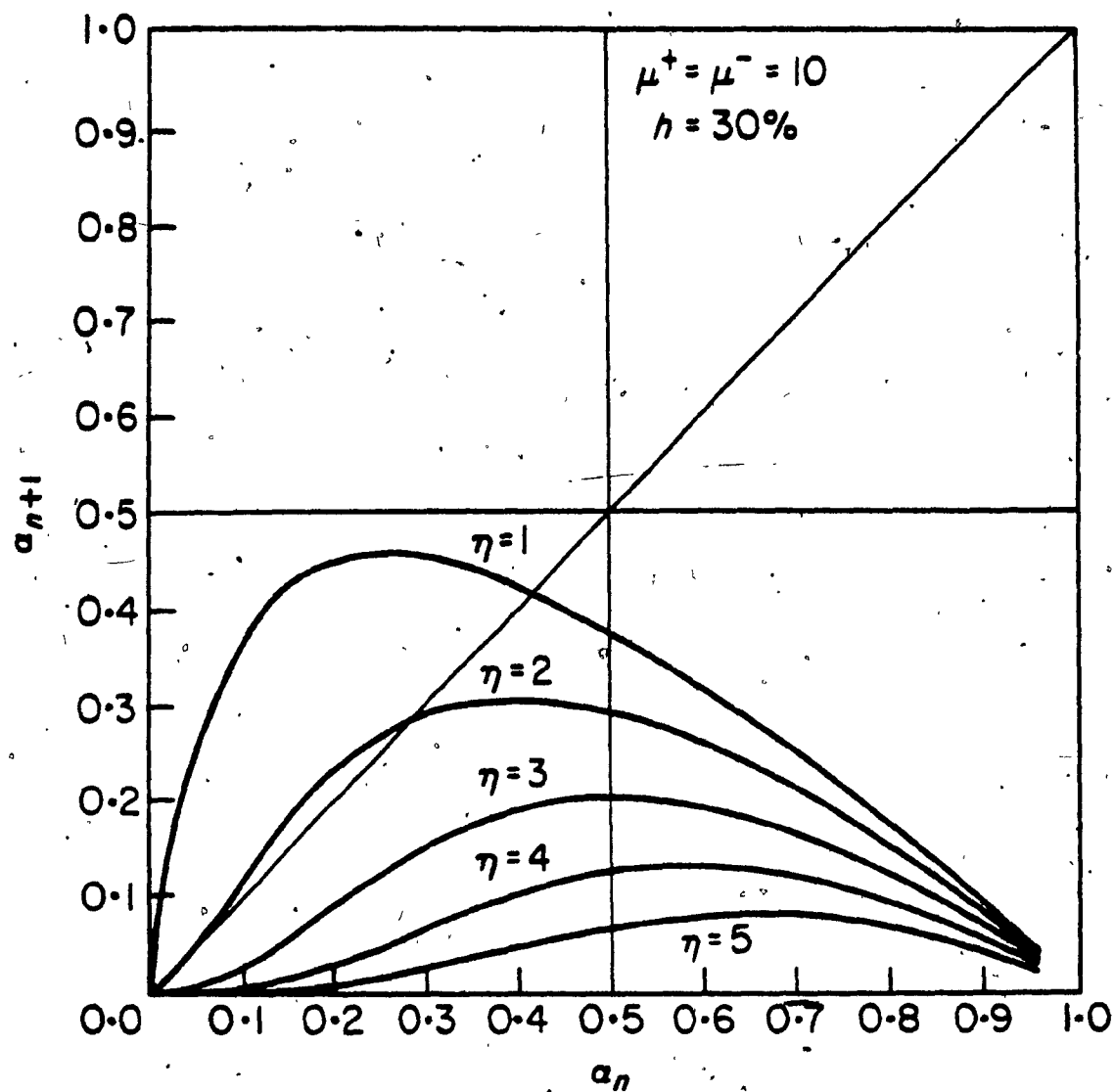


Figure 29. $\langle \alpha_{n+1} \rangle$ vs. α_n for netlets with $h = 0.3$, $\mu^+ = \mu^- = 10$.

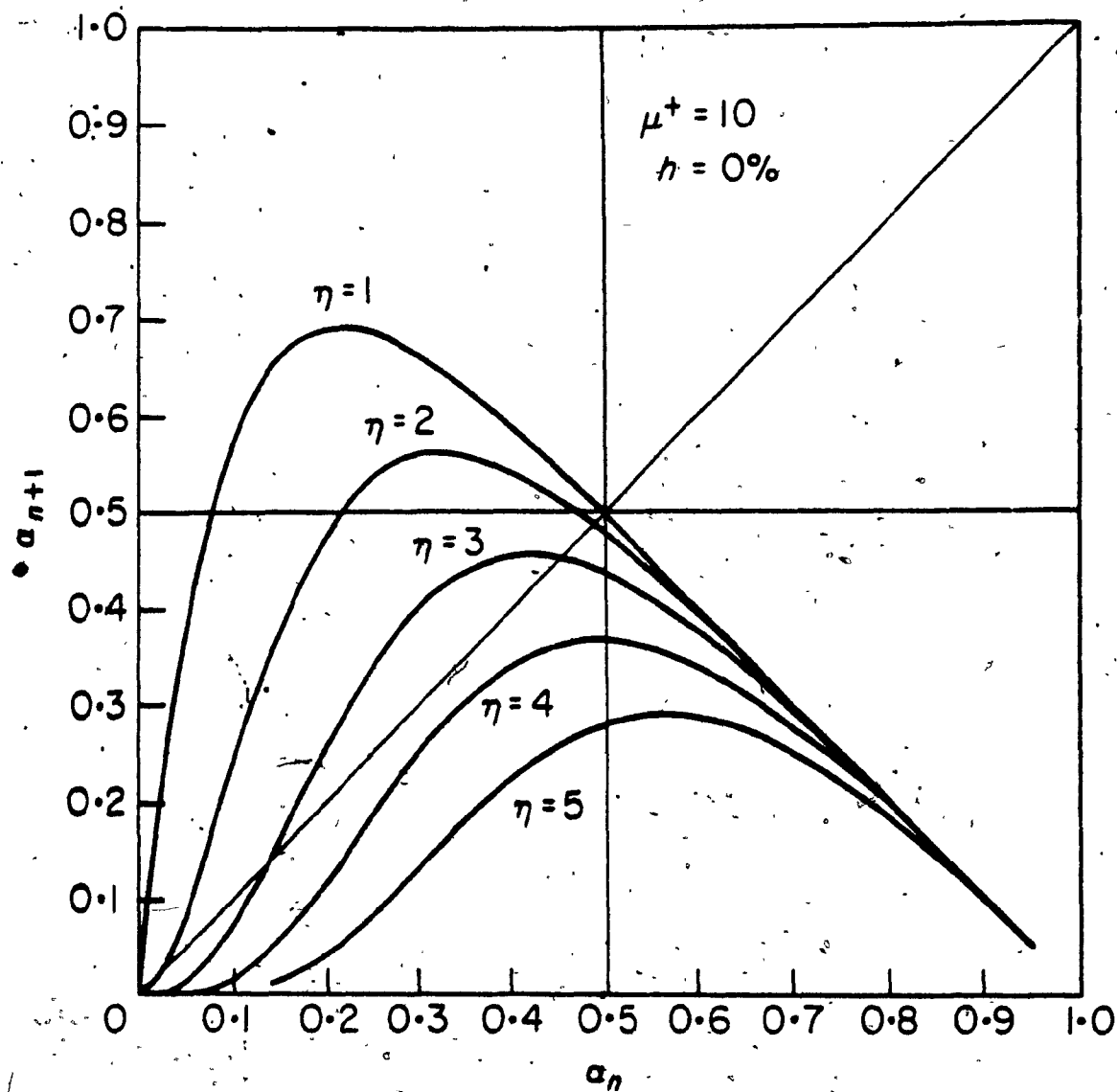


Figure 30. $\langle \alpha_{n+1} \rangle$ vs α_n for netlets with $h=0$, $\mu^+ = 10$.

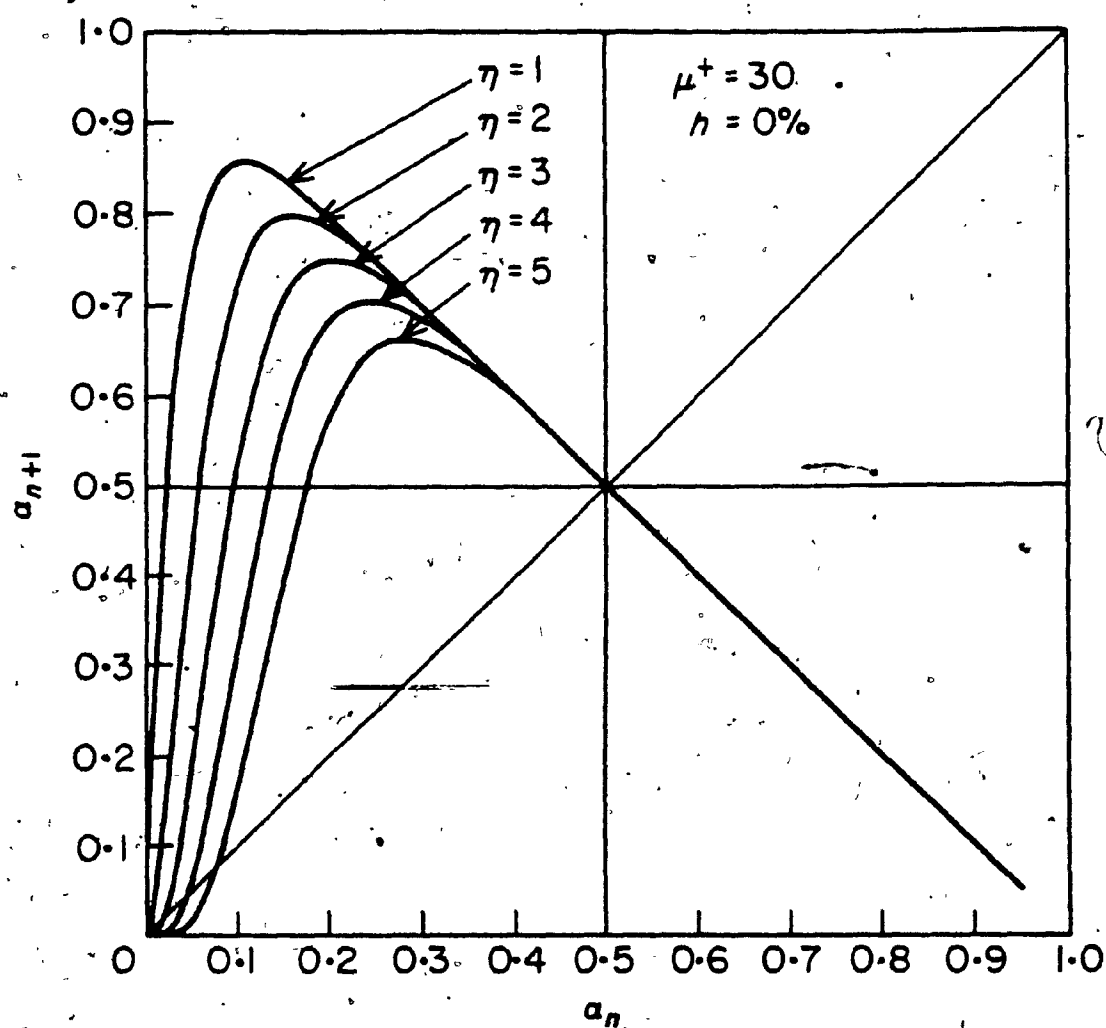


Figure 31. $\langle a_{n+1} \rangle$ vs. a_n for netlets with $h = 0$, $\mu^+ = 30$.

activity is through the plot of the activity α_{n+1} (at $t = (n+1)\tau$) versus the activity α_n of the preceding instant $t = n\tau$. It is evident that if $\alpha_{n+1} > \alpha_n$ the number of active cells increases; when these two values are equal, the activity is stable: there is neither increase nor decrease of the number of active elements. On the other hand, if α_{n+1} is always lower than α_n (i.e. the curve always remains below the 45° diagonal) then for any initial value of activity whatsoever, there are no prospects for sustained activity, since the number of cells active in successive iterations will decrease in monotonous fashion. Thus, study of the curve of α_n versus α_{n+1} may provide important information about the possibility of sustained activity in a given net and, in turn, also allow us to analyse the role of various parameters affecting sustained activity.

These curves have been classified into three groups. These three classes of nets differ in their reaction to an initial input given at time $t=0$. Considering that the nets explored here are noiseless, it is of course necessary to provide some pattern of initial activity at time $t=0$ in order to allow the net to continue its activity. However, the three types of nets differ in their capacity to maintain this initial activity. Because this distinction provides an important measure of the effect of structure on function, these nets are described here in some detail: (Anninos et al, 1970).

Class A nets are those nets for which $\alpha_n=0$ is not a stable point. Consequently, these nets produce sustained activity for all values of initial input to the net (the only exception is when $\alpha_n=1$, i.e. all neurons in the net are activated simultaneously, for in this case all cells will be simultaneously refractory in the next cycle α_{n+1}).

so that no spontaneous activity will follow in subsequent cycles. An example of this mode are the curves for $n=1$ in figures 28 to 31.

Class B netlets lie partly under the 45° diagonal and partly above it; for low values of α_n the ensuing $\alpha_{n+1} < \alpha_n$. However, if α_n exceeds a certain level, as seen for example in the curve $n=2$ figure 28, then $\alpha_{n+1} > \alpha_n$. In other words, these nets show a threshold for sustained activity. If the initial activity is below this threshold, subsequent activity will decay. Unlike Class A netlets which are characterized by having $\alpha_n = 0$ as an unstable point and $\alpha_n \sim 0.5$ as stable one, in Class B netlets both these points are stable. Thus, Class B netlets have the property of highly non-linear decision elements having two stable states and a threshold for excitation from inactive to active state.

Class C netlets lie wholly below the 45° diagonal. Hence, activity will decrease for any initial value of activity. These nets are thus incapable of maintaining any form of sustained activity.

It is clear from the above discussion that the classification of neuronal netlets into three classes allows evaluation of these nets in fairly precise manner with regard to the presence or absence of sustained activity following an initial stimulus. This criterion is of interest not only from a theoretical viewpoint, but also from an experimental point of view as well.

Thus, if one finds empirically that a net is oscillating one can go backwards and determine whether this net belongs to any of the aforementioned types.

(ii) The Universal Curves with Inhibition

The last observation brings us to another way of looking at the activity curves. If in equation (6) one plots not $\langle \alpha_{n+1} \rangle$ versus α_n but one plots γ versus $\alpha_n \mu^+$, one obtains what are known as the universal curves. (See figures 32 to 34). Here $\gamma = \frac{\langle \alpha_{n+1} \rangle}{1 - \alpha_n}$

Observe that each of these curves can handle all multiplicities μ^+ . One must, however, construct a different family of universal curves for different levels of inhibition. The importance of these curves will become apparent when we come to consider overlap behaviour and learning. These curves have been developed by this writer.

(iii) Netlets with Afferent Input

The detailed studies of isolated neural nets are useful as a starting point for study of more complex systems. The simplest case is that of neural nets with steady input. Consider a probabilistic neural netlet of the type discussed previously, except that a cable of afferent fibers is attached to it. Upon entering the netlet, each of the fibers splits and makes synaptic connection with μ_0^+ or μ_0^- different neurons (depending on whether it originates in an excitatory or inhibitory fiber). The specific neurons are selected randomly in accordance with the assumption of local randomness. Let the total number of afferent fibers be A . They may be considered to be axons of A neurons in another part of the network. Let a fraction h of the fibers be inhibitory and let k^+ and k^- be the average coupling coefficients of the synapses made by the afferent excitatory and inhibitory fibers respectively. Further, let σ be the fraction of active fibers, i.e. those carrying action potentials at a particular instant.

Figure 32: Universal curve with no inhibition from equation (6).

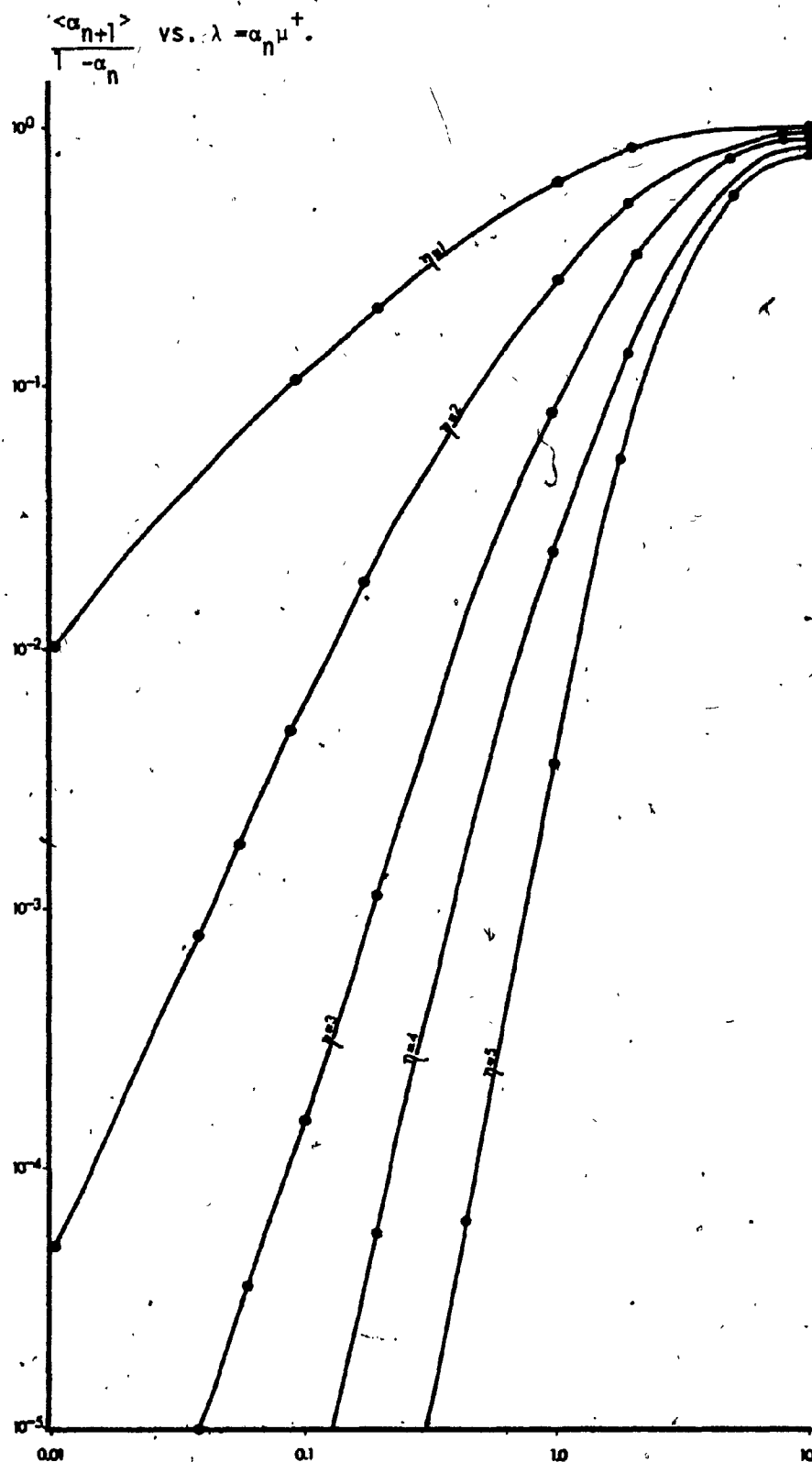


Figure 33. Universal curve with 30% inhibition from equation (6).

$$\frac{\langle \alpha_{n+1} \rangle}{1 - \alpha_n} \text{ vs. } \lambda = \alpha_n \mu^+.$$

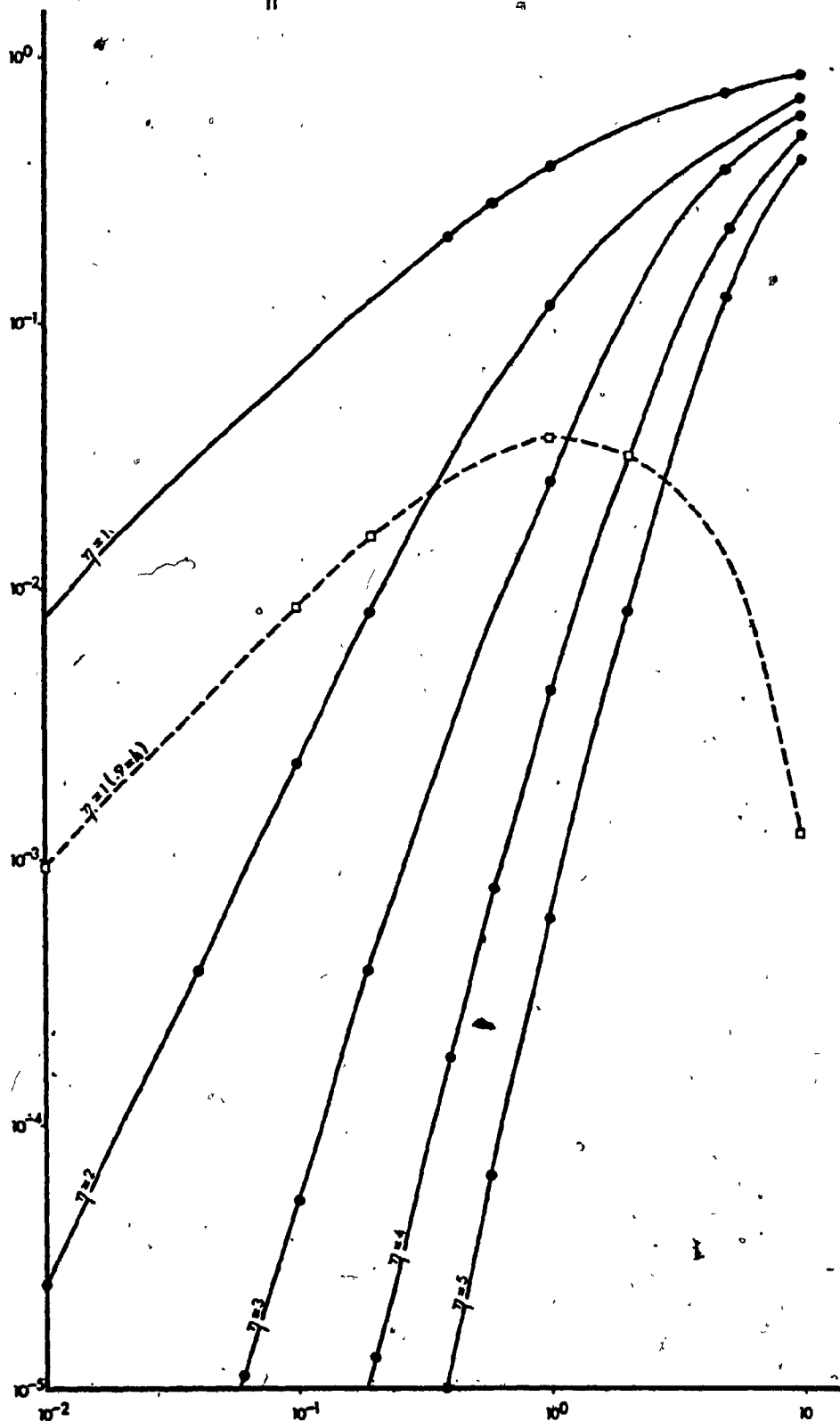
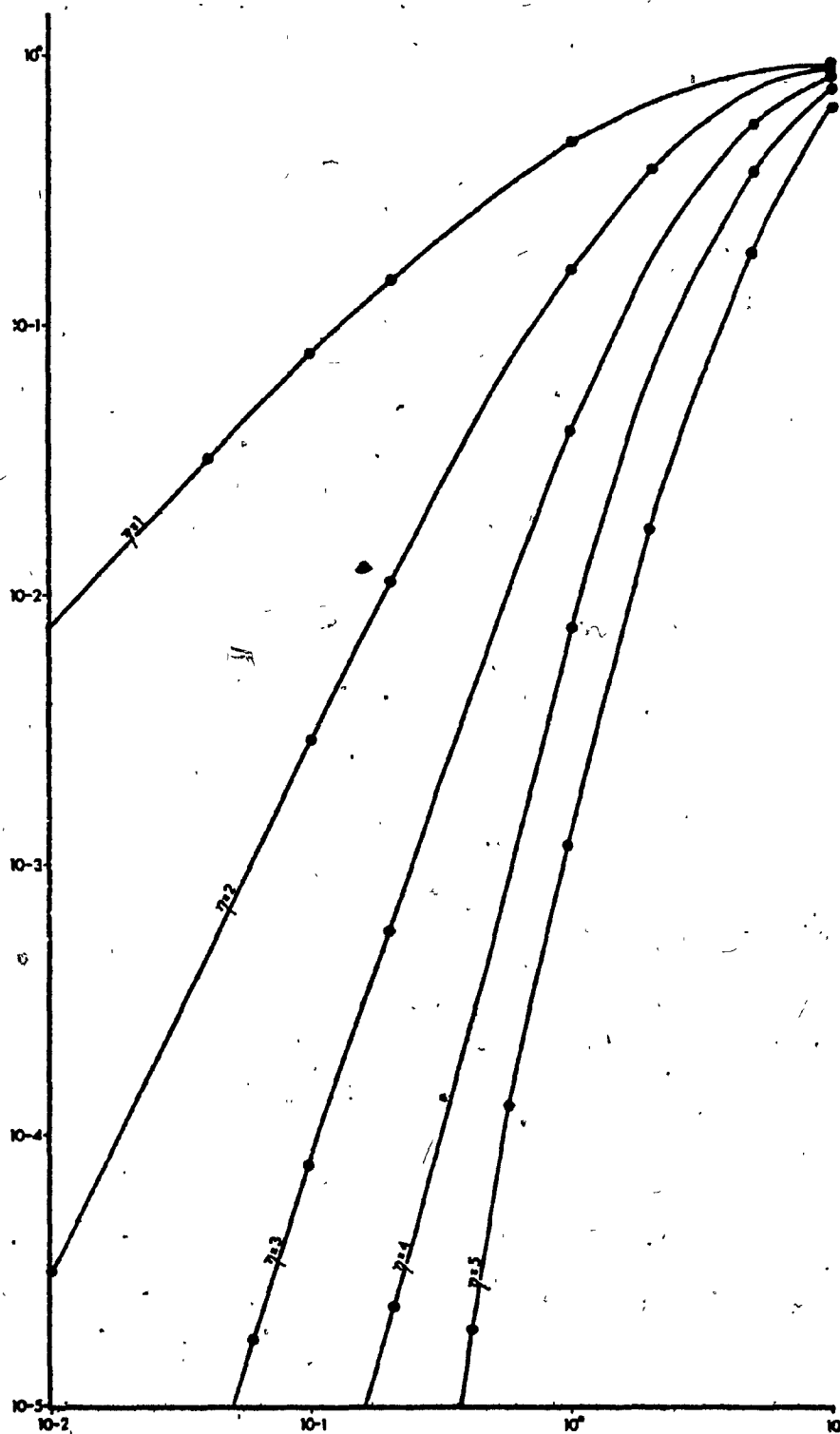


Figure 34. Universal curve with 20% inhibition.



We assumed that σ is constant for each experiment, but the number of active fibers is chosen randomly at each time interval. The expectation value of the activity $\langle \alpha_{n+1} \rangle$ for such steady input was calculated in a way similar to the cases of isolated netlet. The expectation is given by an equation similar to (6); the explicit form of which has been given in (Anninos et al, 1970). Thus:

$$\langle \alpha_{n+1} \rangle = (1 - \alpha_n) P(\alpha_n \sigma) \quad (7)$$

Equation (7) has been used extensively to compute the dynamic properties of netlets with steady inputs for a variety of parameters. A schematic of the general behaviour when a purely excitatory input ($h=0$) arrives into a Class B netlet is shown in figure 35. The form of the curve is seen to be similar to that of an isolated netlet which is now the special case $\sigma=0$. Upon increase of σ , the "ignition" point A is shifted downward while the point of steady activity, B, is raised. At the same time it can be seen from figure 35 that a second state of steady activity exists at 0. This point will move toward zero for low σ . When these equilibrium points are plotted versus σ , we obtain a diagram which we call "phase diagram" (Anninos, 1969; Anninos et al, 1970). Each curve in these diagrams may be thought of as dividing the plane into two areas on the right and on the left of the curve. The right-hand region is one of rising activity, the left-hand of decreasing activity. As may be seen from these curves, we have single-valued curves (solid lines) and multi-valued curves. The multi-valued curves have stable and unstable portions, while the single-values curves only have stable points. Another characteristic of these curves is that there exists a critical value of σ_{crit} such that when

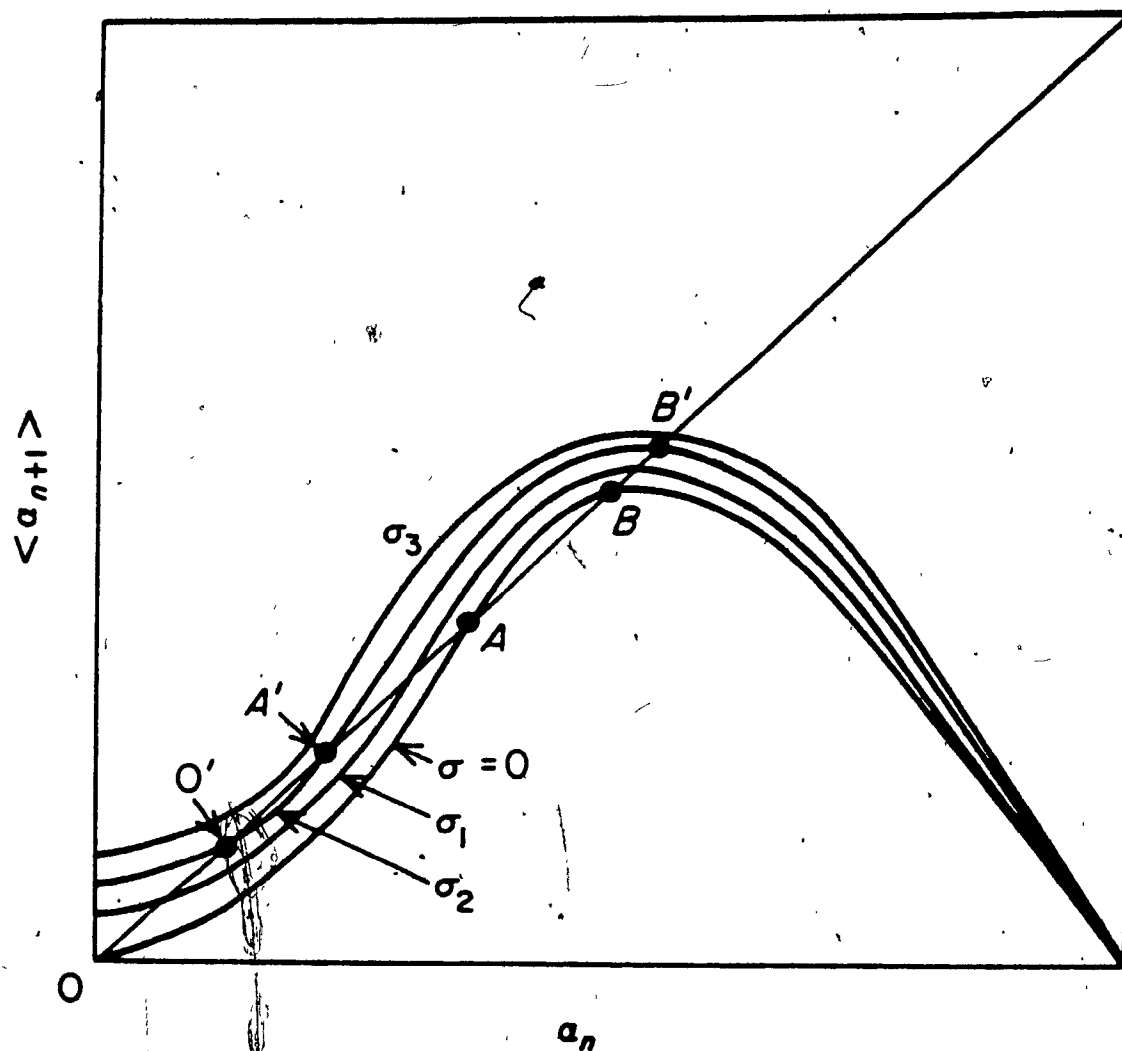


Figure 35. $\langle a_{n+1} \rangle$ vs. a_n for a class B netlet receiving sustained excitatory input (schematic). Note the shift of the three intersection points O , A , B with the 45° line as a function of σ .

the system is at its low value of steady activity, a sharp transition to a high steady state value occurs when σ is made to exceed σ_{crit} . Another such discontinuity occurs at a lower value of σ_0 (figure 36).

The striking aspect of this analysis is the appearance of hysteresis loops. A slow change of the level of afferent inputs leads to irreversible changes in the steady state activity of the netlet. This is best illustrated in figure 36. Here a Class C netlet is subjected to sustained excitatory input σ . The quantity α_{ss} is the steady state value of the netlet obtained by requiring $\langle \alpha_{n+1} \rangle = \alpha_n$. Figure 36 shows what happens if we change the level of σ . The resulting hysteresis loop has a lower and upper reversible portion (solid lines) linked by the irreversible upward transition at $\sigma = \sigma_{crit}$ and the downward transition at $\sigma = \sigma_0$. The steady input model described above thus furnishes a possible mechanism for explaining the tonic effect of one portion of the neural system upon another. The dotted lines in figure 36 represent the threshold for triggering the netlet into a state of higher sustained activity. As may be seen in the figure, this threshold depends very sensitively on the level of steady input. Figure 37a combines in one phase diagram the effects of purely excitatory ($\sigma+$) and purely inhibitory ($\sigma-$) steady inputs. It will be seen in this figure that certain Class C netlets labelled C_2 have no unstable portions and hence no hysteresis. Another interesting feature is that Class A and B netlets require inhibitory inputs in order to extinguish the high steady activity once it is established. The effect of varying the σ 's are shown in figure 37b. Netlets A, B, and C_1 are seen to have hysteresis loops, while C_2 behaves reversibly over the entire range.

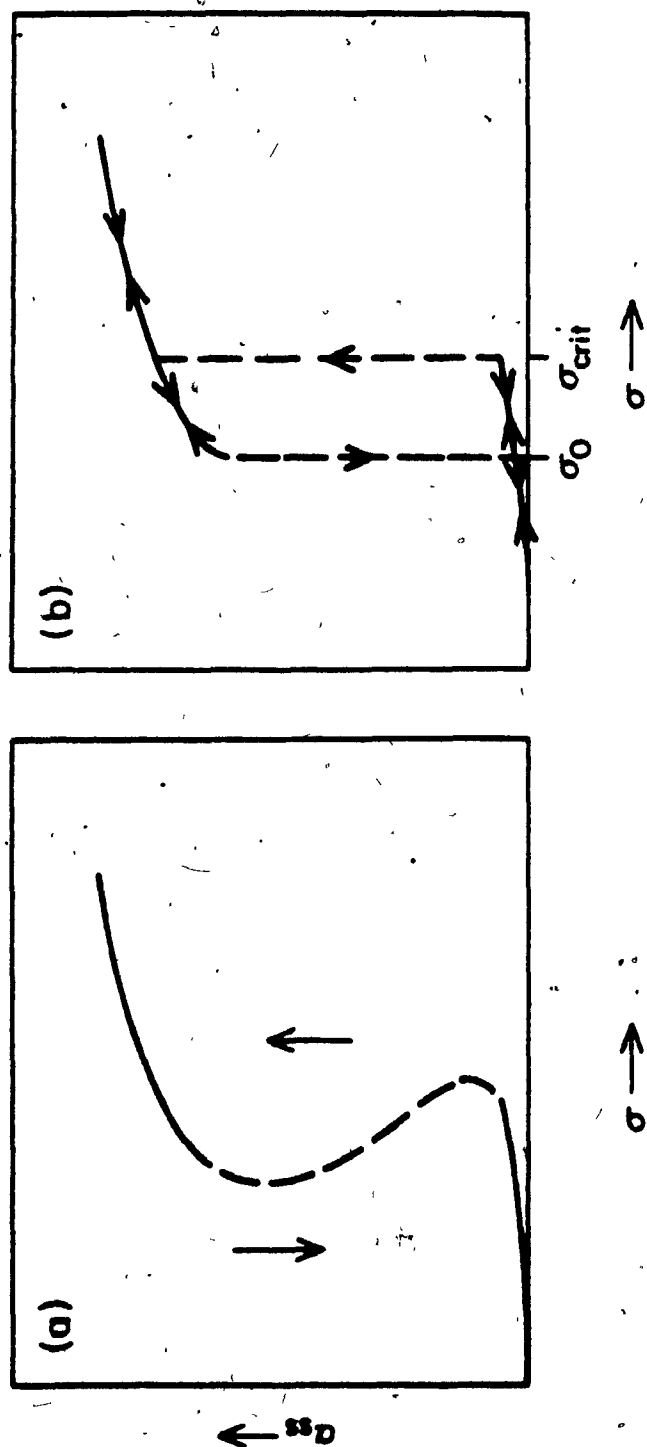


Figure 36. Steady state activities for a netlet with sustained inputs. (a) Steady state values (corresponding to points 0 and B in figure 35) are shown as solid lines, the metastable states (A in figure 35) are the dotted line. Arrows show direction of activity change for non-steady state values of activity. (b) Hysteresis loop generated by slow changes in σ , the fraction of active afferent fibers. Irreversible transitions are shown as dotted lines.

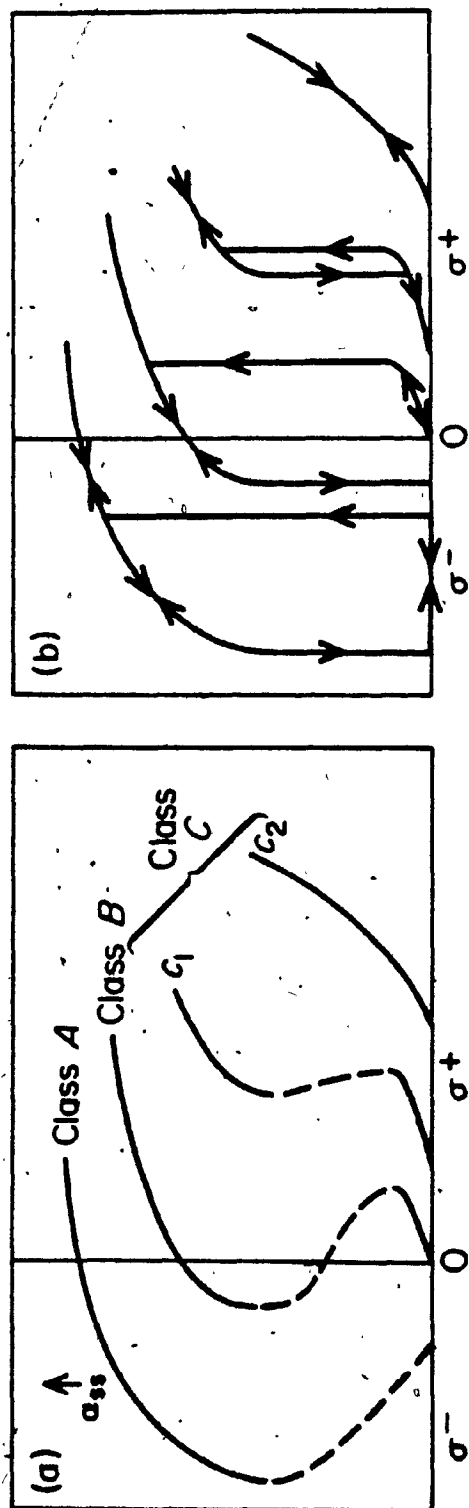


Figure 37. Steady state activities for different netlet classes, subject to sustained inputs. The afferent fibers are either inhibitory (left side of diagrams.) (a) Steady state and metastable state values (dotted) as function of afferent activity. (b) Hysteresis curves corresponding to the curves given in (a).

The results mentioned above are of considerable interest since systems which utilize hysteresis cycles and are characterized by their ability to retain an irreversible imprint of previous experience are known to exist in both physical and biological systems. The behaviour of Class A, B, and C₁ netlets belongs to the category of time-independent irreversible systems which are often invoked as the physical basis for the long term memory. According to this view, the short term memory is represented by sustained activity within a net. Such sustained activity may cause a number of connections to become reinforced (Hebb, 1949; Eccles, 1966). In an analogous manner in artificial nets, if the activity shifts ergodically in time, an increase in synaptic coupling coefficients k_{ij} may be expected. If we accept this mechanism of synaptic facilitation, we may expect that sensory nets stimulated most frequently, become excitable as a result of their increased coupling coefficients and the consequent decrease in the parameter n . Thus, given adaptable nets, the short term memory of an event represented by the sustained activity in nets, as described above, may also generate long term memory consisting of heightened excitability of the same nets, (Anninos, 1972a).

CHAPTER VII

THE PROBABILISTIC NEURAL NET: A SET-THEORETICAL
APPROACH TO THE REPRESENTATION OF NEURAL BEHAVIOUR

A. Introduction

Harth (1966) has devised a recursive algorithm for the behaviour of such nets. We shall use it here to discuss the operation of the net. Harth assumes that between the sensory cortex and the association cortex one has one-way and invariant connections. One assumes that sensory transducers trigger a sensory state \vec{s}_0 , which in turn, because of the connections between the sensory cortex and the association cortex, triggers an association set \vec{a}_0 one synaptic delay later. The mapping $\vec{s}_0 \rightarrow \vec{a}_0$ is invariant and genetically determined. \vec{a}_0 in its turn (the initial cortical state), by virtue of its interconnections with other cortical neurons, will produce \vec{a}_1 and so on. So one has

$$\vec{s}_0 \rightarrow \vec{a}_0 \rightarrow \vec{a}_1 \rightarrow \dots \rightarrow \vec{a}_n \rightarrow \vec{m}_0 \rightarrow \vec{m}_1 \dots$$

where \vec{m}_i are motor states.

Since the cortical states from \vec{a}_0 on depend on the value of the coupling coefficients k_{ij} which belongs to association area synapses and are supposed to exhibit plasticity, one can no longer assume invariance in the mappings of the chain.

B. Learning

We assume following Hebb (1949) that synaptic facilitation is the mechanism by which memory is acquired. In the words of Hebb (1949) "repeated stimulation of specific receptors will lead slowly to the formation of an assembly of association-area cells which can act briefly

as a closed system, after stimulation has ceased, this prolongs the time during which the structural changes of learning can occur and constitutes the simplest instance of a representative process (image or ideal)". In our model we used two modes to accomplish this. The coupling coefficients k_{ij} are increased by an amount δ as a result of learning in the following fashion:

$$I \quad \delta_{ij} = \begin{cases} \delta & \text{iff } a_n^i = a_n^j = 1 \text{ and } k_{ij} > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$II \quad \delta_{ij} = \begin{cases} \delta & \text{iff } a_n^i = a_{n-1}^j = 1 \text{ and } k_{ij} > 0 \\ 0 & \text{otherwise} \end{cases}$$

Rule I, Harth and Edgar (1967) and Csermely (1968), might be considered as exemplifying the synergetic effect of neurons in a collective mode of excitation; it simply says that if state \vec{a} is repeatedly triggered, intrastate connections k_{ij} will be strengthened. Rule II, states that in a sequence of cortical states only those coupling coefficients will be increased which have contributed to the firing of neurons. Rule II is the one that more accurately reflects Hebb's (1949) ideas of synaptic facilitation. Both of these rules were used in our computer simulation of learning behaviour in probabilistic neural nets:

C. Information Processing in the Association Net

Information in our view is assumed to be carried by the binary code form of the cortical state vector \vec{a} . This would fit some ideas expressed by Eccles (1973), Verzeano (1970) and others.

This cortical state can arise as a result of either (i) direct triggering from a sensory state \vec{s} or (ii) from some other cortical state antecedent to \vec{a} . We now examine the effect of applying the learning rules. This has been investigated in the work of Harth (1966), Harth and Edgar (1967) and Csermely (1968). However, we will here consider the problem from a slightly different angle.

For the sake of concreteness let us use rule I: Learning of state \vec{a} from the point of view of rule I means that all intracortical coupling coefficients for which $a_n^{(i)} = 1$ and $a_n^{(j)} = 1$ must be increased. In other words, if $a_n^{(i)}$ and $a_n^{(j)}$ are connected and if $a_n^{(i)} \in \vec{a}$ and $a_n^{(j)} \in \vec{a}$ then

$$K_{ij} \rightarrow \delta_{ij} + k_{ij}$$

The effect of the reinforcement on the connectivity matrix can be pictured according to Harth and Edgar (1967) as follows:

We can first of all, relabel all neurons in the net in such a fashion that neurons belonging to a given state that has been reinforced have consecutive labels. Thus net A can be decomposed as follows:

$$A = \{a^{(1)}, a^{(2)}, \dots, a^{(L)}; b^{(1)}, b^{(2)}, \dots, b^{(M)}; \dots; c^{(1)}, c^{(2)}, \dots, c^{(N)}\}$$

In this way one creates in the connectivity matrix square blocks of coefficients corresponding to the distinct states \vec{a} , \vec{b} , \vec{c} which have been reinforced. In these smaller blocks one finds the reinforced coupling coefficients belonging to the different learned states \vec{a} , \vec{b} , \vec{c} . If the rest of the cortex is naive and if our blocks are cross-hatched then the pattern we get looks like figure 38. Areas that are not cross-hatched contain coefficient k_{ij} which have not changed their

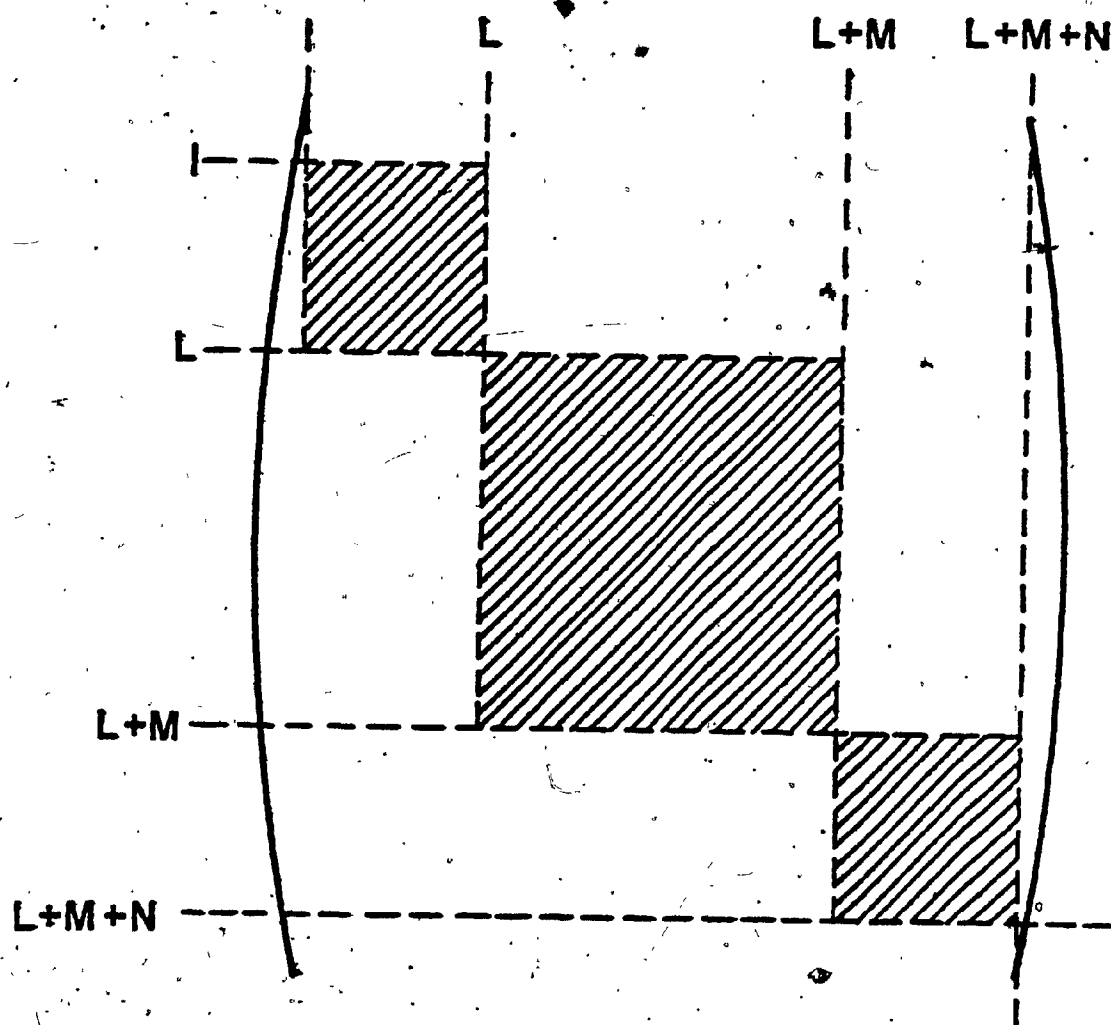


Figure 38. Schematic diagram for the connectivity matrix k_{ij} , with reinforced domains shown as hatched lines.

values. Further we assume that

$$\vec{a} \cap \vec{b} \cap \vec{c} = 0$$

It should also be apparent that although the matrix shows symmetry with respect to its principal diagonal this does not extend to the microstructure of L, M, N since in general $k_{ij} \neq k_{ji}$ since connections in the cortex are supposed to be one way and asymmetrical. Consider now the following two theorems due to Harth (1967) and their implications.

Theorem I. Reinforcements of a cortical state \vec{a} will not change the mapping $\vec{S} \rightarrow \vec{a}$. (assuming \vec{S} originally maps on to \vec{a}).

Proof: Assume that we have the mapping $\vec{S} \rightarrow \vec{a}$. Learning according to rule I means that we search for all coefficients k_{ij} for which $a_n^{(i)}$ and $a_n^{(j)}$ belong to \vec{a} and we adjust k_{ij} so that $k_{ij} = k_{ij} + \delta_{ij}$. Consider now a neuron $a^{(k)} \notin \vec{a}$. Since it was not triggered while \vec{a} was active all coefficients k_{kl} for which $a^{(l)} \in \vec{a}$ will retain their previous value. Since these connections did not become active to the extent of triggering $a^{(k)}$, $a^{(k)}$ will not belong to \vec{a} , either before, or after reinforcement of state \vec{a} .

Theorem II. Consider two maps $\vec{S}_a \rightarrow \vec{a}$ and $\vec{S}_b \rightarrow \vec{b}$ so that $\vec{a} \cap \vec{b} = 0$. If \vec{a} and \vec{b} are reinforced then the disjointness property of \vec{a} and \vec{b} is invariant under reinforcement. Here \vec{S}_a, \vec{S}_b stand for the sensory states that trigger \vec{a} and \vec{b} respectively.

Proof: Take an element $b^i \in \vec{b}$ but $b^i \notin \vec{a}$ before reinforcement. Reinforce \vec{a} . By Theorem I since b^i did not belong to \vec{a} originally it cannot belong to \vec{a} afterwards. The same holds true for every element of \vec{a} with respect to \vec{b} .

What however, is the effect of the learning rule on two states \vec{a} , \vec{b} which have the property that $\vec{a} \cap \vec{b} \neq 0$. Case I: only \vec{a} is reinforced. We have first, the topological schematic shown below.

(Schematic I, figure 39).

We now reinforce \vec{a} (Schematic II, figure 40).

After this we represent the system with \vec{s}_b . (Schematic III, figure 41).

The relationship between \vec{a} and \vec{b} now changes. Consider a point $a^i \in \vec{a}$ in Schematic II, with the additional property that $a^i \in \vec{b}$ of Schematic I. All connections ending on a^i and including those that belong to $\vec{a} \cap \vec{b}$ will be reinforced. Consequently, if now we present \vec{s}_b as per Schematic III the possibility arises that there will exist points $b^j \in (\vec{b} \cap \vec{a})$ for which k_{ij} exists (i.e. a connection exists) and for which $k_{ij} \rightarrow k_{ij} + \delta_{ij}$. In other words, a^i could be triggered. The effect in general would be a growth of \vec{b} into the state \vec{a} that has been reinforced. It will be seen later on, that this property of the boolean overlap of two states, at least one of which has been reinforced, to grow, will have profound consequence for our hypotheses concerning association and memory.

Case II: Both \vec{a} and \vec{b} are reinforced.

This corresponds to the classical conditioning experiment.

Assume: $\vec{s}_a \rightarrow \vec{a}$ and $\vec{s}_b \rightarrow \vec{b}$. Further, suppose we originally have $\vec{a} \cap \vec{b} = 0$. (Schematic IV, figure 42).

Suppose now we consider the map. $(\vec{s}_a \cup \vec{s}_b) \rightarrow (\vec{a} \cup \vec{b})$ and further suppose we reinforce $(\vec{a} \cup \vec{b})$. (Consider Schematic V, figure 43).

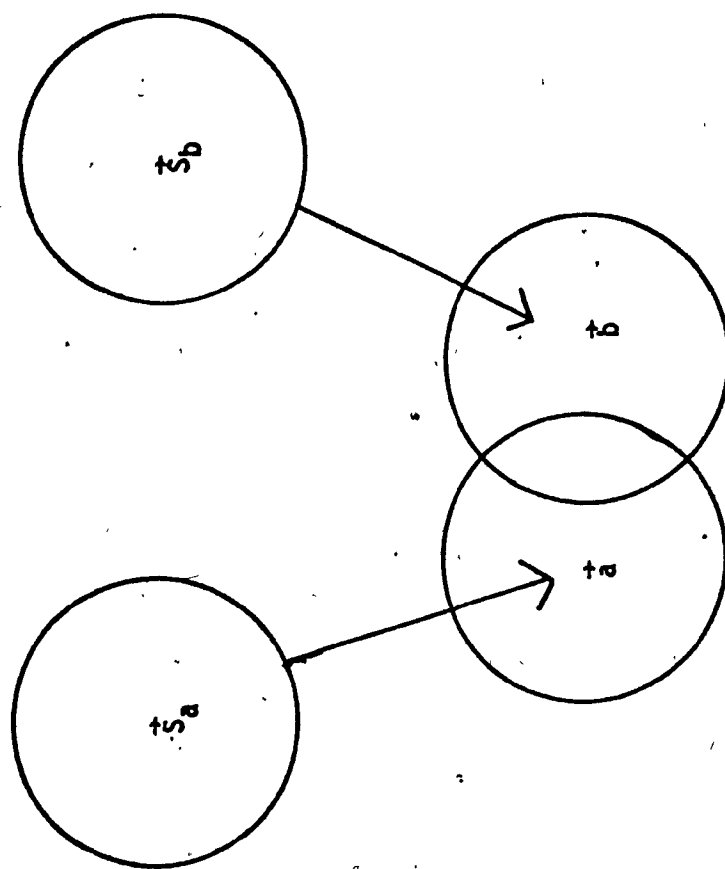


Figure 39. Schematic diagram showing the mapping of stimulus sets \vec{S}_a and \vec{S}_b being mapped onto two intersecting sets \vec{a} and \vec{b} .

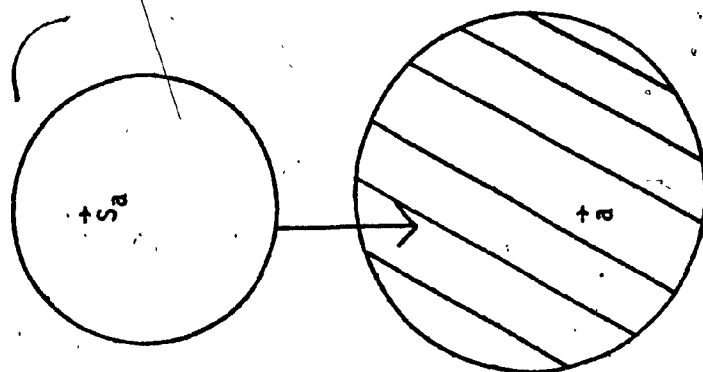


Figure 40. Schematic diagram showing the reinforced set \vec{a} that has 'learned' \vec{S}_a . Darkened areas imply reinforcement.

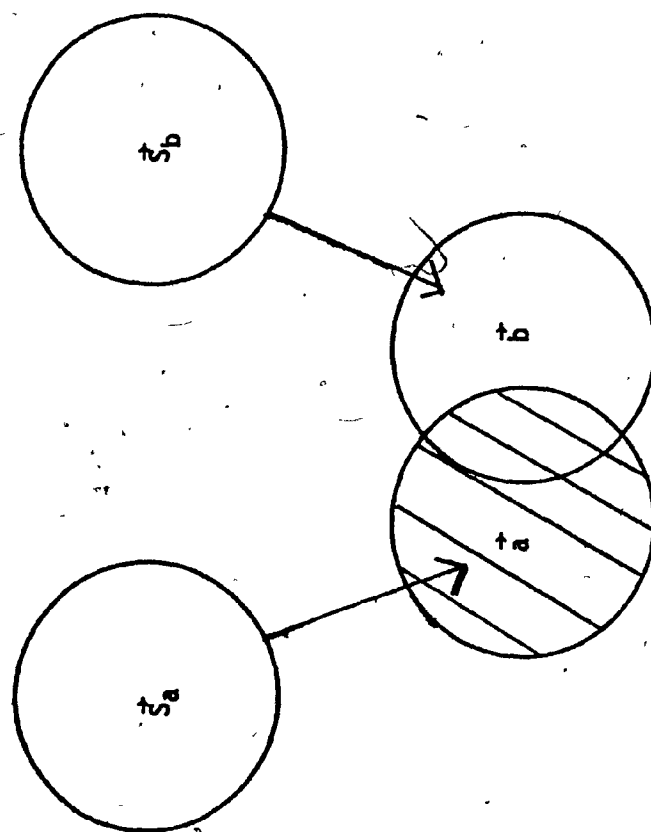


Figure 41. Schematic diagram showing the map of \vec{S}_a on \vec{S}_b .
 Note: $\vec{a} \cap \vec{b} \neq \emptyset$.

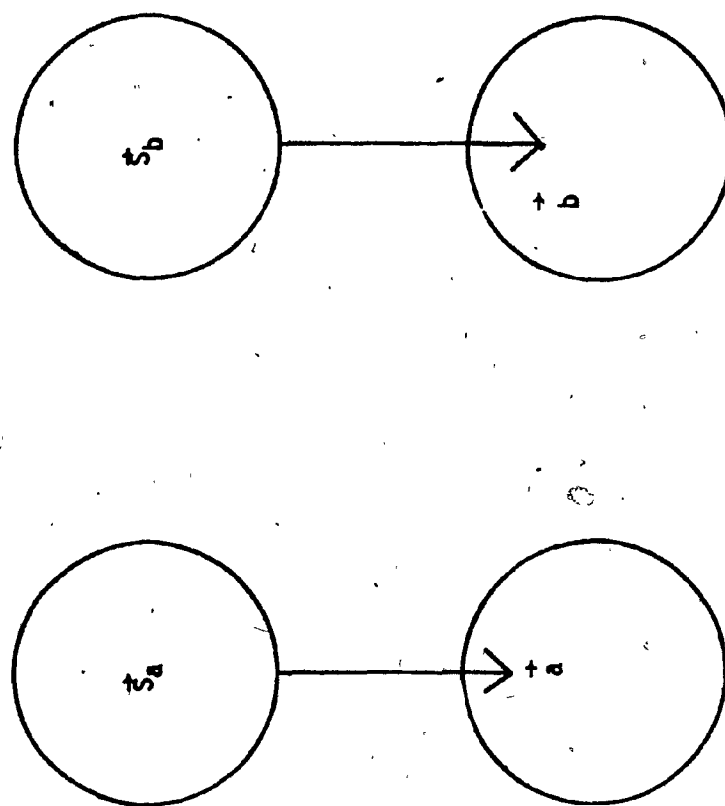
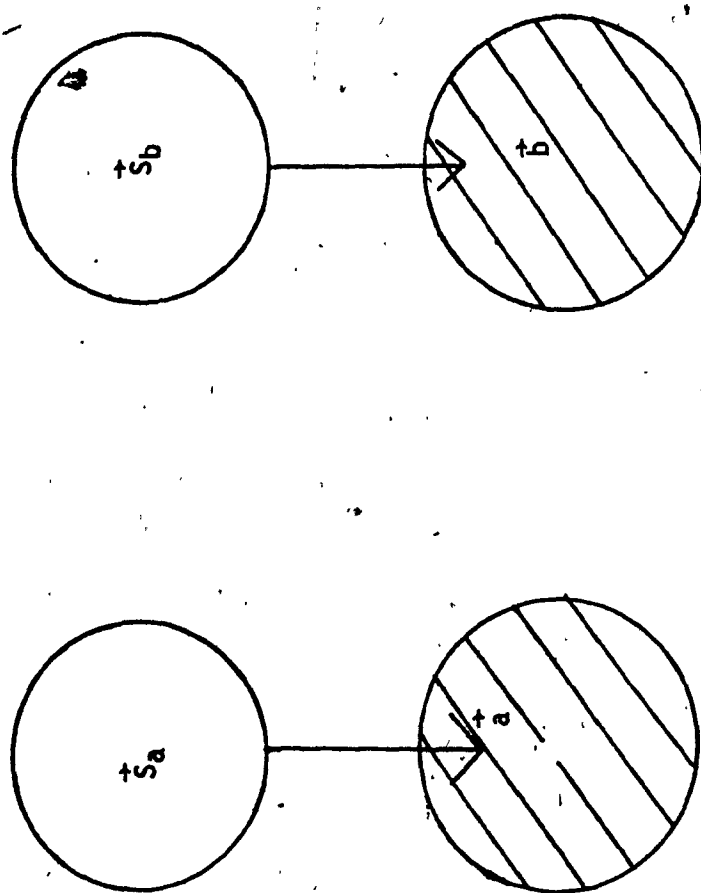


Figure 42. Schematic diagram showing the mapping of two stimulus sets \vec{s}_a and \vec{s}_b being mapped onto \vec{a} and \vec{b} . Note: $\vec{a} \cap \vec{b} = 0$.



• Figure 43. Schematic of the map of figure 42 after reinforcement of \vec{a} and \vec{b} .

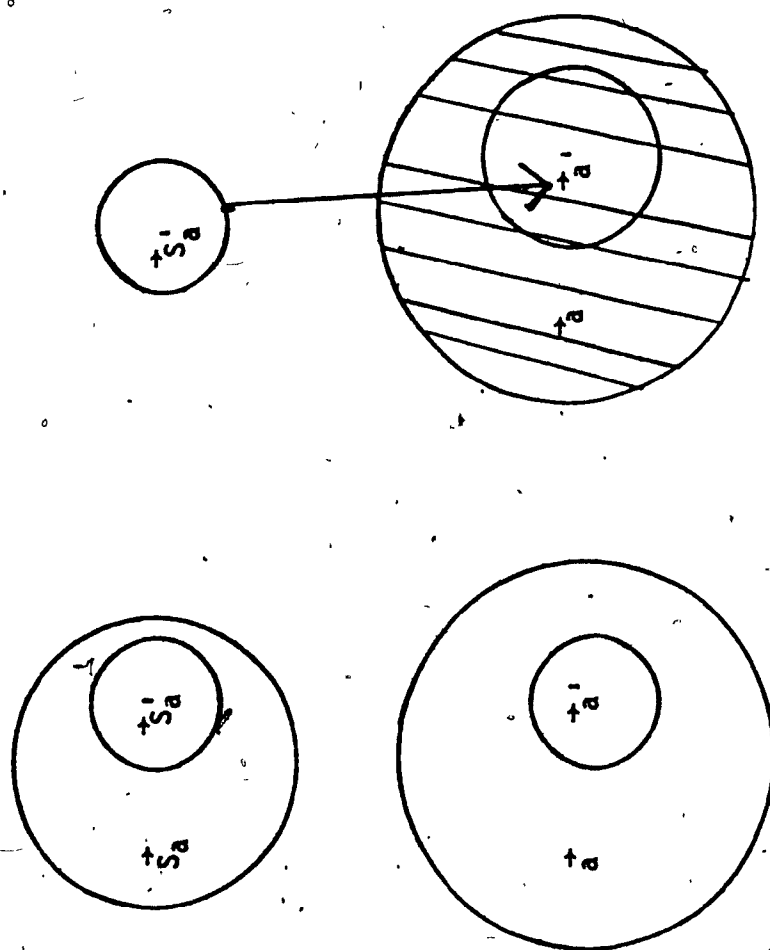


Figure 44. Schematic showing process of pattern completion.

Note: $a' \subset a$.

Because of the topological assumption we have made concerning the connectivity of the brain we can make the following two statements. On the one hand, there exist connections between \vec{a} and \vec{b} joining some elements a^i and b^j belonging to \vec{a} and \vec{b} respectively. This implies that in the connectivity matrix $k_{ij} \neq 0$ for these elements. On the other hand, if the maps $\vec{S}_a \rightarrow \vec{a}$, $\vec{S}_b \rightarrow \vec{b}$ were such that $\vec{a} \cap \vec{b} = \emptyset$ this means that inputs from \vec{a} to \vec{b} and vice-versa would be subthreshold. This, however, need not be the case after reinforcement. In fact, we now have the possibility, that presentation of \vec{S}_b alone would trigger elements in \vec{a} i.e. states that were originally completely disjoint now develop an overlap.

An Important Corollary:

Consider again the map $\vec{S}_a \rightarrow \vec{a}$ and further take an \vec{S}'_a such that $\vec{S}'_a \subset \vec{S}_a$. Obviously the map $\vec{S}'_a \rightarrow \vec{a}$ is such that $\vec{a}' \subset \vec{a}$. If however, the map $\vec{S}_a \rightarrow \vec{a}$ is reinforced, the new \vec{a}' will be larger than the old one. We call this pattern completion. See Schematic V. Figure 44. The above considerations can be correlated with various cognitive functions.

1. Binary Association and the Simple Conditioned Reflex

"Harth and Edgar (1967); Csermely (1968).

Psychological theories seem to imply that cognition is impossible without the ability of the brain to produce association either between various stimuli or between various ideas and concepts. Association is a necessary but not sufficient condition for the solution of the serial order problem of cortical function. See Lashley (1951).

Let us consider just the simple problem of association of two stimuli \vec{S}_a and \vec{S}_b . How would this come about in our scheme? It is obvious that the answer lies in Case II of the previous section.

We have:

$$\vec{S}_a \rightarrow \vec{a}$$

$$\vec{S}_b \rightarrow \vec{b}$$

Simultaneous presentation of \vec{S}_a and \vec{S}_b means that we have a new stimulus

$$\vec{S}_c = (\vec{S}_a \cup \vec{S}_b) \text{ such that}$$

$$\vec{S}_c \rightarrow \vec{c}$$

Learning of \vec{c} is known as binary association of \vec{a} and \vec{b} . If we now present stimulus \vec{S}_a by itself we will not produce \vec{a} but \vec{a}^- where according to case II:

$$\vec{a} \subset \vec{a}^- \text{ and } \vec{b} \cap \vec{a}^- \neq 0$$

In the connectivity diagram we have a situation where instead of the diagonal hatched blocks we have the set up of figure 45. The rectangles become hatched because they represent the cross coefficients between \vec{a} and \vec{b} which now become strengthened. The above kind of logic has been used by Csermely (1968) to interpret the so-called switching reflex of Asratian (1968) and by Edgar and Harth (1967) to produce various types of association phenomena such as language learning and other classes of conditioning.

2. Pattern Completion and Overlap Growth due to Learning

From the point of view of present work however, pattern completion is more important as it gives us a way of understanding some aspects of memory mechanism. As Norman (1968), Schneider and Shiffrin (1977) and others have pointed out, one important aspect of memory is the

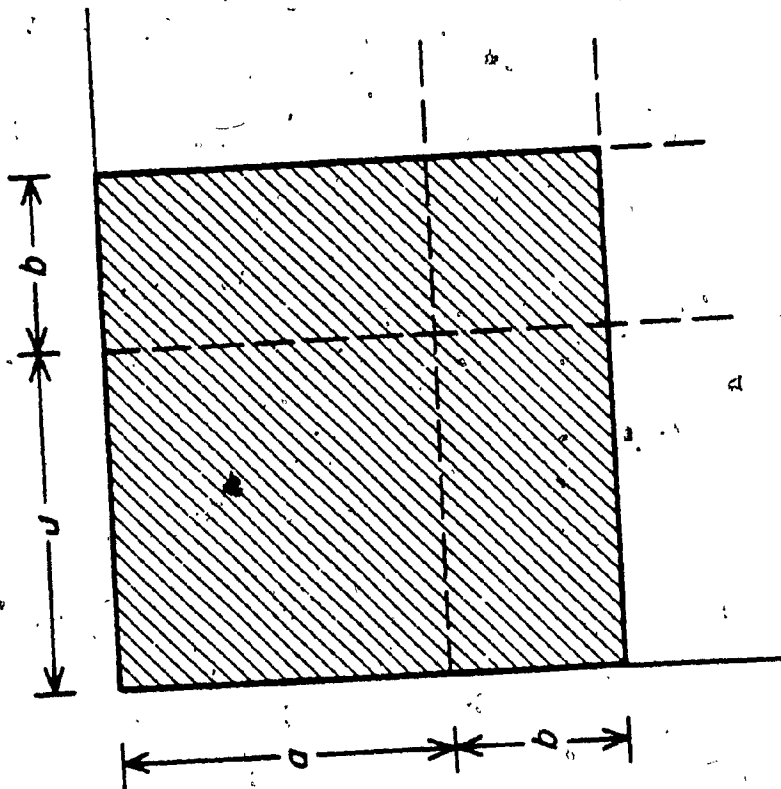


Figure 45. Binary association of two stimuli as reflected on the connectivity matrix.

ability of this system to store new information by implanting it or embedding it into the memory bank. This is done by cross-referencing the new in terms of the old and already acquired material. From the point of view of set theory if X is some new input into the system then cross-referencing is equivalent to producing the following:

$$X = \bigcup_i (X \cap M_i)$$

where M_i are permanent memory elements.

Once this embedding has occurred which from the point of view of Hebb (1949) means that an assembly has been created, one can present even fragments of the new object and bring about recall of the old. One can refer in this regard to Gestalt completion phenomena as well. Milner (1970). Similarly one can obtain recognition of a pattern if a cue is given. More importantly, from the point of view of research under way, however, we should like to discuss pattern completion. Recall that we have two situations of overlap growth. In one case we have two stimuli \vec{S}_1 and \vec{S}_2 such that $\vec{S}_1 \supset \vec{S}_2$. See Schematic V (Figure 44). We found that if $\vec{S}_1 \rightarrow \vec{a}_1$, and $\vec{S}_2 \rightarrow \vec{a}_2$ in such a way that $\vec{a}_1 \supset \vec{a}_2$, then, following reinforcement, if \vec{S}_2 is presented, one obtains the map, $\vec{S}_2 \rightarrow \vec{a}_2$ with the property that $\vec{a}_2 \supset \vec{a}_2$. In other words, the sensory fragment tends to trigger memory units that would correspond to the whole object. The above find empirical verification in the work of Noton and Stark (1971). The reader is referred to the on Hebb for a full description of this experiment.

Of interest in our work is that 70% of the number of original fixations were necessary for pattern completion. This result closely agrees with our computer simulation results. (See Calculations and Results).

We now consider overlap growth when the two stimuli \vec{S}_1 and \vec{S}_2 give rise to two states \vec{a} and \vec{b} such that $\vec{a} \cap \vec{b} \neq 0$. Assume \vec{a} is learned and \vec{S}_2 is again presented. We showed that in general one would get a new state \vec{b}' such that $(\vec{a} \cap \vec{b}') \supset (\vec{a} \cap \vec{b})$.

If \vec{a} is the memory store of stimulus \vec{S}_1 , then the last result would seem to imply that reaction to stimulus \vec{S}_2 would be a mixture of readouts from \vec{a} and \vec{b} . This kind of result finds verification in the experiments of Ruchkin and John (1966). In these experiments, an animal was conditioned to respond to some flicker frequency f_1 . The animal was subsequently tested with another frequency f_2 . The evoked potentials for both f_1 and f_2 frequencies were registered. The initial f_2 evoked potential was found to be quite similar to that of f_1 . John (1972) interprets this as a readout from the memory store for f_1 .

CHAPTER VIII

CALCULATIONS OF NET BEHAVIOURA. Introduction

We found a combination of mathematical analysis based upon equations (1) and (7) and computer simulation to be the best methods for investigating the behaviour of the nets we studied.

The present investigation makes use of the neuronal simulation approach developed in previous research (Csermely, 1968; Anninos, 1969; Anninos et al, 1970). A brief description of the approach is given in the following paragraph.

B. Outline of the Simulation Program

Given a net consisting of A neurons, the first step is to establish a neuronal connectivity matrix $\{K_{ij}\}$. (Warth and Edgar, 1967). Each of the non-zero elements of the connectivity matrix is characterized by three values: the serial number of the neurons on which the connection terminated (i); the number of the neuron from which the connection originates (j); and a coupling coefficient (k), which may be either zero or non-zero. A coupling coefficient of zero denotes the absence of a synaptic link. A coupling coefficient which is either positive or negative indicates that the neuron is excitatory or inhibitory, respectively. Therefore, since the values of k are either zero or non-zero we have in general an asymmetric matrix $\{K_{ij}\}$. In the computer program utilized, the user specifies the limits within which i and j may vary, generating in this way the coupling coefficients for each net.

An additional significant feature is that the number of outgoing connections for each neuron need not be constant; it may be chosen randomly between two limits, μ_{\min}^+ and μ_{\max}^+ for excitatory neurons, μ_{\min}^- and μ_{\max}^- for inhibitory ones. Finally, the fraction of inhibitory neurons in a pool of neurons is specified by a parameter h . The specific neurons which will be inhibitory are determined during the initialization process, before interconnections are established, by randomly selecting the appropriate number of neurons consistent with h . This information is stored and later, when the value of k is computed for a given connection, a negative sign is assigned to k whenever the connection originates on an inhibitory neuron. In establishing the pattern of connections, extensive use is made of a random number generator. This subroutine generates a uniform distribution of random numbers between 0 and 1. An integer number selected by the user serves as first entry; this number determines the specific sequence of random numbers obtained from the program. Thus, when used in setting up the connections, this procedure determines the specific structure of the net. It is clear that by using different initial numbers for the subroutine, the detailed "microscopical" structure of the net may be modified, while the statistical parameters n , μ^+ , μ^- , h and k are unaltered. This feature of the present simulation procedure is of great importance in allowing comparison of the relative significance of the specific "microscopical" net structure and of the statistical parameters of the net.

When all desired interconnections have been established and stored in the computer, the net is activated by specifying the set of neurons which fire at $t=0$. One synaptic delay later, all neurons

which are linked to those initially active are found through reference to the connectivity matrix. These neurons now receive excitatory, or inhibitory input, as the case may be. The algebraic sum of the coupling coefficients characterizing the active incoming connections represents the excitation level in any of these neurons. This information is stored by the program and whenever the sum exceeds the threshold (see below) the neuron would fire and, following one additional synaptic delay, become a source of excitation - or inhibition - for all neurons connected to it. Otherwise, all the active incoming connections return to the initial value in one additional synaptic delay and disappear immediately thereafter. When a neuron fires, its threshold is raised to a maximum value and remains at this level for the specified duration of the absolute refractory period, during which time the neuron receives excitation but cannot fire due to its high threshold. Following the absolute refractory period the neuron returns to its normal threshold and is capable of firing once more, provided that the input is equal to, or greater than, the threshold (i.e. there is neither a relative refractory period nor post-spike facilitation).

Learning in the net is accomplished by either of the two rules. In the simulation program, rule I is accomplished as follows: The program checks and finds if a particular neuron has fired in any of the firing cycles. The result is stored in an array called Join. Learning is accomplished by incrementing the coupling coefficients between those neurons that have fired at least once during the previous cycles. In the case of rule II one checks if a particular neuron has fired in cycle n . It then notes all neurons that fired in cycle $n-1$. If these are connected to the neuron that fired in cycle n , the

appropriate coefficients are incremented. Our program could handle up to 1000 neurons. Selected portions of the net were chosen as stimuli. These were triggered and those neurons that become active were identified for each cycle and were subsequently listed in the output. (One list for each firing cycle). Some of these results are tabulated in table 1. Nomenclature is chosen to conform to the notation of section V.

C. Pattern Completion

As we mentioned before, pattern completion is the result of overlap growth. This was accomplished in our model as follows:

Neurons 1-100 constituted stimulus \vec{S}_1 , whereas neurons 1-50 represented stimulus \vec{S}_2 . We reasoned that a stimulus overlap of 50 neurons would produce enough of an overlap so as to make the statistical fluctuation phenomena of the model not significant. Trigger of \vec{S}_1 caused the \vec{a} neurons to become active. Similarly, the \vec{b} neurons became active as a result of firing of the \vec{S}_2 neurons. The overlap $N(\vec{a}) \cap N(\vec{b})$, representing the number of common neurons between \vec{a} and \vec{b} was noted.

The net we used was highly damped. Activity died after one or two firing cycles. (See Harth and Edgar (1967)). All the coefficients of neurons that became active as a result of firing of \vec{S}_1 , were now incremented by an amount δ the reinforcement parameter. \vec{S}_1 and \vec{S}_2 were now fired again and the first cycle overlap was noted. The reinforcement parameter was changed and the process was repeated. For net I whose connectivity is produced by the random number generator whose initial input is the number 141935623.00 we made 15 runs. (Runs I(1) to I(15)). In runs I(6) to I(11) the results of a 20% inhibition on the net were studied. Both learning rules were applied. We then utilized the

Universal curves to ascertain whether the theoretical predictions agreed with the simulation procedures. Good agreement was obtained as shown in the following sample of calculation. Consider net II in runs 16 and 17. As we see, the overlap increased from 15 to 37 neurons.

This corresponds to an increase of 22 neurons. We use the Universal graph for no inhibition corresponding to equation (6). Here

$\alpha_n = \frac{N(\vec{a} \cap \vec{b})}{N(\vec{a})} = 15/52$, $\lambda = \alpha_n \bar{\mu} = 1.44$. From the graph we can have $\gamma = .38$ corresponding to an $n=2$ net after reinforcement or $\gamma = .7$ corresponding to an $n=1$ net. Utilizing the latter we obtain $\alpha_{n+1} = (52) (.7) (1-.29) = 25$.

Computer results as well as theoretical predictions are shown in table 2.

From the parameters of the net we see that we started with an $n=3$ net.

Upon learning the net became either an $n=2$ or an $n=1$. I.e., starting

with a C-type net one ends up with either an A or B type net. If we

call n the phase constant of the net we see that learning involves a

phase change. As we stated before the properties of A and B types

include the property of oscillation. In our computer runs oscillation

was observed as a result of triggering the reinforced \vec{a} net via its overlap $\vec{a} \cap \vec{b}$, which itself was the inevitable result of triggering \vec{b} .

In runs I(12) to I(15) as well as in runs II(20) to II(21) we tried

to simulate the experiments of Noton and Stark mentioned earlier on.

Neurons 1 to 100 stood for the original stimulus in the learning phase. These neurons were triggered and the set of neurons that

became active as a result had the appropriate K_{ij} 's strengthened. The

object was now "degraded" by the process of triggering only neurons

1 to 70. We found that the "degraded" object generated between 94 and

96 percent of the image of the original object. A special subroutine

was developed by this writer to generate the universal curves.

D. The Simple Conditioned Reflex:

Our results closely match those obtained by Harth and Edgar (1967) and by Csermely (1968). Neurons 1 to 100 stood for the US (Unconditional Stimulus), whereas 501 to 600 stood for the CS (Conditional Stimulus). These two sets were fired singly and in combination and the response sets they produced were duly noted. The net was then "taught" the combination 1 to 100 and 501 to 600. Following this we again triggered the 501 to 600 set. The response set that was produced exhibited cyclic behaviour. Of the 52 neurons that constituted the UR (Unconditioned Response) of the 1 to 100 net, 51 were found to belong to the two state oscillating cycle of CR (Conditioned Response), 18 were in one of the states and 32 were in the other. In other words, the CR activity oscillates within the space of the neurons of the UR as well.

E. Some Topological Aspects of Memory Acquisition in the Probabilistic Neural Net Modeling of the Association Cortex

Anninos (1969) has shown that sustained input, into A, B and C₁ type nets leads to hysteresis effects which are interpreted as the basis for memory acquisition by these nets. It will be recalled that these nets have thresholds for stable oscillation. When this threshold is reached the net begins to reverberate in various modes of oscillation. Reverberation within a net is the mechanism, postulated by Hebb (1949), leading to synaptic facilitation and hence to the formation of stable cell assemblies.

Csermely (1968) and Harth et al (1970) have proposed a netlet theory of brain dynamics capable of providing a mechanism for this sustained input. They postulate the existence of connections from

TABLE 1 PATTERN COMPLETION

Effects of Reinforcement state \vec{a} on growth of state \vec{b} . Net parameters are $A = 1000, K^+ = K^- = 10, \theta = 30, (n=3), \mu^+ = \mu^- = 5, N(\vec{S}_1) = 100$

Net	Random Number Generator	h	Learning Rule	$N(\vec{S}_1) \cap N(\vec{S}_2)$	$N(\vec{a}) \cap N(\vec{b})$	δ	$N(\vec{a})$	$N(\vec{b})$	Oscillation Type
1 (1)	141935623.00	0	2	50	15	0	55	15	None
2 (2)					47	20	55	47	2-state
3 (3)					47	40	56	47	"
4 (4)					48	60	55	48	"
5 (5)					47	80	55	47	"
6 (6)		20			12	0	33	12	None
7 (7)					27	20	33	27	2-state
8 (8)					27	30	33	27	"
9 (9)					27	40	33	27	"
10 (10)					27	60	33	27	"
11 (11)	145368237.00	0	1	70	27	80	33	27	"
12 (12)					31	0	52	31	None
13 (13)					49	30	52	49	2-state
14 (14)					31	0	55	31	None
15 (15)					51	30	55	51	2-state
16 (16)		0			15	0	52	15	None
17 (17)					37	60	52	37	2-state
18 (18)					37	40	52	37	2-state
19 (19)					30	30	52	30	"
20 (20)					30	0	52	30	None
21 (21)			2	70	50	30	52	2-state	

Table II Overlap Growth due to Learning

Net	Computer Result	Universal Curve
I (2)	30	30
I (3)	30	30
I (5)	30	30
I (17)	22	25
I (18)	22	25
I (19)	15	14
I (13)	18	19

neuron pools, known to exist in the sensory cortex to netlets in the association cortex. The association cortex outside the netlets is very damped so that no spillover from the embedded netlets can occur. In other words, no information is lost. Csermely's (1968) model involves the creation of high excitability netlets in the association cortex via oscillation of the sensory cortex netlets which are already formed. Anninos (1972b) has used a suggestion by John (1967) to interpret the pattern of reverberation of a particular net as a representation of a given stimulus. Anninos (1972a) has also elaborated a possible mechanism of read-in of stimuli. In other words, the reaction of the net to particular stimuli.

Little and Shaw (1975) and more recently Shaw (1978) using what they call a modified Hebb hypothesis also identify the reverberation patterns of various assemblies as representations of external stimuli. More precisely, a specific, cyclic, firing pattern, such as $x^a \rightarrow x^b \rightarrow x^c \rightarrow x^a$ would represent a given stimulus S. This activated pattern, incidentally, would also represent the short term memory of the event.

It is obvious however, that if one considers our results in pattern completion and our modeling of classical conditioning one has another mechanism for triggering sustained oscillations in nets. Basically, one is confronted with two problems. If one already has a high sensitivity net one merely has to trigger once enough neurons near threshold for the oscillation to occur. From that point on, reverberations will go on continuously until the net becomes inhibited by some sort of central mechanism.

Secondly, one must concern oneself with the mode of formation of high sensitivity nets in the association cortex. One possible mechanism is that proposed by Csérmely (1968). There must exist other mechanisms. Reverberating netlets in the association cortex itself must somehow be able via their connections with other sets of neurons in the association cortex to form these neurons into cell of assemblies. Further elaboration of the above ideas is warranted by the following considerations: Burns and Smith (1962) studying specificity of behaviour of single neurons with microelectrodes found that every cell which was monitored could be demonstrated to alter its firing pattern as a consequence of any arbitrary stimulus which was presented. Similar results were obtained by Gerstein (1969) who concluded that the firing pattern of a single neuron could not carry enough information to account for the discriminative and integrative activity of a given animal; rather these functions must be mediated by simultaneous activity in ensembles of neurons, organized into functional groups with boundaries and relationships that shift from task to task. Cells responsive to inputs from multiple sensory modalities have been reported to range from 18 to 92 percent in various brain regions. (O'Brien and Fox, 1969, Morell, 1969). In the words of John (1972): "...Those polysensory cells which originally responded to both the CS and the US (are the ones) that seem most likely to alter response during conditioning".

More recently Bechtereva (1978) has cited evidence to the effect that a given neuronal pool may partake in more than one cell assembly. This consideration will obviously have the effect of increasing the memory capacity of given aggregate of neurons. (Shaw, 1978).

At the level of psychological theories of memory, as we saw earlier on, one comes across theories that postulate an organization of memory whose neurological implications are in agreement with the above findings. Norman (1968), Meyer (1970), Schneider and Shiffrin (1977), in considering the problem of referencing incoming information by already existing information in the memory store, invoke the existence of structures (categories) that recall to one's mind the so-called logical Venn diagrams.

On the basis of the above considerations we propose the following hypothesis concerning the topology of memory in the association cortex: One can consider memory as being a collection of high excitability nets with a definite structure. We could call these nets memory units and denote one of these units by the subscripted variable M_i or M_j . Using the notation of set theory we could represent the space of all memory by the set S , which may be taken as the union of all the M_i 's. In other words, $S = \bigcup_i (M_i)$. In general $M_i \cap M_j \neq 0$. In other words, the memory units must have a non-vanishing Boolean overlap. This requirement seems to be warranted by the ability of our memory system to form associations. The memory units, or combinations thereof, are capable of various modes of sustained oscillation. In their quiescent (or non-oscillating mode) state the M_i 's constitute the permanent memory store (PMS). While oscillating they constitute active memory or short term memory (STM). The oscillating mode corresponds to the readout process of John (1972). The idea that both PMS and STM are modes of the same structure is due to Norman (1968). The memory units constituting the memory space S can only be nets of the types A, B, or C_1 . At the early stages of development of this system, one possible mechanism of its formation is that postulated by Csermely (1968). However, once these

excitable nets have been formed in the association cortex, they can now act as sources of steady input to the ensembles to which their axonal projections feed. Hebb's mechanism can now again be invoked to explain further growth of this structure. Cross-referencing incoming new information, say X , means establishment of the various $M_i \wedge X$. In general, $X = \bigcup_i (X \wedge M_i)$ and, therefore, to the extent that X overlaps with various units M_i of the memory space, triggering of X leads to excitation of the M_i 's. If input is at the threshold for any given M_i , then this particular M_i will oscillate.

As we saw X may be represented by the set of oscillations that become generated due to the overlap of $X \wedge M_i$. Little and Shaw (1975) have also identified access to a given memory by the measure of the overlap a given stimulus has with established structures.

In spite of its abstractness our model has certain appealing features:

- (1) It allows for both local and non-local memory storage. Our memory units M_i which are basically assemblies can be either diffuse or local assemblies. Braitenberg (1978) analyzing the known response characteristic of various areas of the cortex hypothesizes the existence of three types of aggregations of neurons. Thus his A system corresponding to a diffuse assembly would be the equivalent of a Hebb assembly representing some term such as say "bird". His B system is required to represent aspects or properties of a given object. B systems are said to be localized cell assemblies. Thus, line and angle features would be represented by such topographically localized populations. Finally, C systems would represent sequences of interconnected assemblies. These sequences when activated would give coherence to neural activity. Basically, Braitenberg (1978) categorizes

the information that must be learned into two types: On the one hand, we have type I signals that occur together and thus represent properties or aspects of things or situations. The succession in time of such signals (or aggregations thereof) are called type II signals. As he very aptly puts it "Abstractly, I are the terms and II their (e.g. causal) relations, or I is the dictionary and II is the syntax.

Information of type I must be represented in the brain by the connections of the cells within a cell assembly, that of type II by the connections between one cell assembly and the next". This last remark of Braitenberg brings us again to considerations having to do with control of brain function. They seem to imply that control is a structuro-dynamic property of duly connected neuronal pools. In our work one might correlate control with the highly damped character of the net. In spite of its abstractness and hence its lack of analytical power in the details we believe that further work along our own model could be fruitful even from the point of view of control mechanisms.

2. A single memory mechanism is postulated for all phenomena of memory. Craik and Lockhart's hypotheses mentioned in earlier parts of this work can be fully covered via our overlap mechanism. Obviously his claim that elaboration and rich encoding of a stimulus make it more likely to be recalled than some other stimulus which is not as elaborated simply means that access to an elaborated signal can be attained more easily due to its multiple overlaps with other established memories. Further, keeping processing of information at a fixed level (short term memory) seems to imply that only a given level in the hierarchy of assemblies is activated at this time.

3. Robustness of the network is guaranteed by our demand that activity rather than specific neuronal configurations within the netlet are the important parameters to consider. This gives our representation degeneracy properties, (in the sense of quantum mechanics) and hence, statistical reliability of firing patterns.

4. The memory domains I propose have been given a topological characterization. They are fuzzy, non metric and overlapping topological spaces. The consequences of this overlap property are quite drastic. This can be seen in Figures 46, 47 and 48. Figures 46 and 47 show the results of pattern completion for both excitatory and inhibitory nets. The parameter ρ stands for the fraction of the stimulus that is used to produce the cortical state. They may be interpreted as simulations of the Notion and Stark experiments. Similarly Figure 48 illustrates the ability of the overlap which is near threshold to set the net into sustained oscillation. A series of overlapping nets of this type can be seen to correlate with Hebb's phase sequence mentioned earlier on.

5. The necessity and ability of the CNS to form associations is also seen to depend on the overlapping domains. We have correlated these psychological variables to neurophysiological variables i.e., Szentagothai's demand that only a finite number of synaptic stations separate any two neurons in the association cortex. Finally we have correlated this neurophysiological finding to the overlap and fuzziness property of the topology of memory domain.

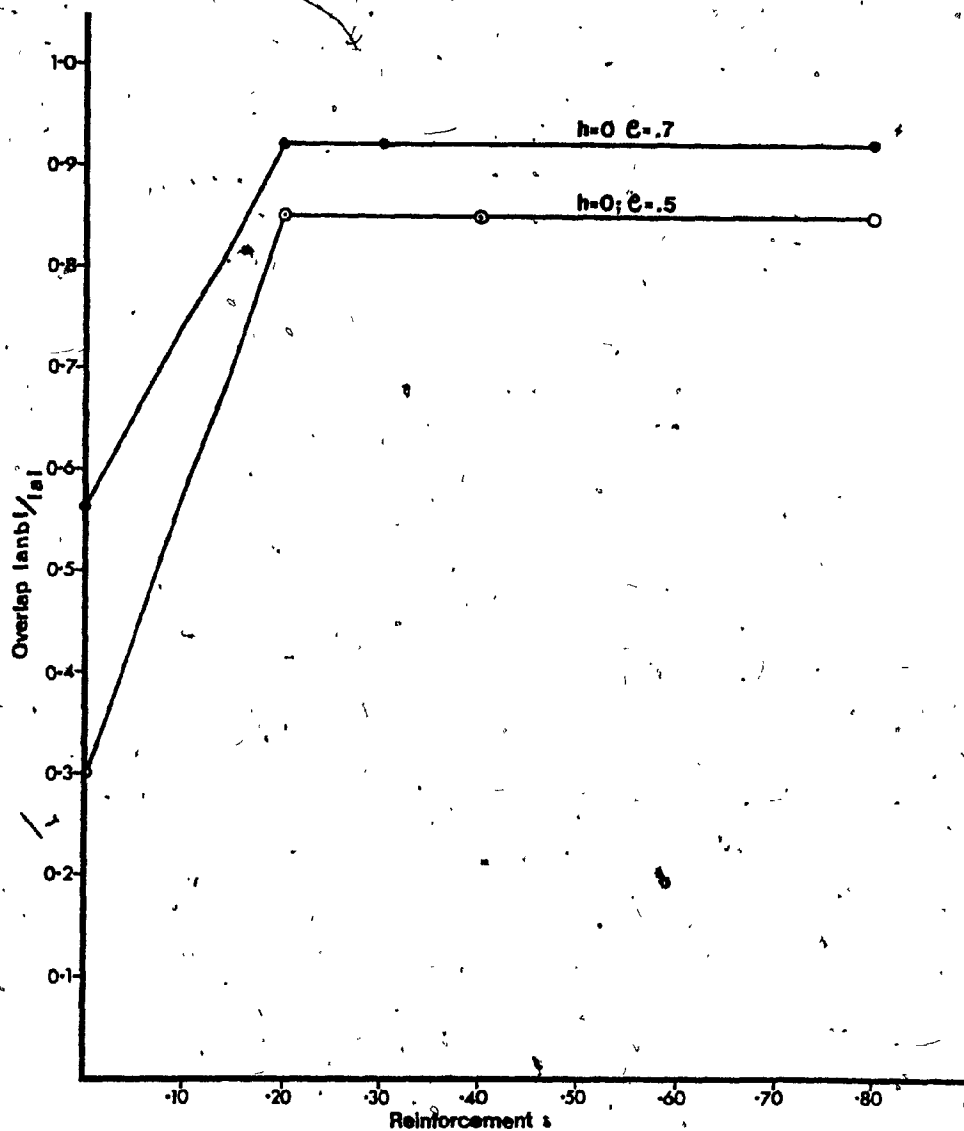


Figure 46. Pattern Completion

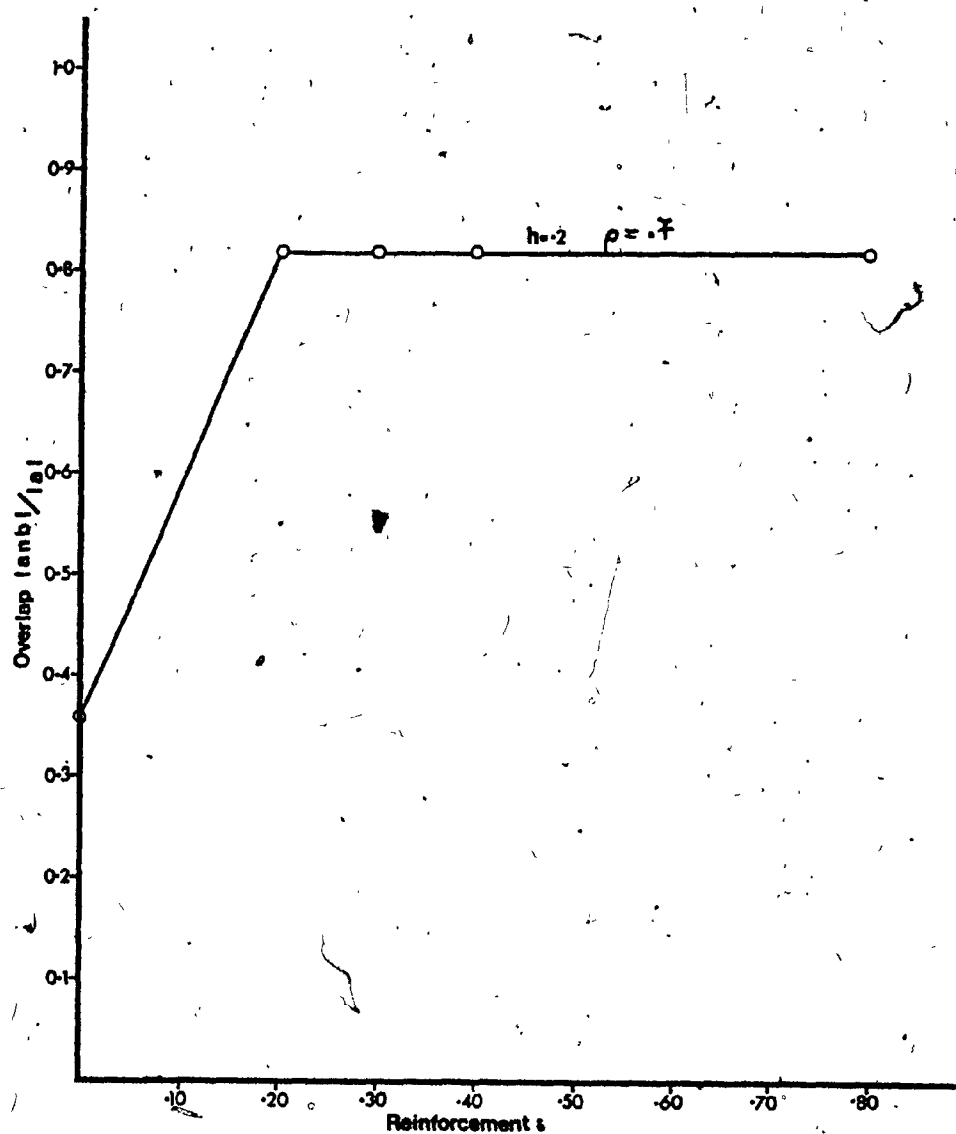


Figure 47. Pattern Completion with Inhibition

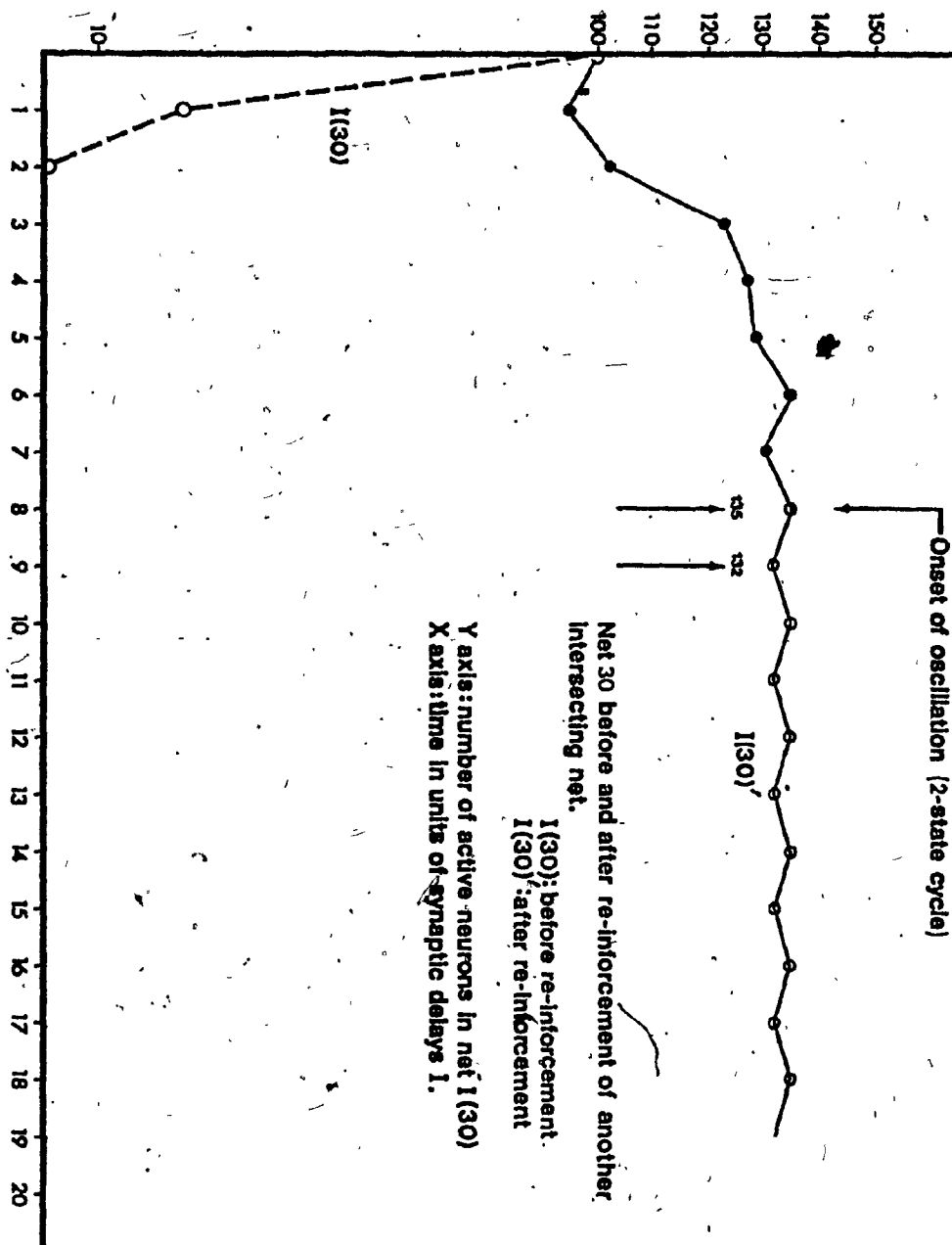


Figure 48. Firing of an Overlap near Threshold for Sustained Oscillation

REFERENCES

- Allanson, J.T., 1956, Some Properties of a Randomly Connected Neural Network, In Information Theory (Third London Symposium), C. Cherry, (ed.), Butterworth and Co., Ltd., London.
- Anasuma, H. and Sakata, H., 1967, J. Neurophysiol. 30:35.
- Anninos, P.A., 1969, Dynamics and Function of Neural Structures, Doctoral Dissertation, Syracuse University, Syracuse, New York, (unpublished).
- Anninos, P.A., 1972a, Kybernetik, 10:165.
- Anninos, P.A., 1972b, Kybernetik, 11:5.
- Anninos, P.A., Beek, B., Csermely, J.J., Harth, B.M. and Pertile, G. 1970, J. Theor. Biol. 26:121.
- Anninos, P.A. and Ellul, R., 1971, Biophys. J. 11:242.
- Anninos, P.A. and Ellul, 1974, Biophys. J. 14:8.
- Anninos, P.A., 1975, Progress in Neurob., 4:57.
- Arvanitaki, A., 1942, J. Neurophysiol., 5:89.
- Asratian, E.A., 1960, Compensatory Adaptation and Reflex Activity in the Brain. Pergamon Press, Oxford.
- Atkinson, R.C. and Shiffrin, R.M., 1968, Human Memory: A proposed system and its control processes. In the Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 2. Academic Press, New York.
- Bechtereva, V.P., 1978. The Neurophysiological Aspects of Human Activity. Oxford University Press, New York.
- Bennet, B.L. and Rosenzweig, M.R., editors, 1976. Neural Mechanisms of Learning and Memory. MIT Press, Boston.
- Beurle, R.L., 1956, Trans. Roy. Phil. Soc., London, Ser. B. 240:55.
- Bourdois, P.S., McCandles, D.L. and MacIntosh, F.C., 1975. Can. J. of Physiol. and Pharmacol. 53:155.
- Braitenberg, V., 1968. Cortical Architectonics: General and Areal. In Architectonics of the Cerebral Cortex. Brazier, M.A.B. and Petsche H., editors. Raven Press, New York.
- Browning, M., 1979. Science. 203:60.
- Burns, B.D. and Smith, G.K., 1962. J. Physiol. London 164:238.

- Colon, B.J., 1968. The Study of Brain Functions: Some Approaches and Techniques. In *The Mind; Biological Approaches to its Functions*. W.C. Corning and M. Balaban, editors, Interscience, Wiley, New York.
- Colonnier, M.L., 1966. The Structural Design of the Neocortex. In *Brain and Conscious Experience* J.C. Eccles, (ed.) Springer-Verlag, New York.
- Craik, F.I.M. and Masani, P.A., 1969. *British J. of Psych.* 60:315.
- Craik, F.I.M. and Lockhart, R.S., 1972. *J. of Verbal Learning and Verbal Behaviour*, 11:671.
- Craik, F.I.M. and Tulving, E., 1975. *J. of Exper. Psych. General*, 104 (3):28.
- Csermely, T.J., 1968. Analysis of Neural Net Dynamics by Computer Simulation. Doctoral Dissertation, Syracuse University, Syracuse, New York (unpublished).
- Eccles, J.C., 1964. *The Physiology of Synapses*. Springer-Verlag, New York.
- Eccles, J.C., 1966. Conscious Experience and Memory. In *Brain and Conscious Experience*. J.C. Eccles, (ed.). Springer-Verlag, New York.
- Eccles, J.C., 1972. Possible Synaptic Mechanisms Subservicing Learning. In *Brain and Human Behaviour*. J.C. Eccles and A.G. Karczmar, (eds.). Springer-Verlag, New York.
- Eccles, J.C., 1973. *The Understanding of the Brain*. McGraw Hill, New York.
- Eccles, J.C. and Popper, K., 1978. *The Self and its Brain*. Springer-Verlag, New York.
- Farley, B.G. and Clark, W.A., 1961. Activity in Networks of Neuron-Like Elements. In *Information Theory (Fourth London Symposium)*. C. Cherry, (ed.). Butterworth and Co. Ltd., London.
- Ferrando, F.R., 1962. *Persp. Biol. Med.* 6:296.
- Freeman, W.F., 1972. Waves, Pulses and the Theory of Neural Masses. In *Progress in Theoretical Biology*, Vol. 1. John Wiley and Sons, New York.
- Gerstein, G.L., 1969. *Int. Biophys. Congress 3rd*. Cambridge, Mass.
- Gottlieb, G. (ed.) 1976. *Neural and Behavioural Specificity*. Academic Press, New York.

- Griffith, J.S., 1963. Bull. Math. Biophys. 25:111.
- Griffith, J.S., 1965. Bull. Math. Biophys. 27:187.
- Guyton, A.C., 1976. Organ Physiology: Structure and Function of the Nervous System. W.B. Saunders, Philadelphia.
- Harth, E.M., 1966. Brain Models and Thought Processes. In Automatic Theory, E.R. Caianiello, (ed.). Academic Press, New York.
- Harth, E.M. and Edgar, S.L., 1967. Biophys. J. 7:689.
- Harth, E.M., Csermely, T.J., Beek, B. and Lindsay, R.D., 1969. J. Theor. Biol. 26:93.
- Hebb, D.O., 1949. The Organization of Behaviour. John Wiley and Sons, New York.
- Hebb, D.O., 1959. A Neuropsychological Theory. In Psychology a Study of Science. S. Köch (ed.), Vol. 1. McGraw Hill, New York.
- Hebb, D.O., 1963. American Psychologist 18:16.
- Hebb, D.O., 1968. Psych. Rev. 75:466.
- Hebb, D.O., Lambert, W.E. and Tucker, G.R., 1971. Modern Language J. 55:212.
- Hebb, D.O., 1972. A Textbook of Psychology. Saunders, Philadelphia.
- Hoopen, M., 1965. Cybernetics of Neural Processes. E.R. Caianiello, (ed.). CNDR, Rome.
- Horn, G. and Hinde, R.A. (eds.), 1970. Short Term Processes in Neural Activity and Behaviour. Cambridge University Press, Cambridge.
- Hubel, D.H. and Wiesel, T.N., 1961. J. Physiol. London. 155:385.
- Hubel, D.H. and Wiesel, T.N., 1962. J. Physiol. 160:106.
- Hubel, D.H. and Wiesel, T.N., 1963. J. Physiol. 165:559.
- Hubel, D.H. and Wiesel, T.N., 1965. J. Physiol. 28:229.
- Hubel, D.H. and Wiesel, T.N., 1968. J. Physiol. (London) 195:215.
- Jeannerod, M., Gerin, P. and Pernier, J., 1968. Vision Research. 8:81.
- John, E.R., 1967. Mechanisms of Memory. Academic Press, New York.
- John, E.R., 1972. Science. 177:850.

- Kandel, E.R., 1970. Nerve Cells and Behaviour. Scientific American, San Francisco.
- Lashley, K.S., 1950. In Search of the Engram. Symp. Soc. Exp. Biol. 4:454.
- Lashley, K.S. 1951. The Problem of Serial Order in Behaviour. In Cerebral Mechanisms in Behaviour. Ch. A. Jeffreys, (ed.), John Wiley, New York.
- Lashley, K.S., 1958. Proc. Ass. Res. Ment. Disor. 36:1.
- Lindsay, P.H. and Norman, D.A., 1972. Human Information Processing. Academic Press, New York.
- Little, W.A. and Shaw, G.L., 1975. Behav. Biol. 14:115.
- London, I.D., 1960. Amer. J. Psych. 73:478.
- Lorente de No, R., 1938a. J. Neurophys. 1:195.
- Lorente de No, R., 1938b. J. Neurophys. 1:207.
- Lund, R.D., 1978. Development and Plasticity in the Brain. Oxford University Press, New York.
- McCulloch, W.S. and Pitts, W. H., 1943. Bull. Math. Biophys. 5:115.
- MacGregor, R. and Lewis, E., 1977. Neural Modeling. Plenum Press, New York.
- Meyer, P.M., 1970. Cogn: Psych. 1:42.
- Milner, P.M., 1957. Psych. Rev. 64:242.
- Milner, P.M., 1970. Physiological Psychology. Holt, Rinehart and Winston, New York.
- Morell, F., 1967. Electrical Signs of Sensory Coding. In The Neurosciences; A Study Program. Quarton G.C., T.C. Melnechuk and F.O. Smitt, (eds.), Rockefeller University Press, New York.
- Mountcastle, V.B., 1957. J. Neurophys. 20:408.
- Norman, D.A., 1968. Psych. Rev. 75:522.
- Norman, D.A., 1970 (ed.), Models of Human Memory. Academic Press, New York.
- Norman, D.A., 1976. Memory and Attention. John Wiley and Sons, New York.
- Noton, D. and Stark, L., 1971. Science. 171:308.

- O'Brien, J. and Fox, S.S., 1969. *J. Neurophysiol.* 22:3.
- Papadopoulos, M.S., Anninos, P.A. and Zenone, S., 1979. *J. Theo. Biology.* 80:505.
- Penfield, W. and Rasmussen, T., 1955. *The Cerebral Cortex of Man.* The MacMillan Press, New York.
- Peterson, L.R. and Peterson, M.J., 1959. *J. Exp. Psych.* 58:193.
- Phillips, W.A., 1974. *Perception and Psychophysics.* 16:283.
- Popper, K.R., 1959. *The Logic of Scientific Discovery.* Hutchinson, London.
- Pribram, K., 1966. Some Dimensions of Remembering: Steps Toward a Neuropsychological Model of Memory. In *Macromolecules and Behaviour*. John Calto, (ed.). Appleton-Century Crofts, New York.
- Pribram, K., 1969. *Scient. American.* 1:220.
- Pritchard, R., 1971. Stabilized Images on the Retina. In *Contemporary Psychology*. R.C. Atkinson, (ed.). W. Freeman and Co., San Francisco.
- Quintilianus, M.F. *The Institutio Oratoria* (M. E. Butler, Trans.) Putnam, 1921. New York.
- Rapoport, A., 1952. *Bull. Math. Biophys.* 14:35.
- Reynolds, A.G. and Flagg, P.W., 1977. *Cognitive Psychology*. Winthrop Publishers, Cambridge, Mass.
- Rochester, N., Holland, J.H., Halbt, L.H. and Duda, W.L., 1956. *IRE Trans. Inform. Theory.* 2:80.
- Rodan, G.A., 1978. *Science.* 199:
- Ruckin, D. and John, E.R., 1966. *Science.* 153:209.
- Rugh, T.C. and Patton, H.D., 1965. *Physiology and Biophysics.* W.B. Saunders, Philadelphia.
- Schmidt, R.F. (ed.), 1978. *Fundamentals of Neurophysiology.* Springer-Verlag, New York.
- Schneider, W. and Shiffrin, R.M., 1977. *Psych. Rev.* 84:1.
- Shaw, G.L., 1978. *Brain Res. Bull.* 3(2):107.
- Shiffrin, R.M. and Atkinson, R.C., 1969. *Psych. Rev.* 56:179.
- Shiffrin, R.M., 1970. Memory Search. In *Models of Human Memory*. D.A. (ed.). Academic Press, New York.

- Shiffrin, R.M. and Schneider, W., 1977. Psych. Rev. 84:127.
- Sholl, C.A., 1956. The Organization of the Cerebral Cortex. Methuen, London.
- Simon, H.A. Science. 183:482.
- Simon, H.A., 1976. The Information-Storage System called 'Human Memory'. In Neural Mechanisms of Learning and Memory. M.R. Rosenzweig and E.L. Bennet, (eds.). MIT Press, Cambridge.
- Smit, J., 1968. Doctoral Dissertation. University of Amsterdam, Holland. (cited by T.J. Csermely).
- Smith, D.R. and Davidson, C.H., 1962. U. Assoc. Comput. Math. 9:268.
- Sperling, G., 1960. Psychol. Monographs. 74.
- Sperling, G., 1963. Human Factor. 5:19.
- Stevens, C.F., 1966. Neurophysiology: A Primer. John Wiley, New York.
- Stoney, S.D. Jr., Thompson, W.D. and Anasuma, H., 1968. J. Neurophysiol. 31:659.
- Szentagothai, J., 1969. Architecture of the Cerebral Cortex. In Basic Mechanisms of the Epilepsies. H.H. Jasper, A.A. Ward and A. Pope (eds.). Little, Brown and Co. Ltd., Boston.
- Szentagothai, J., 1972. The Basic Neuronal Circuit of the Neocortex. In Synchronization of EEG Activity in Epilepsies. A. Petsche and M.A.B. Brazier (eds.). Springer-Verlag, Berlin.
- Szentagothai, J., 1973. Synaptology of the Visual Cortex. In Visual Perception in Neurophysiology. R. Jung (ed.) In Handbook of Sensory Physiology, Vol. VII/3A. Springer-Verlag, Berlin.
- Szentagothai, J., 1971. Memory Functions and the Structural Organization of the Brain. In Biology of Memory. G. Adam, (ed.) Plenum Press, Budapest.
- Szentagothai, J., 1974. From the last skirmishes around the Neuron Theory to the Functional Anatomy of the Neuron Network. In The Neurosciences: Paths of Discovery. F.G. Worden, J.P. Swazey and G. Adelman, (eds.). MIT Press, Cambridge.
- Szentagothai, J., 1975. Brain Research. 95:475.
- Szentagothai, J., and Arbib, M.A., 1975. Conceptual Models of Neural Organization. MIT Press, Cambridge.
- Trucco, E., 1952. Bull. Math. Biophys. 14:365.

Underwood, B.J., 1969. Psych. Rev. 76:559.

Verzeano, M., Laufer, M., Spear, P. and McDonald, S., 1970. The Activity of Neuronal Networks in the Thalamus of the Monkey. In Biology of Memory. K. Pribram and S. Broadbent (eds.). Academic Press, New York.

Wilson, H.R. and Cowan, J.D., 1972. Biophys. J. 12:1.

Yarbus, A.L., 1967. Eye Movements and Vision. Plenum Press, New York.