A HIERARCHICAL MODEL OF VISUAL PERCEPTION

by

JAMES L. EILBERT

A thesis submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

DEPARTMENT OF STATISTICS, BIOMATHMATICS PROGRAM

RALEIGH

1980

APPROVED BY:

Jack W. Shernoff

[Signature]

H. van der Vaart

[Signature]

CHAIRMAN OF ADVISORY COMMITTEE
ABSTRACT

EILBERT, JAMES LAWRENCE. A Hierarchical Model of Visual Perception. (Under the direction of HARVEY J. GOLD)

A hierarchical neural model of visual recognition has been developed. The recognition process has three sources of input: optical signals entering the retina, stored past experience, and the current brain state. A neural model of recognition must characterize the signals generated by each of these inputs and explain how they interact. Physiological evidence indicates that the simplest neural model that can account for both the local processing and the global spread of neural signals is a four cell type model (composed of small excitatory interneurons, small inhibitory interneurons, relay cells, and large inhibitory cells) with variable parameters.

The model was constructed by combining three recent observations, concerning the connections among these four types of cells, with previous concepts of neural modeling. First, midbrain neurons have diffuse axon terminations that can fill a large portion of a whole nucleus, and their neurotransmitters can cause long lasting changes in the response of their targets. Variable parameters are included to account for the effects of these neurons. Second, the axon terminations of relay cells are distributed through a cylindrical column 300 microns in diameter, at least in the cerebral cortex. Third, within one cortical column, only the two types of interneurons seem to have reciprocal connections. These last two observations establish a
hierarchical decomposition of the model, with one level describing the single interneuron interactions within a cortical column, and the other level describing the interactions of columns within the visual system, mediated by the relay cells. The two levels are connected by the transfer of output from the interneuron network onto relay cells of the same column. The processing within the interneuron network model was shown capable of separating visual qualities due to the axonal field asymmetry of certain inhibitory neurons.

The author has proposed that self-sustaining activity in each level of the hierarchy plays a key role in cognitive processes. Self-sustaining activity demonstrated in the interneuron network has properties that could allow the type of storage of particular visual qualities found in iconic memory. Self-sustaining activity among a set of columns, designated a turned-on active circuit, has signal processing properties resembling those in the long term storage of information in memory. For a particular set of parameters, each column participating in an active circuit must receive a strong, synchronized input from some minimal number of other columns in the active circuit. Synchronization requires that signals circulating through an active circuit have certain optical-like properties, which allow these signals to be considered neural waves. Neural waves can carry information about input, past experience and brain state; thus they can characterize the inputs to the recognition process. It is also proposed that a progressive series of synaptic modification allows the connection strengths between columns to be modified, thus changing the columns participating in an active circuit or combining several active circuits into one.
The definition of neural waves which emerged from the four cell type model presents a method for developing a mathematical basis for the holographic analogy of neural recording and retrieval advocated by Pribram and others. Neural waves and modifiable active circuits were used to demonstrate a morphism between a holographic and a neural scene analysis system. Using the analogy as a heuristic guide, a preliminary model of neural recording which assigned the performance of particular processes to specific nuclei was constructed. However, differences in the number of steps required for recording and retrieval, and the way in which stored information is distributed, make it unlikely for the mathematics governing holography to apply to the neural system.

A survey of psychological studies led to a description of the recognition experience as an extraction of basic visual qualities, followed by a separation of figure and ground, and finally a slow integration of basic units into synthesized scenes. The turn-on of an elemental active circuits by a visual input, followed by their integration into a complete active circuit, possesses the major characteristics of recognition. Thus, it provides a model for the neural processes by which recognition is achieved.
A HIERARCHICAL MODEL OF VISUAL PERCEPTION

by

JAMES L. EILBERT

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy
DEPARTMENT OF STATISTICS, BIOMATHEMATICS PROGRAM

RALEIGH
1980

APPROVED BY:

[Signatures]

[Signature]
CHAIRMAN OF ADVISORY COMMITTEE
James Lawrence Eilbert was born in Pittsburgh, Pennsylvania on April 12, 1950. His family moved to Washington, D.C., Arkansas, West Germany, and back to Pittsburgh, before settling in Yonkers, N.Y. He graduated Roosevelt High School, Yonkers in 1968. He received a Bachelor of Science degree with a major in Physics from the State University of New York at Stony Brook in 1973, and a Master of Science with a major in Applied Mathematics from New York University in 1975.

In September 1975 the author enrolled in the Biomathematics Program at North Carolina State University. He chose neural models of perception as his area of research. He was supported in his studies at North Carolina State University by a NIH Fellowship, and later by a Teaching Assistantship from the Statistics Department.

The author is married to Loren R. Lasky. She graduated from the Colorado School of Mines with a Master of Science degree in Geology in 1980.
ACKNOWLEDGEMENTS

I would like to express my appreciation to Professor Harvey J. Gold, the chairman of my advisory committee, for his guidance, advice and patience during the course of this study. I would also like to thank the other members of my committee, Professors H. R. van der Vaart, Jack W. Silverstein, and Thomas E. LeVere for constructive criticism and assistance.

I would also like to extend special thanks to my father, Dr. Leo R. Eilbert, and my wife, Loren, for their help and sacrifices in the preparation of this dissertation.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF EQUATIONS</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF MATHEMATICAL SYMBOLS</td>
<td>xiii</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 The recognition process</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Neural models</td>
<td>6</td>
</tr>
<tr>
<td>1.3 The approach to modeling</td>
<td>12</td>
</tr>
<tr>
<td>2. A GENERAL MODEL OF THE VISUAL SYSTEM</td>
<td>18</td>
</tr>
<tr>
<td>2.1 A model of neurons within a column</td>
<td>28</td>
</tr>
<tr>
<td>2.1.1 Densities, shapes, and connections of the cell types</td>
<td>29</td>
</tr>
<tr>
<td>2.1.2 The model</td>
<td>38</td>
</tr>
<tr>
<td>2.1.3 Data supporting the belief in parameter changes</td>
<td>51</td>
</tr>
<tr>
<td>2.1.4 Decomposing equation 2.1 into a hierarchical model</td>
<td>54</td>
</tr>
<tr>
<td>2.2 The interneuron network</td>
<td>56</td>
</tr>
<tr>
<td>2.2.1 Deriving the time coarse grained model</td>
<td>56</td>
</tr>
<tr>
<td>2.2.2 Dynamics of the within column interneuron network</td>
<td>61</td>
</tr>
<tr>
<td>2.2.3 Self-sustaining activity and related states</td>
<td>64</td>
</tr>
<tr>
<td>2.2.4 The formation of quality differences among cells in a column</td>
<td>71</td>
</tr>
<tr>
<td>2.3 The model of columns within a pathway</td>
<td>75</td>
</tr>
<tr>
<td>2.3.1 Relay cell axon distributions</td>
<td>76</td>
</tr>
<tr>
<td>2.3.2 The basic connections of the</td>
<td></td>
</tr>
</tbody>
</table>
relay system model ........................................ 79

2.3.3 Data based restrictions on
relay system processing ................................... 83

2.3.4 Changing the sensitivity of relay
and large inhibitory cells ................................. 89

2.3.5 Modifiable synapses .................................. 95

2.4 Neural waves ............................................ 99

2.4.1 Properties of optical waves ......................... 103

2.4.2 Properties of neural waves ......................... 106

2.4.3 The transformation of optical waves ............... 112

2.4.4 The transformation of neural waves ............... 114

3. MEMORY .................................................... 124

3.1 Introduction ............................................. 124

3.2 Iconic memory .......................................... 125

3.3 Short term memory ...................................... 128

3.4 A model of short term memory ......................... 132

3.5 Long term memory ...................................... 134

3.6 A neural model of long term memory ................. 141

3.7 The effects of sensory deprivation ................... 144

4. AN ANALOGY BETWEEN HOLOGRAPHY AND
NEURAL SCENE ANALYSIS ................................ 147

4.1 A morphism of the I/O spaces holographic
and a neural scene analysis system .................... 147

4.1.1 Introduction ........................................ 147

4.1.2 The basic I/O pairs in a holographic system .... 154

4.1.3 The I/O spaces of the neural
scene analysis system ..................................... 156

4.1.4 The morphism ....................................... 159

4.1.5 Comments about the spatial distributions
associated with the g and k maps ....................... 162
4.1.6 Discussion and conclusions ......................... 164

4.2 The state transition functions of the
    holographic and neural scene analysis systems ........ 166
    4.2.1 Introduction ................................... 166
    4.2.2 The state transition and output
        functions of holography ......................... 170
    4.2.3 The state transition and output functions
        for the visual scene analysis system ............ 175
    4.2.4 Conclusions .................................... 184

5. THE ROLE OF ATTENTION IN ACHIEVING PERCEPTION ........ 187
    5.1 Cases of perceptual processing ....................... 190
       5.1.1 Complex, simple, and very familiar objects .... 190
       5.1.2 Cases with no specific expectations
            vs. no specific expectations ................... 192
    5.2 Previous models of attention in the
        extraction of sensory information ................. 198
       5.2.1 Kahneman's model ............................ 199
    5.3 A new model of attention combining
        interaction and synthesis .......................... 205
       5.3.1 Basic components of the model ................. 205
       5.3.2 Templates and attention channels ............. 206
       5.3.3 Synthesized scenes (events) and the
            interaction of attention channels ............. 211
    5.4 The role of specific brain regions
        in attention ...................................... 217
       5.4.1 The frontal effects on the
            dimensions of attention ....................... 218
       5.4.2 The posterior effects on the
            dimensions of attention ....................... 221
    5.5 A neural model of attention ........................ 223

6. DISCUSSION AND CONCLUSIONS ........................... 226
    6.1 Assumptions in the four cell type model ............ 226
6.2 Models of recognition, attention, and memory based on the active circuit concept 229

6.3 Experimental testing 233

BIBLIOGRAPHY 237
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The form of the connection matrices</td>
<td>46</td>
</tr>
<tr>
<td>2.2</td>
<td>The form of the matrix $A_{21}$</td>
<td>48</td>
</tr>
<tr>
<td>2.3</td>
<td>The form of the matrix $A_{34}$</td>
<td>49</td>
</tr>
<tr>
<td>5.1</td>
<td>Characteristics of attention for different combinations of scenes and expectations</td>
<td>191</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.1</td>
<td>Neuron shapes</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Types of synapses</td>
<td>22</td>
</tr>
<tr>
<td>2.3</td>
<td>The myelin sheath and nerve conduction</td>
<td>23</td>
</tr>
<tr>
<td>2.4</td>
<td>Excitatory connections in one area 17 column</td>
<td>24</td>
</tr>
<tr>
<td>2.5</td>
<td>Inhibitory connections in one area 17 column</td>
<td>26</td>
</tr>
<tr>
<td>2.6</td>
<td>Axonal arrangements in area 17</td>
<td>30</td>
</tr>
<tr>
<td>2.7</td>
<td>The cerebral cortex</td>
<td>31</td>
</tr>
<tr>
<td>2.8</td>
<td>The sigmoidal function</td>
<td>59</td>
</tr>
<tr>
<td>2.9</td>
<td>The area of axon influence</td>
<td>66</td>
</tr>
<tr>
<td>2.10</td>
<td>Hysteresis</td>
<td>69</td>
</tr>
<tr>
<td>2.11</td>
<td>The active transient mode</td>
<td>69</td>
</tr>
<tr>
<td>2.12</td>
<td>A steady state showing edge enhancement</td>
<td>69</td>
</tr>
<tr>
<td>2.13</td>
<td>A localized limit cycle</td>
<td>69</td>
</tr>
<tr>
<td>2.14</td>
<td>Different levels of pyramidal cell inhibition</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>due to the asymmetry of large basket cells</td>
<td></td>
</tr>
<tr>
<td>2.15</td>
<td>An active circuit composed of seven columns and requiring four connections per column</td>
<td>87</td>
</tr>
<tr>
<td>2.16</td>
<td>An Airy Pattern</td>
<td>116</td>
</tr>
<tr>
<td>4.1</td>
<td>Two algebraically equivalent systems</td>
<td>150</td>
</tr>
<tr>
<td>4.2</td>
<td>I/O function observation morphism</td>
<td>150</td>
</tr>
<tr>
<td>4.3</td>
<td>A holographic system</td>
<td>155</td>
</tr>
<tr>
<td>4.4</td>
<td>A block diagram of the visual system</td>
<td>157</td>
</tr>
<tr>
<td>4.5</td>
<td>The I/O systems morphism</td>
<td>168</td>
</tr>
</tbody>
</table>
5.1 Supply and demand of attention .................................. 189
5.2 Kahneman's model of perception and attention ................. 199
5.3 A new model of attention and perception .......................... 206
LIST OF EQUATIONS

2.1 A within column model for area 17 ................. 40
2.2 A continuous approximation for the
      interneuron network ........................... 57
2.3 The conversion of the axon hillock potential
      to a firing rate ............................... 58
2.4 The time coarse grained interneuron equations .... 60
2.5 Wilson and Cowan's interneuron equations ........ 60
2.6 A between column model for the visual system .... 82
2.7 The optical wave definition ........................ 105
2.8 The neural wave definition ........................ 111
2.9 The transformation of optical waves ............... 114
4.1 Recording a hologram ............................ 172
4.2 Reconstructing the image of the original object ... 174
LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMP</td>
<td>adenosine monophosphate</td>
</tr>
<tr>
<td>AP</td>
<td>action potentials</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>CNV</td>
<td>contingent negative variation</td>
</tr>
<tr>
<td>EEG</td>
<td>electroencephalogram</td>
</tr>
<tr>
<td>EP</td>
<td>evoked potential (EEG)</td>
</tr>
<tr>
<td>EPSP</td>
<td>excitatory post-synaptic potential (microelectrode)</td>
</tr>
<tr>
<td>IPSP</td>
<td>inhibitory post-synaptic potential (microelectrode)</td>
</tr>
<tr>
<td>IT</td>
<td>inferotemporal cortex</td>
</tr>
<tr>
<td>l.b.c.</td>
<td>large basket cell</td>
</tr>
<tr>
<td>LGN</td>
<td>lateral geniculate nucleus of the thalamus</td>
</tr>
<tr>
<td>PRF</td>
<td>pontine reticular formation</td>
</tr>
<tr>
<td>SC</td>
<td>superior colliculus</td>
</tr>
<tr>
<td>s.e.i.</td>
<td>small excitatory interneuron</td>
</tr>
<tr>
<td>s.i.i.</td>
<td>small inhibitory interneuron</td>
</tr>
</tbody>
</table>
LIST OF MATHEMATICAL TERMS

t, \tau = time
x, y = position in a plane
z = position in the direction of signal propagation
\tau_0 = time delay for an AP in going from a s.e.i. axon hillock to a target axon hillock.
\tau_1 = time delay for an AP in going from a relay cell axon hillock to a target axon hillock.
\tau_2 = time delay for an AP in going from a s.i.i. or l.i.c. axon hillock to a target axon hillock.
\mu_{ex} = the decay rate of an EPSP
\mu_{in} = the decay rate of an IPSP

EQUATION 2.1

A_{11} = connection strength matrix from s.e.i. to s.e.i.
A_{12} = connection strength matrix from X-type LGN pyramidal cells to s.e.i.
A_{1X} = connection strength matrix from s.i.i. to s.e.i.
A_{21} = connection strength matrix from s.e.i. to s.i.i.
A_{22} = connection strength matrix from s.i.i. to s.i.i.
A_{31} = connection strength matrix from s.e.i. to pyramidal cells
A_{32} = connection strength matrix from s.i.i. to pyramidal cells
A_{34} = connection strength matrix from l.b.c. to pyramidal cells
A_{3Y} = connection strength matrix from Y-type LGN pyramidal cells to pyramidal cells
A_{33(cj)} = connection strength matrix from pyramidal cells in column j to pyramidal cells in column c
A_{41} = connection strength matrix from s.e.i. to l.b.c
A_{42} = connection strength matrix from s.i.i. to l.b.c
$A_{43}(c_j)$ = connection strength matrix from pyramidal cells in column $j$ to l.b.c. in column $c$

$I$ = midbrain and other input which causes long lasting response changes in target neurons

$\alpha(t)$ = the vector of s.e.i. axon hillock potentials in a column

$\beta(t)$ = the vector of s.i.i. axon hillock potentials in a column

$\gamma(t)$ = the vector of pyramidal cells axon hillock potentials in a column

$\delta(t)$ = the vector of l.b.c. axon hillock potentials in a column

$S(t)$ = the vector of s.e.i. momentary firing rates in a column

$B(t)$ = the vector of s.i.i. momentary firing rates in a column

$P(t)$ = the vector of pyramidal cells momentary firing rates in a column

$L(t)$ = the vector of l.b.c. momentary firing rates in a column

$T_1(t)$ = the vector of X-type LGN relay cell momentary firing rate

$T_2(t)$ = the vector of Y-type LGN relay cell momentary firing rate

$\alpha_0(t)$ = the no input vector of s.e.i. axon hillock potentials in a column

$\beta_0(t)$ = the no input vector of s.i.i. axon hillock potentials in a column

$\gamma_0(t)$ = the no input vector of pyramidal cells axon hillock potentials in a column

$\delta_0(t)$ = the no input vector of l.b.c. axon hillock potentials in a column

$S_0(t)$ = the no input vector of s.e.i. momentary firing rates in a column

$B_0(t)$ = the no input vector of s.i.i. momentary firing rates in a column

$P_0(t)$ = the no input vector of pyramidal cells momentary firing rates in a column

$L_0(t)$ = the no input vector of l.b.c. momentary firing rates in a column
$G_S$ = the transformation of the s.e.i. axon hillock potential vector into the s.e.i. momentary firing rate vector

$G_B$ = the transformation of the s.i.i. axon hillock potential vector into the s.i.i. momentary firing rate vector

$G_P$ = the transformation of the pyramidal cell axon hillock potential vector into the pyramidal cell momentary firing rate vector

$G_L$ = the transformation of the l.b.c. axon hillock potential vector into the l.b.c. momentary firing rate vector

EQUATION 2.2

$S(x,y,t)$ = a continuous approximation to the s.e.i. momentary firing rate vector in a column

$B(x,y,t)$ = a continuous approximation to the s.i.i. momentary firing rate vector in a column

$S_0(x,y,I)$ = a continuous approximation to the s.e.i. no-input momentary firing rate vector in a column

$B_0(x,y,I)$ = a continuous approximation to the s.i.i. no-input momentary firing rate vector in a column

$Z_S$ = input to the s.e.i. population from outside the column

$Z_B$ = input to the s.i.i. population from outside the column

$a_{ee}(I)$ = amplitude of the connection strength from the s.e.i. to the s.e.i. population

$a_{ei}(I)$ = amplitude of the connection strength from the s.i.i. to the s.e.i. population

$a_{ie}(I)$ = amplitude of the connection strength from the s.e.i. to the s.i.i. population

$a_{ii}(I)$ = amplitude of the connection strength from the s.i.i. to the s.i.i. population

$w_{ee}(I)$ = spatial distribution of the connection strength from the s.e.i. to the s.e.i. population

$w_{ei}(I)$ = spatial distribution of the connection strength from the s.i.i. to the s.e.i. population

$w_{ie}(I)$ = spatial distribution of the connection strength from the s.e.i. to the s.i.i. population
\( w_{ii}(I) \)  = spatial distribution of the connection strength from the s.i.i. to the s.i.i. population

**Equation 2.3**

\( O_S(I) \)  = position of the maximum slope of the function \( G_S \)

\( O_B(I) \)  = position of the maximum slope of the function \( G_B \)

\( v_S(I) \)  = four times the slope of \( G_S \) at \( O_S \)

**Equation 2.4**

\( S(x,y,t) \)  = time coarse-grained s.e.i. firing rate distribution

\( B(x,y,t) \)  = time coarse-grained s.i.i. firing rate distribution

**Holographic recording**

\( u_r(t) \) and \( x_r(t) \) = the reference wave at the source and in the hologram plane respectively

\( u_o(t) \) and \( x_o(t) \) = the object wave at the source and in the hologram plane respectively

\( (x_o + x_r)(t) \) = the combined object and reference optical waves in the hologram plane

\( y \) = the recording made of the interference of the object and reference waves in the hologram plane averaged over their temporal coherence

**Holographic reconstruction**

\( u_1(t) \) and \( x_1(t) \) = the reconstruction wave at its source and immediately in front of the hologram plane respectively

\( x_{12}(t) \) = the modified reconstruction wave directly behind the transmittance in the hologram plane

\( y_1(t) \) = the part of the diffracted reconstruction wave in the image plane that is identical up to a constant to the original object wave

**Holographic transform**

\( bp \) = the transformation of an optical wave in traveling between planes

\( dif \) = the transformation in an optical wave in passing through a transmittance

**Visual scene analysis recording**
(Note that since a neural recording may take place over many neural events, the subscript \( i \) indicates a specific neural event.)

\( u_{oi}'(t) \) and \( x_{oi}'(t) \) = the object neural wave at the retina and the inferotemporal cortex (IT) respectively

\( u_{ri}'(t) \) and \( x_{ri}'(t) \) = the reference neural wave where it is permanently stored (probably in some combination of the pontine reticular formation (PRF), the lateral posterior nucleus (LP), the pulvinar (PUL), areas 18 and 19 of the cortex, and the suprasylvian sulcus (SSS)) and from where it is projected to the entorhinal cortex (ENT) respectively

\( (u_{ri}'+x_{ri}') \) = the combined object and reference neural waves in the ENT

\( y' \) = the recording made of the interference of the object and reference neural waves in the nuclei where the permanent record is made

\( z_1' \) = the difference between the object and reference neural waves in the hippocampus averaged over one neural event

\( z_2' \) = the activity of the cells in the reticular formation that are postulated to put the visual system into the record mode (It is a 0 or 1 variable.)

\( x_{ei}'(t) \) = the combined object and reference waves in the region where the neural recording is made on the \( i \)th neural event from a scene

**Visual scene analysis reconstruction**

\( u_{1}'(t) \) and \( x_{1}'(t) \) = the reconstruction neural wave at its source and in the visual cortex respectively

\( x_{1,2}'(t) \) = the reconstruction neural wave after passing through the storage nuclei

\( y_{1}'(t) \) = the activity in the motor part of the central nervous system that results from the reconstruction neural wave if it is the same as the activity caused by the original object neural wave

**The visual scene analysis transformations**

(Note that all the scene analysis transforms in Figure 4.5A and 4.5B involve movement between nuclei. However, this is not indicated over those arrows.)

\( 1_w \) = this is an indicator function on the hippocampal signal to the PRF, which takes the value 1 when novelty is recognized.
1- INTRODUCTION

The psychological process through which an animal understands its environment is defined as perception (Zimbardo 1978). The perceptual process must deal with an overwhelming amount of information entering the central nervous system (CNS) on nerves coming from various sensory receptors. To reduce this information to a level that can be managed, man and other mammals employ a number of processing techniques. These include many types of filtering and regrouping to reduce the information load. However, the most important technique for handling sensory information is comparison with stored sensory experience. The difficulty in analyzing a visual input from scratch is illustrated by examining a photograph of a small piece of an object without any knowledge of scale, orientation, or surroundings. This confusion is prevented by the recognition experience that results from the agreement between the current and the stored information. The recognition experience releases the full power of previously stored information, by placing the current input in a context and building up expectations about the structure of the information that is to follow. Thus, the sensory input is not analyzed from scratch, but in comparison to something else.

A basic hypothesis of physiological psychology is that activity in the central nervous system (CNS) mediates sensory experiences, and that these experiences can be used to guide the animal through interactions with the environment. It is difficult to believe that a complicated
subjective experience such as recognition is generated solely by the electrical activity in the central nervous system (CNS). Yet, the discovery of ever more refined methods of controlling and storing neural activity are beginning to make this idea plausible. Perhaps the strongest support for the belief that neural activity underlies consciousness comes from radical changes in the cognitive processes of patients with brain lesions (Luria 1973). There is no assumption made here that neural activity is sufficient to explain all experiences and behaviors. The question to be addressed is what do they explain. In particular, this paper explores that part of neural processing, primarily in the visual modality, which leads to the recognition experience. A note of caution at the outset of this task is that recognition is part of a matrix of experiences and behaviors generated by perception. However, recognition provides a central point of view around which the present study is organized.

1.1- THE RECOGNITION PROCESS

The definition of recognition implies contributions from at least two information sources, including some subset of information from the sensory input and some subset of stored information about past experience. These two sources are insufficient to encompass the variability of the recognition experience. What is implied when a person states that he or she recognizes a scene? The knowledge entailed in the recognition experience in response to a presentation of a scene is quite variable in humans even when the relevant experience remains
fixed. Recognition certainly implies there is a separation of objects from background. Also required is a feeling of familiarity elicited by the scene. Some of the attributes of a scene that might or might not be known are: its context and location (i.e. what is around it); the objects composing it, their shapes, sizes, uses, and names; and what changes in the scene have been made since the last viewing. This variability in the knowledge entailed in the recognition experience must come from a source beside the visual input or past experience.

It is proposed that this third source of information is the current brain state as characterized by arousal and attention. Zimbardo (1978) defines arousal and attention as follows: arousal— to arouse attention, or to excite into activity, attention — the concentration of the mind upon an object. Thus, it may be seen that arousal is a general readiness to react, which is a prerequisite for attention. Furthermore, attention is the specific filtering of the available information, which is a prerequisite for analyzing a scene.

The brain state interacts with both sensory input and the past experience. The minimum intensity of a stimulus that can be noticed is a decreasing function of arousal. It is easily observed that the intensity at which a stimulus is just noticeable rises as an animal goes from wide awake, to drowsy, and then to sleep. The amount of information gained through perception of a visual scene is contingent on the amount of detail reaching threshold intensity. The specific allocation of attention also affects whether a subject will become aware
of an object. The threshold intensity in waking is fairly variable, apparently depending on the direction of the animal's attention (Fribram 1971). Similarly, whether a recognition experience occurs may depend on verbal instructions given before the scene is presented (Kahmeman 1973 p.61). To study the details of these effects, the targets of attention must be identified. In section 5.4, the allocation of attention as well as the way the allocation influences recognition is explored.

Note that the store of past experience must be considered a changing set. As more experiences are added to the store, there are changes in the recognition experience resulting from the same sensory input. Classical conditioning is a good example of the variability of response to the same retinal input over time. This implies that the past experience on which recognition draws is an ongoing phenomenon, not based totally on innate programming.

To achieve a workable model of recognition requires representations of each of the three sources of information, as well as a method for describing their interactive dynamics. Such a model must eventually give correct answers to the following questions about storage of experiences and attention. What part of the information contained in a visual input can be stored in the brain? To what kinds of changes in future perception can the storage of the current perceptual experience lead? How is the point at the center of attention chosen as it varies over the examination of a scene?
One approach to modeling recognition is to find or construct a system that will store information, direct attention, and combine stored information with visual input to achieve recognition in the same way as a live subject. If this could be done, would the resulting system give the correct connections and response properties of neurons in the visual system (see Figure 4.4)? There is a theoretical problem that arises in trying to deduce the structure of a system from its function, also known as the top-down or inverse modeling approach. When the functions of a system are characterized by the input-output pairs, Padulo and Arbib (1975) have pointed out that there are many systems that can realize the given set of input-output pairs. Note that there are many systems that can reproduce a single-sensory function, such as that of orienting toward a bright stimulus, found in different members of the animal kingdom. These systems may bear little structural resemblance to each other (see Figure 4.1 and subsection 2.2.3). To build one model of a real organism with a top-down approach, a list of every function that could be found would have to be made. From the list of functions, one could also be chosen and one of the models capable of realizing this function could be selected. The other functions could then be added to the model one at a time. Hopefully, a simple modification of the original model would allow the performance of the additional functions. More likely, the model of the system would have to be rebuilt from scratch many times to achieve the joint functions in a single model. It should be noted that such programs can be quite useful if the objective is not to model a real organism, but to build a device that can perform
a given set of functions as in Artificial Intelligence projects.

1.2- NEURAL MODELS

The controversy over whether the substrate of nervous activity is composed of single nerve cells or is an undifferentiated network was settled in favor of single nerve cells around the turn of the century, through the work of people like Ramon y Cajal and Waldeyer. Since that time, a great deal has been learned about the types of neurons in the brain, how they make connections, and how they interact. In addition, over the past century considerable research has been done on the psychophysics of higher brain functions like recognition, perception, memory, and cognition. Nonetheless, questions like why one sees a face as a face still cannot be explained by single cell mechanisms (Poppel 1979 p.324). A number of authors have come to the conclusion that the bridge between single neuron mechanisms and psychophysics is based on cooperative effects that emerge from the long range interactions of neural masses (Edelman 1979, Thatcher & John 1977 p.85, Katchalsky, Rowland, & Blumenthal 1974). In fact, there are two levels of neural organization between single neurons and the whole central nervous system, one structurally defined and the other functionally defined. The functionally defined level are columns which respond to a particular quality of the input such as line orientation (Hubel & Wiesel 1977, Mountoastle 1979). The units of the structurally defined level are nuclei which are bounded collections of cell bodies, such as the lateral geniculate nucleus (LGN) (Thompson 1967). An underlying assumption in this paper is that a behavior of a structural unit at one level is
supported by the global activity of a large number of units at the next lower level of organization. In particular, a whole system CNS behavior such as recognition should be supported by activity at the nuclear level. Stated in another fashion, it is assumed that activity of single units at the neuronal or columnar levels would not determine the scene that is being recognized without specifying the interactions among these units. This leads to the hierarchical notion that the scene being recognized could be determined by the activity of units at the nuclear level and perhaps their interactions.

Support for this concept comes from the fact that single neurons respond to many visual objects in any brain area studied (Gross 1974). Powerful evidence for the claim that the individual activity of neural or columnar level units do not support the higher brain functions has been obtained from data on recovery after lesion. The months it sometimes takes animals to become proficient at visual recognition tasks after extensive damage to their visual systems argues for a gradual take over of gross functions (Pribram 1971 p.119), such as response to luminosity gradients. In almost all cases, there are areas which if left intact lead to recovery of function, but which do not severely affect the function when they alone are damaged. Note that the function may be carried out in a very different way after the 'recovery'.

The most successful approaches to explaining a global behavior of a system, on the basis of some property of a vast number of individual components that make up the system, have arisen in physics. First, a
set of variables describing the global behavior in a simple system must be found. Then the properties of the individual particles that effect each global variable are located. Finally, a rule for producing these effects through some combination of individual particle properties is established. An example of how this approach has been applied, taken from the kinetic theory of gases is the derivation of the pressure on a box from the motion of the individual particles within it.

The search for general rules of neural operation in simple nervous systems is being carried out in a number of studies, for example Kandel's (1979) studies of the snail, Aplasia. In such an animal, there is no columnar level of neural organization and fewer cells in the whole organism than in a single column of the mammalian cortex. Single neurons may play a role that is performed by whole nuclei of higher animals. However, related mechanisms of neuroregulation probably exist in mammals. Thus the understanding of higher brain functions may eventually result from the synthesis of neural rules found in simple animals.

The primary objective in this study is to link the subjective experience of recognizing a scene to the intricate patterns of neural activity that are simultaneously occurring in the CNS. Some characterization of these patterns and the substrate on which they occur must be given before the part relevant to recognition can be extracted. The basic model considered in this paper seeks to characterize the changes in electrical potential on single neurons. Three properties of neurons which are incorporated in the model are:
1) the different functional cell types to which a particular neuron can belong,
2) the connections among the cell types,
3) and the response properties of each cell type to an input.

A neural model is defined by a list of neurons specifying position and cell type of each, the spatial distribution of each neuron's connections, plus a description of the response function of the target neurons. The main sources of data for this part of the research are derived from anatomical studies on synapses, neurons and nuclei. Additional data come from studies on single cell response to stimulation, results of which are expressed either in terms of post synaptic potentials (PSP) or changes in firing rate, i.e. the number of action potentials (AP) per unit time.

How much of the vast amount of available detail should be included in the neural model? There is no clean cut answer to this question. At a general level, one answer is the simplest model whose set of mechanisms explain the organism's behavior to the satisfaction of the modeler. The method of obtaining this answer is generally iterative, with the exploration of the properties of one model leading to ideas on what must be included in the next model.

Automata theory offers a theoretical basis for the claim that neural circuits are sufficient to explain the logical functions of the brain. Turing showed the existence of a finite automata with an infinitely extendible tape, such that any program performed by any other automata
(i.e. any logical sequence) can also be performed by his 'Turing machine' if it is correctly programmed. McCulloch and Pitts demonstrated that a Turing machine could be constructed from a network of formal neurons (von Neuman 1966 p.49). Unfortunately, there are far too many cells and connections for a 1-1 computer simulation of the model. Therefore, any computer simulation of the interaction of single cells must look at a small segment of the system and assume that any other portion works about the same way as the modeled segment. This assumption is untenable if, as accumulated evidence suggests, the most important factors in recognition are concerned with communication between distant regions of the visual system. The alternative to leaving out the long range effects is to work with a hierarchical model that treats a group of neurons as the basic processing unit.

Another practical limitation on deducing function from structure comes from the quality of the available data, which although vast are neither consistent or complete. Some of the deficiencies in the data on which neural models must be based are the following:

A- While most of the connections or tracts among nuclei of the visual system have been uncovered in the last few years, the nature of the connections among various cell types (see Figure 2.1) is largely unknown (Szentagothai 1978).

B- The classical method of spreading neural information, i.e. signals entering on the dendrites, summing at the axon hillock, and being transmitted over the axon, accounts for only part of the total spread. Other sources of spread are:
1) effects of interneural hormones on information processing and retrieval (De Wied & Bohus 1979);

2) extracellular waves of metal ion concentration which are significant in controlling neural firing rates (Adey 1979);

3) and variations and anomalies in the structure, boundaries and synaptic arrangement of some neurons. Some neurons have no real axon. There are also many nonstandard synaptic arrangements: often axons terminate on other axons, dendrites have synapses onto other dendrites, and other permutations occur (see Figure 2.2; Gray 1974, Shepherd 1974).

C- Some neurotransmitters have long lasting (minutes to days) effects on the firing rate of their target cells (Inversen 1979).

D- In those neural models that hope to describe recognition, it is vital to incorporate a mechanism for the storage of visual information. Most neurophysiologists think that Hebb's (1949) theory of synaptic modification must play a role in permanent storage of information. Therefore, the efficacy of models which deduce function from structure is severely limited by the lack of knowledge of the conditions under which synaptic modification takes place.

In spite of automata theory, a neural network with unrealistic connections and incorrect parameters is not likely to perform the desired physiological functions.

Besides the problem of choosing connections and parameters, the data presents one further problem in interpreting a neural model. The output of a neural model is the predicted activity of neurons in a region in
response to a particular input. On the other hand, the behavior of the organism is given in terms of movements and descriptions of subjective experiences. There is a very weak correlation between single cell firings and overt behavior (Fox 1970). Thus, to relate these two kinds of data a third kind of data is required. The combined cellular activity over an extended neural group that occurs during an animal's behavior can be measured using EEG methods. (In this paper, EEG refers to any macroelectrode measurement that is being influenced by more than a few hundred cells.) Fox (1970) showed the relation of EEG to single cell measurements is fairly good, while the relation of EEG data to overt behavior is much better than that of single cell measurements (Regan 1973). Thus EEG data can be used to bridge the gap between single cell activity and overt behavior.

1.3 THE APPROACH TO MODELING

This paper is not an attempt to develop the characteristics of a particular model, but rather an attempt to devise a framework for modeling neural processes leading to recognition. Therefore, only qualitative properties of the models proposed are developed. The decision as to what type of model is appropriate to various types of neural processing is the main goal which was sought.

During this project several interrelated postulates were formulated which guided the construction of models.
1) Any higher cognitive process combines sensory input, past experience and momentary intentions.

2) The neural substrate for these inputs must lie in widely separated regions of the brain. This implies that the long range interactions between regions of the brain must be considered in any model of cognitive processes.

3) Overt and cognitive behavior results from the coordination of activity, rather than from activity in individual neurons, columns, or nuclei. This implies that a neural model should distinguish the class of relay cells which provide the communication between nuclei.

4) Neural models must account for the variety of cell types in the nervous system based both on response characteristics as well as output distribution and effects.

5) There is a great deal of noise in the nervous system, so it is postulated that activity which leads to cognitive behavior must arise in a coordinated way building into self-sustaining activity.

The main goal can be restated as clarifying these postulates and finding a reasonable level of neural organization at which they can be applied to a model.

It was observed that models involving single neurons in the literature fall mostly in two classes: those that examine the properties of arbitrary networks of neurons and those that model the response properties of individual neurons. To understand behavior, one must associate response properties with neural activity. One model that combines response properties with the concept of interactions among
neurons was presented by Marr (1976). However, the interaction was not based on physiological connections. A model that deals with the interaction of nuclei is the holographic analogy discussed by Pribram and others. The most attractive feature of the holographic analogy is that it incorporates involvement of a large part of the CNS in memory formation and utilization, plus a need for cooperative activity among different nuclei. A neural hologram requires a concept of a neural wave with properties analogous to those of an optical wave. Although some evidence supporting the existence of neural waves is available, it is not clear that the assumptions required for neural waves are generally true. Neural waves were assumed to exist, so that the mathematical formalism of holography could be applied directly to the nervous system to find a method of deriving the activity in one nucleus from the activity in another. There are several questions intrinsic to applying a holographic model to recognition, including the significance of individual neuron response characteristics and the nature of neural waves.

In order to incorporate the importance of cell type and the response of individual neurons in various nuclei into a model, a specific neural model was developed. The mathematical study of idealized neurons was begun by McCulloch and Pitts (1943). Beurle (1955) incorporated the idea of inhibitory and excitatory populations into a neural model. Most of the models since that time have considered the processing properties or conditions for synaptic modification in networks with an inhibitory and an excitatory population. The neural model introduced here uses differential and integral equations to describe the distribution of
firing of one cell type as a function of the previous firing of of the other cell types, as did the previous models. However, it is argued that a two population model is inadequate for modeling higher brain functions. Three sources of data suggest that a more elaborate model is required.

1) Two pairs of excitatory and inhibitory cells are suggested by Szentagothai's anatomical studies of area 17. Hence, a model that includes small excitatory interneurons (s.e.i.), small inhibitory interneurons (s.i.i.), large inhibitory neurons, and excitatory relay cells is proposed. The axons of the first two cell types remain in the same nucleus and are the prime recipients of sensory input, while the second two cell types have axons that leave the nucleus or at least travel extended distances and provide the communications between nuclei. Some of the connections among the cell types and their spatial distributions are known and these are included in the model.

2) A columnar axonal distribution for relay cells is suggested by the data of Szentagothai, Mountcastle, and Hubel & Wiesel. It seems that cortical afferents spread their axons throughout a 300 micron column and a number of axons may enter the same column (see Figure 2.6).

3) Variable model parameters controlled by midbrain and/or association cortex inputs are suggested by data of Singer, Bloom, and Inversen. To incorporate these neural properties, a four population model with variable parameters is advocated. The known connections in area 17 lead to a model that can be decomposed into a two level hierarchy, which greatly simplifies the analysis of this model. This decomposition is
based on the lack of connections from the relay and large inhibitory cells onto the interneurons in the same column. The comparative distances and time over which processing in the two subsystems take place allows the interneuron network to be treated as continuous in space and time. The interneuron network is shown to be very similar to a neural network model analyzed by Wilson and Cowan (1973). Some of their results are applied to this model.

The model assumes the physiological restrictions that the effect of any single cortical input is small and that a convergence of inputs is required to change the firing rates of target neurons. When there is a convergence of actively firing relay cells on a column, the result would be stimulation of most of the relay cells in the column. This configuration leads to a very interesting type of self-sustained activity. It is shown that this activity, here termed a turned-on active circuit has the required properties for a neural wave. When the concept of synaptic modification in regions of high activity under special circumstances is added to the active circuit concept, the result yields a plausible method of long term information storage. Finally, the active circuit idea can be related to the control of attention.

The comparison of the neural model of the relay system to memory and attention required reasonable models of these processes. The models used were synthesized from a number of models in the literature. The model of memory is essentially a catalogue of the varying characteristics of stored information. The model of attention during perception is a breakdown of the steps involved in the process. The end
The result of this program is three related models at different levels of the hierarchy of neural organization.

The neural model is presented before the discussion of holography, since the concept of an active circuit motivates the assumptions about the properties of neural waves which are critical to neural holography. Since neural holography is primarily a model of long term storage of information, a discussion of the characteristics of information storage in the brain precede it. The controversy over the properties of memory makes it critical to specify exactly which characteristics of memory are being assumed. The discussion of attention was left for last since it is the most complicated, and makes use of concepts introduced in earlier chapters. Here also, the controversy over the nature of attention required specifying the assumption made, and in fact formulating a psychological model of attention.
2- A GENERAL MODEL OF THE VISUAL SYSTEM

The CNS is a highly complex structure composed of billions of cells, each of which is a sophisticated organic entity. That portion of the CNS which responds to visual input is shown in Figure 4.4. Information is transmitted through the CNS in signals traveling over neurons by passing through specialized membrane structures called synapses. There are a number of classification schemes for neurons based on various physiological response or anatomical characteristics. In area 17, one example of a physiological classification of cell types is simple, complex, and hypercomplex neurons described by Hubel and Wiesel (1977). Ramon y Cajal (1899) gave another classification of cell types based on their anatomical shapes for almost every nucleus in the CNS. An anatomic rather than a physiological basis for the neural model is chosen, because the processing properties are the primary concern. A group of neurons in the auditory cortex with the same connections as a group in the visual cortex would have the same processing properties, but a totally different physiological response. The anatomic model must satisfy two important criteria: a) it should be so simple that it applies to most nuclei, b) it can nonetheless support the processes underlying recognition when the response properties of individual neurons are incorporated. Perhaps the most satisfactory breakdown admits four cell types: small excitatory interneurons (s.e.i.), small inhibitory interneurons (s.i.i.), large inhibitory cells, and relay cells (see Figures 2.4 and 2.5). Szentagothai (1978) has classified the various neurons in area 17 that belong to each of these cell types.
Figure 2.1 shows the shape of the different cell types, including the axonal and dendritic branching.

Although neurons make up about half the mass of the CNS, they account for only 10% of the cells in the CNS (Shepherd 1974). The rest of the cells, except for some blood vessels and connective tissue, are support cells called neuroglia. The most important known function of the neuroglia is performed by a type known as Schwann cells, which wrap themselves around the long axons of relay cells and form the myelin sheath (Figure 2.3), which is critical for the conduction of action potentials (AP). Other types of neuroglia surround large groupings of neuron cell bodies and form the boundaries of nuclei (Thompson 1967).

Most of the nuclei in the visual system can be thought of as thin sheets or a number of thin sheets one on top of the other. Most of the cell types within a nucleus are randomly distributed with respect to the surface of these sheets and are consistent in terms of the distribution of connections that they make with their own and other cell types. Section 2.1.1 has information about the connections among and within nuclei.

In carrying out the proposed function from structure approach to building a model capable of supporting the processes leading to visual perception, the first decision to be made is the choice of the underlying units. There are three levels of anatomic structures in the visual system that seem to play a key role in the transmission and
processing of information. Synapses, the units at the lowest level, are part of the cell membrane of neurons, while neurons are the units at the middle level. Neurons in turn are parts of nuclei, the units at the highest level of neural organization below the whole visual system. The distribution of lines of communication among synapses and among neurons give rise to two more levels of functional organization.

One non-structural, functional unit is inferred from data showing there are sequential and reciprocal synapses in many parts of the CNS (Shepherd 1974). Some of these are shown in Figure 2.2. There may be closed circuits that process information involving only small pieces of several neurons. Another type of processing—involving a small group of synapses is seen in glomeruli which are found in many nuclei (Szentagothai 1970). Rakic (1976) calls these kinds of processing units local circuits.

The other type of non-structural, functional unit is inferred from the axonal distribution of relay neurons. These relay axons end on a 300 micron column containing on the order of ten thousand neurons. Localized lesions and tracer techniques have shown that axons from different regions of the cortex are distributed throughout the same 300 micron column (Szentagothai 1978). Therefore, the levels of information processing in the functional heirarchy are: synapses, subneural local circuits, neurons, columns, pathways, and the whole visual system.
FIGURE 2.1- NEURON SHAPES

Neurons A, C, F and H could be classified as relay cells, neurons D and E are small excitatory interneurons, neuron G is a large inhibitory cell, and neuron B could be either a small inhibitory or excitatory interneuron (From Bailey's Textbook of Histology).
FIGURE 2.2: TYPES OF SYNAPSES

A: Cerebral cortex — pyramidal cell. B: Serial synapse in the spinal cord. C: Serial synapse in LGN. D: Mitral cell connections of the olfactory bulb. E: Amaurotic contacts of the vertical lobe. F: Synapses from a central node of Ranvier. H: Electrical synapse (fish brain). I: Electrical synapse showing the morphology of a chemically transmitting synapse. J, K, L, M: Various aspects of gap junctions involved in electrical transmission. N: An electro-receptor unit (fish skin). O: Receptor synapse of the retina. The processes invaginated into the presynaptic bag are two horizontal processes (lateral) and one bipolar cell process (medial). Abbreviations: db.: dense body in sensory cell; den.: dendrite; fe.: filamentous contact; gr.: granule cell; is.: initial segment; lc.: large cell body; mc.: mitral cell; msf.: medium superior frontal fibre (axon); my. myelin sheath; ca.: olfactory axon; per.: perikaryon; s.: synaptic ribbon; sba.: synaptic bar; sp.: spine; sv.: synaptic vesicles. (From Gray 1974)
FIGURE 2.3: THE MYELIN SHEATH AND NERVE CONDUCTION

Diagram 2.3A is schematic diagram of a mammalian relay cell with a medulated nerve axon. The medullary sheath consisting of myelin, is interrupted at regular intervals by the nodes of Ranvier. Diagram 2.3B shows cross sections through a medullated (left) and an unmedullated nerve axon. (From Schmidt 1974)
FIGURE 2.4: EXCITATORY CONNECTIONS IN ONE AREA 17 COLUMN

See legend page 25. (From Szentagotai 1978)
Legend For FIGURE 2.4

Modular arrangement of excitatory connections and of assumed excitatory interneurons. The large cylinder of diameter 300 microns corresponds to the space of termination of a cortico-cortical (ipsilateral association or contralateral callosal) afferent, apart from lamina I, where the horizontal spread of the terminal branches may be considerably larger. The flat cylinder of the same diameter would correspond to the termination space of a specific (sensory) afferent. Two different types of spiny stellates are shown as monosynaptic target cells of the specific afferents: ss, has both an ascending and a descending axonal strand, while ss, has only one ascending strand. Microgliaforms cells (mg) have more generally descending axon strands of similar diameter (around 30 microns). They are also potential monosynaptic recipients of specific afferents, but there is no direct evidence for this. A typical 'cellule a double bouquet' of Ramon y Cajal (c.d.b.) is shown at upper left, giving rise to a long vertical axon strand of even smaller diameter.

Legend For FIGURE 2.5

Putative (conventional) inhibitory interneurons and the spatial distribution of their axonal arborizations. Main part of the diagram shows the various types of interneurons in full black, while the inset diagrams (a), (b), (c) at right indicate the electron microscopic evidence upon which the assumption of the inhibitory nature (s.f. type synapses of fibers indicated in dark) and local origin (persistence in isolated cortex) of the corresponding synapses is based. The large basket cells (l.b.c.) terminate in thin parallel vertical slices of layers III-V on cell bodies of pyramidal cells. The small basket cells (s.b.c.) of lamina II have ranges of 100 microns, while columnar basket cells (c.b.c.) may have descending or ascending (or both) axons that terminate within cylindric spaces of not more than 50 microns in diameter. Axonal tuft cells (a.t.c.) of lamina II contact spines of apical dendrites (see inset (a)) with s.f. type terminals in double contacts, (mainly in lamina I, but similar double a.r.-s.f. contacts are present also in the deeper layers of the cortex). Chandelier cells (c.h.c.) contact mainly the apical dendrite shafts of pyramidal cells (see inset (b)) with s.f. type synapses, while the spines have mainly a.r. type (excitatory) contacts.
FIGURE 2.5: INHIBITORY CONNECTIONS IN ONE AREA 17 COLUMN

See legend page 25. (From Szentagothai 1978)
The activity of a synapse, is characterized by the membrane potential on the receiving side of the synapse, and the electronic spread of potential determines the effect of this synapse on other synapses entering the same neuron. A neuron's activity, in terms of its communication with other neurons, can also be characterized by a single variable, its momentary average firing rate. To define a momentary firing rate it is assumed that many action potentials (AP) arrive during the period of a post synaptic decay (PSP), and that the firing rate over such periods varies slowly. Very slow rates of firing contribute little to the firing rate of target neurons, so their accuracy is not critical.

At each level of the hierarchy there is an organizational principle which allows the activity of the lower level units to be combined to give a single value which characterizes the higher level activity. In going from synapses to neurons, the principle that allows this massive reduction in information is the summation of synaptic potential into a single axon hillock potential. This potential is converted, in a non-linear way, into average firing rate. In going from neurons to a column, the principle seems to be the in-register termination of the extended axonal fields of the relay cells. The input from cortical columns causes the potential on neurons across a whole column to shift by roughly the same amount and possibly shifts their firing rate as well.

Theoretically, the operation of the neuron level could be built into a synaptic level model. This would require accounting for the influence of synapses in different neurons on each other via signals transmitted
over axons, as well as the electrotone spread of potential between
synapses on the same neuron. Since there are a few hundred to a few
thousand synapses per neuron (Shepherd 1974), such a model would be at
least three orders of magnitude more complicated than a neural model (as
defined in subsection 1.2). At the neural level it would give no more
information than the firing rate of neurons would give. Thus, the
amount of complexity makes it vital to study one level of the hierarchy
at a time, and within that level to decompose the model into groupings
based on the units at the next higher level. Therefore, models of the
synapses in a neuron, neurons in a column, and columns in a pathway,
i.e. strongly connected nuclei, should each be explored before global
models at any level are attempted.

2.1- A MODEL OF NEURONS WITHIN A COLUMN

A number of general features of synapses are needed for the
construction of a neural model. The two basic structural types of
synapses are: (a) asymmetric synapses with round vesicles (a-r
synapses), (b) symmetric synapses with flattened vesicles (s-f synapses)
(see the insert in Figure 2.5). The a-r synapses are often found on
dendritic spines, a fair distance up the the dendrite, while the s-f
synapses tend to end close to the cell body of the target neuron (Gray
1974). Evidence indicates that the two structural types also differ
functionally, the a-r synapses being excitatory, and the s-f synapses
being inhibitory (Popper & Eccles 1977). A second important functional
difference among synapses, beyond whether they are excitatory or
inhibitory, is the type of neurotransmitter they contain in their vesicles. Several monoamine neurotransmitters known to play an important role in the control of attention and arousal, are produced largely in localized regions of the midbrain (Jouvet 1967, Inversen 1979).

Area 17 will be used as an example in which to study the distribution of the four major cell types listed above. Though these cell types are essentially uniformly distributed with respect to the surface, different cell types are found at different depths (see Figure 2.6, Thompson 1967 p.303).

2.1.1- Densities, shapes, and connections of the cell types

Some estimates of neuron density and of densities of the axonal fields in area 17 are available that prove useful in determining the model parameters. The human cerebral cortex is a convoluted sheet of tissue approximately 500,000 sq mm in area and 2 mm in depth (Shepherd 1974, see Figure 2.7). The estimates of the number of neurons in the brain have increased over the last few years, from several billion to around 40 billion (Peele 1976). There are only slight differences among the neuron densities among the various cerebral regions, but the occipital region is relatively high (Szentagothai 1978). As a rough estimate we will take two billion as the number of neurons in area 17, 25,000 sq mm as its surface area, and 50,000 cu mm as its volume in humans. This yields a neuron density of 40,000 neurons per cu mm in
FIGURE 2.6 AXONAL ARRANGEMENTS IN AREA 17
FIGURE 2.7: THE CEREBRAL CORTEX: THE LATERAL AND MEDIAL SURFACES
area 17. Of this number, an estimate is needed as to the percentages that belong to each of the four major cell types included in the model.

Evidence that there are far more excitatory than inhibitory synapses on relay cells in the cortex (Guillery 1965), implies more excitatory than inhibitory neurons. Shepherd (1974) claims that pyramidal or relay cells make up on the order of ten percent of the neurons in the primary sensory regions of the cortex. Estimates for the motor and association areas are even higher (Sloper, et al 1978). From these data and attempts at cell counts (Powell 1978), the following rough estimate of cell types in area 17 can be made; 70% small excitatory interneurons (s.e.i.), 15% small inhibitory interneurons (s.i.i.), 10% pyramidal cells, and 5% large basket cells (l.b.c.). These percentages are used to estimate the strength of connections between cell types and the number of pyramidal cells leaving or entering a column.

The sizes and shapes of the axonal fields of the different cell types are quite different. The following discussion refers to Figure 2.4 for the excitatory neurons, to Figure 2.5 for the inhibitory neurons, and to Figure 2.6 in general. The most common type of excitatory interneuron, the spiny stellate cells of layers 3 and 4, have a cylindrical axonal field 30-50 microns in diameter and extend over most of the 2 mm depth of the cortex (see Figures 2.4 and 2.6). The inhibitory interneurons have the most diverse shapes of the four populations with particular shapes concentrated at particular depths. More anatomical data is required before it can be determined whether the cortical neurons are working directly or through an intermediary. The most common shape is
probably roughly spherical and 100 microns in diameter, but some are similar to the spiny stellate in shape. There are others that extend primarily in the horizontal direction (see Figures 2.5 and 2.6). The large basket or large inhibitory cells are found in the middle depths of the cortex, with axonal fields that form slabs 500 mm high and 50x1000 sq.mm. in area 17 (see Figures 2.4 and 2.6). Finally, the pyramidal cells are found at all depths with the greatest concentration of small pyramidal cells in layers 2 and 3, and the greatest concentration of large cells in layers 5 and 6 (Thompson 1967 p.303). The pyramidal cells have a myelinated axon that carries signals to other regions of the brain where it forms a cylindrical column the full depth of the cortex or subcortical target which is 300 microns in diameter. Some axons may be concentrated in horizontal plexi in particular layers. As noted above, the cortical afferents from different areas tend to terminate on the same discrete 300 micron column and this seems to be the basis for a higher level of neural information processing (see Figure 2.6). In addition to the main axon, most pyramidal cells have an average of ten collaterals that leave the initial segment of the axon and extend to a columnar region 1-3 mm in diameter. These collaterals seem to have a primarily inhibitory effect on the surrounding region. Note that the LGN afferents to area 17 are also 300 microns in diameter, but are confined to a small part of layer 4 rather than filling all the layers like the cortical afferents. Whether the columns defined by the LGN input are in register like the cortical afferents is unknown, but unlikely (Szentagothai 1975).
An additional point about the axons of the different inhibitory interneurons is that they are found at different depths in the cortex, and thus make most of their connections at a particular position on the relay cells. Thus some inhibitory interneurons neutralize excitation only from the upper part of the apical dendrite, others affect the cell body and stop all firing, while another type ends on the upper segment of the axon (Szentagothai 1978). This configuration may allow signals to be transmitted over some of the collaterals, but not on the main axon. Complementing this finding is the evidence that inputs from various parts of the brain end at different levels on the apical dendrite (Schiebel & Schiebel 1970). So, different input may be neutralized selectively by one type of inhibitory interneuron.

Let us consider the overlap of axonal fields within a single 300 micron diameter by 2 mm high cylindrical column in area 17, since such a column is the next level of information processing. In accordance with previous estimates of surface area and cell density in area 17, there are about $[25,000 \text{ sq mm} / 0.1 \text{ sq mm per column}] = 250,000$ columns in area 17. Since the volume of such a cylinder is 0.2 cu mm, there should be about 8000 cells in it. Combining this with the estimate presented above for the percentage of cells of each type gives 5600 s.e.i., 1200 s.i.i., 800 pyramidal cells, and 400 large basket cells per column. The overlap of the axonal fields within a single 300 micron diameter by 2 mm high cylindrical column in area 17 is needed to calculate the connective strengths of the various cell types.
The volume of the axonal field of each s.e.i. is about .001 cubic mm. Therefore, 200 of these axonal fields would fill a column if they did not overlap. Thus, 5600 of them fill the column 28 times (see Figure 2.6). Note that the volume of the actual axon material is negligible compared to the volume it pervades. The s.i.i. have axonal fields of about the same size, and thus would have an overlap of 1200/200 = 6 times. However, their circular shape gives them a larger gross subsectional area as seen from the surface. The axonal field of the large basket cell is about .025 cubic mm. and 8 of them would fill a whole column if there was no overlap. Actually, they are so large that they would overlap the column where their cell bodies lie. The 400 of them in a column would overlap their axonal fields 50 times. The main axon of the pyramidal cell does not terminate in the same column at all. Since the axons of most cortical pyramidal cells end in the cortex (Mountcastle 1979), 50% of the axons leaving a column is a very conservative estimate for the number of cortical afferents entering it. Since each cortical afferent fills a column, this would mean an overlap of 400 times.1

A last observation about overlap is that there are on the order of a million axons going from LGN to area 17 (which is also the number of fibers reaching the LGN from either the retina or area 17 (Singer 1977)). Since little branching is seen on stains of these fibers, an

1- Mountcastle (1978) has pointed out that there is an input isolation effect in the primary sensory areas. The LGN input dominates in area 17, there very few u-shaped pyramidal axons returning to area 17, and there may be fewer or less influential cortical afferents than in the deeper regions of the CNS.
estimated 8 LGN fibers should end on each area 17 column.

These overlaps in the various axonal fields can be seen in Figure 2.6. A question that must be considered when constructing a model is the purpose of this massive overlap. In general, the dendrite field of these cells is much smaller than their axonal fields. For instance, in the pyramidal cells, the axon field is 300 microns in diameter, whereas the apical dendrite branching is fairly scarce below layer 1. This would allow a pyramidal cell to sample only a small fraction of the interneurons in a column. It is likely that the narrow apical dendrite of a pyramidal cell would be effective in selecting particular categories of information. Thus, different qualities of the input within a column can be separated and sent to different targets. Section 2.2.3 discusses the basis for the separation of input qualities by interneurons within a column.

In constructing a neural model it is important to determine which cell populations are connected and how strong their influence is. Not enough is known about this critical subject. Some progress in this area has recently been made. Somogyi (In press) went through the difficult process of combining degeneration with electronmicroscopy and used computer reconstruction to trace the synaptic connection on several neuronal types. He destroyed the LGN and then thin-subsectioned area 17 looking for the degenerating fibers in the sequences of subsections. He was able to show that most of the LGN input went to the spiny stellate cells in layer 4, although some ended on the basil dendrites of layer 3 pyramidal cells. Kelly and Essen (1974) used an electrode containing
dye to record single cell responses from area 17 neurons in cats, and later located the stained cells from which they had recorded. Due to the timing they claimed that Y-type LGN fibers (also called fast or transient) had inputs directly to the pyramidal cells which incidentally responded like the Hubel and Wiesel complex cells. It was also claimed that X-type LGN fibers (or sustained or slow) had inputs to spiny stellate cells that responded like the Hubel and Wiesel simple cells. It is still unknown whether LGN input directly contacts inhibitory cells or the apical dendrites of pyramidal cells. However, Somogyi also found that the spiny stellate cells made multiple contacts (due to their elongated shape) with the apical dendrites of the pyramidal cells in their axonal field. Although not directly confirmed, it seems highly probable that these excitatory interneurons must also contact the s.i.i. Observations indicate that the large basket cells may end exclusively on the cell bodies of the pyramidal cells. The pyramidal cells seem to be a target for all the other major cell types. This is logical since all the internuclear communication must go through them.

The greatest lack of information at present for neural modeling concerns the nature of the inputs to the two classes of inhibitory cells. About the only definite thing known is that pyramidal cell collaterals inhibit their surround by stimulating some inhibitory cells (Schiebel & Schiebel 1970). Figures 2.4 and 2.5 show the connections assumed to exist in the present neural model. Note that the connection strength between neurons is determined by the number of shared synapses, multiplied by the individual synaptic strengths.
Some regions of the visual system have relay cells whose axonal fields do not end in 300 micron columns. Included are some regions involved in the control of the brain state and in the integration of vision into the internal milieu, such as the reticular formation of the midbrain (RF) and the hypothalamus. The output from these nuclei pervades several nuclei with diffuse contacts, and single axons may contact neurons in many nuclei. Note that some of the input to these regions with diffuse projections comes from cortical areas whose axons have the typical 300 micron in diameter axonal fields. Thus columns based on cortical afferents may be defined in these regions also.

2.1.2- The model

In section 1.2, it was stated that the output of a neural model is the activity of neurons in a region in response to a given input. Neurons carry on many metabolic activities that have been shown to have an effect on information transfer and storage in the CNS. They have rates at which they produce neurotransmitter, actively pump metal ions across the cell membrane, and add proteins to their cell membranes. However, only the electrical activity that leads to changes in the electrical activity of other neurons is included in the present neural model. Neural electrical activity has two distinct forms. The first is the degree of polarization at each synapse on the neuron, and the other is the number of AP per unit time passing some point on the axon, i.e. the firing rate. (For a discussion of how an AP is transmitted and post synaptic potentials (PSP) are generated by a neurotransmitter, see Aidley 1971).
One of these two forms of activity must be chosen to characterize neural activity. The input enters the cell in the form of changes of potential at the synapses (see Figure 2.3 or 2.5). If depolarization occurs the change in the potential due to the input is called an excitatory post synaptic potential (EPSP). If hyperpolarization occurs the change is called an inhibitory post synaptic potential (IPSP) (Thompson 1967). The effects of both kinds of PSP from all the synapses on the neuron sum at the axon hillock at each moment. At the axon hillock there is a non-linear conversion of the potential into an average firing rate (Freeman 1975). When an AP reaches the end of an axon it influences the cell membrane on the other side of the synapse by causing the release of some neurotransmitter.

Knowing the exact time that all the action potentials (AP) entering a neuron arrive will not allow the calculation of the exact time that AP are generated at the axon hillock of that neuron. This is because the firing rate is a stochastic process (possibly Poisson, see Perkel 1967), rather than a deterministic one. An additional problem is calculating the change in firing rate of a target cell due to changes caused by altering the firing rate of one input neuron. Since the transformation of the axon hillock potential into a firing rate is non-linear, the initial potential must be known to calculate this change. Therefore, given a potential on the axon hillock of a neuron, there is still no way to know exactly when an AP will arrive at the synapse with the target neuron.
EQUATION 2.1: A WITHIN COLUMN MODEL FOR AREA 17

\[ \alpha(t) - \alpha_0(I) = \int_0^t \left\{ \exp(\xi-t/\mu_{ex})A_{11}S(\xi-\tau_0) - \exp(\xi-t/\mu_{in})A_{12}B(\xi-\tau_1) \right. \]
\[ + \exp(\xi-t/\mu_{ex})A_{13}T_1(\xi-\tau_2) \]
\[ + \text{frontal + posterior + midbrain} \right\} d\xi \]

\[ S(t) - S_0(I) = G_S[\alpha(t), I] \] (2.1a)

\[ \beta(t) - \beta_0(I) = \int_0^t \left\{ \exp(\xi-t/\mu_{ex})A_{21}S(\xi-\tau_0) - \exp(\xi-t/\mu_{in})A_{22}B(\xi-\tau_1) \right. \]
\[ + \text{frontal + posterior + midbrain} \right\} d\xi \]

\[ B(t) - B_0(I) = G_B[\beta(t), I] \] (2.1b)

\[ \gamma(t) - \gamma_0(I) = \int_0^t \left\{ \exp(\xi-t/\mu_{ex})A_{31}S(\xi-\tau_0) - \exp(\xi-t/\mu_{in})A_{32}B(\xi-\tau_1) \right. \]
\[ + \exp(\xi-t/\mu_{ex})A_{33}L(\xi-\tau_2) + \exp(\xi-t/\mu_{ex})A_{34}T_2(\xi-\tau_2) \]
\[ + \sum_j \exp(\xi-t/\mu_{ex})A_{33}(cj)P^j(\xi-\tau_2) \]
\[ + \text{frontal + posterior + midbrain} \right\} d\xi \]

\[ P(t) - P_0(I) = G_P[\gamma(t), I] \] (2.1c)

\[ \delta(t) - \delta_0(I) = \int_0^t \left\{ \exp(\xi-t/\mu_{ex})A_{41}S(\xi-\tau_0) - \exp(\xi-t/\mu_{in})A_{42}B(\xi-\tau_1) \right. \]
\[ + \exp(\xi-t/\mu_{ex})A_{43}(cj)P^j(\xi-\tau_2) \]
\[ + \text{frontal + posterior + midbrain} \right\} d\xi \]

\[ L(t) - L_0(I) = G_L[\delta(t), I] \] (2.1d)
### DEFINITIONS OF VARIABLES

<table>
<thead>
<tr>
<th>CELL TYPE</th>
<th>FIRING RATE</th>
<th>AXON HILLOCK POTENTIAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- Small excitatory interneurons</td>
<td>$S(t)$</td>
<td>$\alpha(t)$</td>
</tr>
<tr>
<td>2- Small inhibitory interneurons</td>
<td>$B(t)$</td>
<td>$\beta(t)$</td>
</tr>
<tr>
<td>3- Pyramidal cells</td>
<td>$P(t)$</td>
<td>$\gamma(t)$</td>
</tr>
<tr>
<td>4- Large inhibitory cells</td>
<td>$L(t)$</td>
<td>$\delta(t)$</td>
</tr>
</tbody>
</table>

**Delays in traveling from a source to a target axon hillock**

\[
\tau_0 = 3 \text{ ms} \quad \tau_1 = 20 \text{ ms} \quad \tau_2 = 8 \text{ ms}
\]

**Decay rates**

\[
\mu_{ex} = 20 \text{ ms} \quad \mu_{in} = 200 \text{ ms}
\]
The present model is composed of two equations for each cell type. One converts from the probability density for the arrival of an AP from each input to that cell type into a potential at the axon hillock. The other equation converts the axon hillock potential into a firing rate at the same point. In equation 2.1, which defines the neural model within a single column, the variables $\alpha(t)$, $\beta(t)$, $\gamma(t)$, $\delta(t)$ are vectors of the potential at each axon hillock of the whole s.e.i., s.i.i., pyramidal cell, and large basket cell populations respectively. The variables $S(t)$, $B(t)$, $P(t)$, and $L(t)$ are vectors of the average momentary firing rates of the same ordering of the cell types.

Equation 2.1A gives the conversion of the average firing rates of each of the populations sending input to the s.e.i. population into a vector of potential at the axon hillocks of this population. It is assumed that the potentials due to the inputs from each source add in a linear way at the axon hillocks (Freeman 1975). The individual contributions from the inputs are composed of three factors integrated over time. One factor is the already mentioned firing rate of the input population. A second factor is the connection strength matrix $A_{kl}$, where $k$ is the receiving population and $l$ is the sending population. The exact form of these matrices is discussed below. The third factor is the decay rate of the PSP for a particular type of input.

The use of the negative exponential to describe the time course of the PSP needs some justification. The actual shape of either an EPSP or an IPSP involves a fast rise time followed by a slow decay. The typical
EPSP in the CNS has a rise time of 2 ms and a decay time of 20 ms, while an IPSP has a rise time of 20 ms and a decay time of 200 ms (Thompson 1967). Since the rise time of the PSP is an order of magnitude faster than the decay in either case, the contribution of the rising part of the PSP is neglected. The negative exponential is a fair approximation of the decay phase. The excitatory decay constant, $\mu_{\text{ex}}$, is taken to be 20 ms, and the inhibitory decay constant, $\mu_{\text{in}}$, is taken to be 200 ms.

The last type of parameter in the contribution of an input to the potential at the axon hillock is the delay, $\tau_1, \tau_2, \tau_3$. These delays are composed of three components: the time it takes an AP to travel the length of the typical input neuron, the time in crossing the synapse, plus the rise time of the PSP. For example in equation 2.1A, the delay in going from one neuron in the s.e.i. population to another is 3 ms; a half ms travel time, a standard half ms crossing the synapse, and a 2 ms rise time (Shepherd 1974, Fukunda & Stone 1974).

The three factors in each term of equation 2.1A can thus be interpreted in the following way. The effects of an AP at the axon hillock of an input neuron begin causing a potential change at the axon hillock of its target neurons after an appropriate delay. A single AP has a particular effect on the axon hillock potential, which is determined by the number of shared synapses formed with the target neuron and the individual effect of these synapses on the axon hillock. The effects of the input begin to decay exponentially as soon as they arrive.
In equation 2.1A, there are only two contributions to the potential vector of the s.e.i. population from within the same column. There is feedback onto itself (although not from the same neuron onto itself), and a contribution from the s.i.i. population. The s.e.i. inputs which are not involved in the within-column processing can be subdivided into a number of components. There is a well documented input from LGN X fibers, and more speculative inputs from midbrain, frontal cortex, and posterior association cortex.

Note that on the left hand side of equation 2.1A there is a parameter, $q_o(I)$, which is subtracted from the axon hillock potential. This parameter reflects the fact that in the no input situation, the potential does not fall to zero, but to a resting potential. This parameter is taken to be a function of the input from the midbrain and the association cortices. Certain types of midbrain inputs may actually alter the resting potentials and connection strengths of cells of a particular type. Most neurons have some intrinsic firing rate at each midbrain-dependent resting potential. Thus significant levels of stimulation may occur in the 'resting state'. The role of the midbrain input in changing the processing mode is discussed below.

The other equation governing the activity in the s.e.i. population, 2.1a, gives the conversion of the potentials at the axon hillocks into firing rates there. The $G_s$ function is sigmoidal due to the threshold for AP at low levels of axon hillock potential, and to refractory periods of the s.e.i. cells at high firing rates. A specific choice of sigmoidal function is made below. The particular choice is not critical
as long as there are parameters for controlling its shape and the rate at which the firing rate climbs to large values as potential decreases. These parameters will also be assumed to be functions of the midbrain and association cortices inputs. The structure of the other three pairs of equations that make up the neural model in equation 2.1 are identical to those for the s.e.i. population.

As a consequence of the way these equations are written, most of the information processing is hidden in the matrices, $A_{k1}$. These matrices describe how the activity of one cell type distributes its influence to another cell type population in a column. The balance of excitatory and inhibitory input at any position and time is controlled by these matrices. The position of the non-zero values represents the configuration of axonal terminations of the input cell type, l, on the target cell type, k. The size of the value of the component $a_{ij}(k1)$ in the matrix represents the strength of the connection from neuron $j$ of type $l$ to neuron $i$ of type $k$.

The matrices in Tables 2.2 and 2.3 were derived as follows:

1- Pick the desired matrix, say $A_{21}$, which converts the input of the s.e.i. population to an input for the s.i.i. population.

2- Locate the number of each relevant cell type in a column (5600 s.e.i. and 1200 s.i.i.) and the number on a diameter across the column (74 s.e.i. and 34 s.i.i.) in Table 2.1A.

3- Find the layout of the relevant cell types within the column in Table 2.1B.
### Table 2.1: Numerical Data for the Connection Matrices

#### 2.1A: Statistics for the Four Cell Types

<table>
<thead>
<tr>
<th>Cell Type</th>
<th>Number</th>
<th>Number along a diameter</th>
<th>Number filling a column diameter</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) small excitatory interneurons</td>
<td>5600</td>
<td>74</td>
<td>100</td>
<td>56</td>
</tr>
<tr>
<td>2) small inhibitory interneurons</td>
<td>1200</td>
<td>34</td>
<td>20</td>
<td>60</td>
</tr>
<tr>
<td>3) pyramidal cells</td>
<td>800</td>
<td>28</td>
<td>1 (different column)</td>
<td>500</td>
</tr>
<tr>
<td>4) large basket cells</td>
<td>400</td>
<td>20</td>
<td>4</td>
<td>100</td>
</tr>
</tbody>
</table>

#### 2.1B: Neuron Positions Within a Column

<table>
<thead>
<tr>
<th>Numbering of s.e.i. in the column</th>
<th>Numbering of s.i.i. in the column</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1_1 1_2 \ldots 1_{74}$</td>
<td>$2_1$</td>
</tr>
<tr>
<td>$2_1$</td>
<td>$1_1 1_2 \ldots 1_{34}$</td>
</tr>
<tr>
<td>$\ldots$</td>
<td>$2_1$</td>
</tr>
<tr>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$74_1 74_{74}$</td>
<td>$34_1 34_{34}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Numbering of Pyramidal cells in the column</th>
<th>Numbering of l.i.c. in the column</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1_1 1_2 \ldots 1_{28}$</td>
<td>$2_1$</td>
</tr>
<tr>
<td>$2_1$</td>
<td>$1_1 1_2 \ldots 1_{20}$</td>
</tr>
<tr>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$28_1 28_{28}$</td>
<td>$20_1 20_{20}$</td>
</tr>
</tbody>
</table>
2.1C: Templates for s.e.i. and l.b.c.

Each s.e.i.'s axon field fills 0.001 of the area of the cortical column in which it lies. Therefore, 0.001 of the 1200, or 12, s.i.i. lie within the axon field of a single s.e.i.. The positions of the s.i.i. contacted by a s.e.i. centered at the position marked by the x are given by the following template:

```
...
..x..
...
```

Each large basket cell's axon field fills 0.25 of the area of a cortical column, but only half of its axon field lies in the column in which it originates. Therefore, 1/8 of the 800, or 100, pyramidal cells in a cortical column lie within the axon field of a single l.b.c.. The positions of the pyramidal cells contacted by a l.b.c. centered at the position marked by the x are given by the following template:

```
   ..................
   ............x....
   ..................
   ..................
```
### TABLE 2.2: THE FORM OF THE MATRIX $A_{21}$

The positive connections of the $A_{21}$ are indicated by a `+`.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>75</th>
<th>149</th>
<th>228</th>
<th>297</th>
<th>5600</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+++</td>
<td>+00</td>
<td>000</td>
<td>0000</td>
<td>0000</td>
<td>0</td>
</tr>
<tr>
<td>75</td>
<td>+++</td>
<td>+++</td>
<td>0+0</td>
<td>0000</td>
<td>0000</td>
<td>0</td>
</tr>
<tr>
<td>149</td>
<td>+++</td>
<td>0++</td>
<td>0+0</td>
<td>0000</td>
<td>0000</td>
<td>0</td>
</tr>
<tr>
<td>228</td>
<td>0+0</td>
<td>000</td>
<td>0+0</td>
<td>0000</td>
<td>0000</td>
<td>0</td>
</tr>
<tr>
<td>297</td>
<td>000</td>
<td>000</td>
<td>000</td>
<td>0000</td>
<td>0000</td>
<td>0</td>
</tr>
<tr>
<td>5600</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

35

<table>
<thead>
<tr>
<th></th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>+++</td>
<td>+00</td>
<td>+00</td>
<td>+000</td>
<td>000</td>
<td>0000</td>
</tr>
<tr>
<td>69</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>0+0</td>
<td>000</td>
</tr>
<tr>
<td>103</td>
<td>0+0</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
<td>0+0</td>
<td>000</td>
</tr>
<tr>
<td>137</td>
<td>000</td>
<td>000</td>
<td>000</td>
<td>000</td>
<td>0+0</td>
<td>000</td>
</tr>
<tr>
<td>1200</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(Note that each column shown in this matrix should be repeated four times.)
### Table 2.3: The Form of the Matrix $A_{34}$

The positive connections of the matrix $A_{34}$ are indicated by a +.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>21</th>
<th>41</th>
<th>61</th>
<th>81</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>++++</td>
<td>++++</td>
<td>+0</td>
<td>+0</td>
<td>+0</td>
<td>000</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>+0</td>
<td>+0</td>
<td>+0</td>
<td>+0</td>
<td>+0</td>
<td>+0</td>
</tr>
<tr>
<td>85</td>
<td>000</td>
<td>000</td>
<td>000</td>
<td>000</td>
<td>000</td>
<td>000</td>
</tr>
<tr>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>141</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>169</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>197</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4- From the number of input neuron axon fields that fill a column (see Table 2.1A), find the number of target neurons that a typical input neuron contacts. It takes 100 s.e.i. axons to fill a column, so 1200/100 = 12 s.i.i. are contacted by a typical s.e.i.

5- From the shape of the axonal field of the input neuron described in subsection 2.1.1 draw a template of the positions of the contacted target neurons around the cell body of a given input neuron. This is done for s.e.i. contacting s.i.i. at the top of Table 2.1C.

6- The position of the target cell in the same position in the column as the input cell is found using table 2.1B. The target neuron whose position best corresponds to the position of the given input neuron is selected.

7- The template from step 5 is then centered on the target neuron from step 6 to locate the rest of the target neurons contacted by that particular input neuron. This is done for each input neuron and the non-zero components are entered in the matrix.

For the example of s.e.i. contacting s.i.i. neurons, there are about four times as many input neurons as target neurons. Thus, in matrix \( A_{21} \), every column is repeated four times. So, the s.e.i. at positions \( 1_1, 1_2, 2_1, 2_2 \) all contact exactly the same s.i.i.

The positions of the non-zero components in the matrix \( A_{21} \) are given in Table 2.2. The sequence of steps given above was also used to find the matrix \( A_{34} \) given in Table 2.3, along with the template of a large basket cell axon on the field of pyramidal cells in a column. The
assumption was made in getting the templates of the contacted neurons that the spread of the dendritic field is small compared to the axonal field. This is essentially correct in the case of the pyramidal and l.b.o. (Szentagothai 1978). However, for the interneurons the spread of the dendrites might require the use of slightly larger templates, in which case the number of cells contacted by the axons terminating on the interneuron populations would be greater. Note that the parameters in this model depend on the proportion of each cell type assumed in subsection 2.1.1.

The exact synaptic strengths are functions of past experience, and the current brain state determined by midbrain and association cortex inputs. Thus even if the connections are fairly uniform to begin with, heterogeneity should be introduced by experience. In fact, in a very basic learning model van der Malzberg (1973) showed that orientation columns seemed to evolve among neurons whose orientation sensitivity is random with respect to position, if oriented stimuli were repeatedly shown. In the present neural model there are four parameters that are functions of neural input: the connection strength in the matrices, the resting potential, and the parameters controlling the shape and the position of the conversion function. Each of these could modify the response to an input and thus contribute to learning or filtering.

2.1.3- Data supporting the belief in parameter changes

Parameter changes should manifest themselves as changes in the
sensitivity of neurons to input. In one case the detailed mechanism by which the sensitization of a synapse is carried out is known. Kandel (1979) has shown that the presynaptic termination of a serotonin releasing interneuron can sensitize synapses in the snail. This is due to the activation of a cyclic AMP reaction sequence, which in turn leads to an increase in the calcium current in the axon end foot when an action potential (AP) arrives. The extra calcium entering the axon causes a greater number of synaptic vesicles to release their neurotransmitters into the synaptic cleft. The sensitization caused by this mechanism has been found to last for days. Kandel also found that the opposite of sensitization, that is habituation, can also occur at these synapses. A repeated irrelevant stimulus can cause the calcium current initiated by an AP to decrease, in some cases for days.

In higher animals, there are indications that monoamine neurotransmitters can initiate metabolic reaction sequences rather than simply cause a short increase in the sodium or potassium conductivity through the membrane (Bloom 1979). The specific mechanisms by which these 'second messenger' neurotransmitters work may include altering selected ionic currents passing through the neural membrane. These effects may include altering selected ionic currents passing through the neural membrane over extended periods, and more speculatively, the formation of new synapses (Moilwain 1979).

What are the data supporting the changes in sensitivity of interneurons? Groves (1976) found that habituation to a click was caused by a decreased response of interneurons in the RF. The
presentation rate of these clicks was much too slow to be causing habituation due to fatigue, so some type of sensitization change mediated by a neurotransmitter is implicated. Such a neurotransmitter might work by changing the calcium or potassium current flow through the membrane of relay neurons along the sensory pathway. If sensitized s.i.i. receive input from the faster receptor fibers, such as the Y retinal ganglion fibers, they might be able to wipe out the main reaction to a stimulus.

Singer (1977) has proposed another type of mechanism that might work on the s.i.i. in LGN. He found that the heightened activity of the LGN following cortical stimulation was due to the inhibition of these inhibitory interneurons in a very specific point-to-point fashion. This implies that cortical stimulation could cause a readiness for particular input patterns. Such localized activation of s.i.i. might be used in a directed search for a particular object. Whether the sensitization of interneurons is completely controlled by cortical afferents, or there is some midbrain system involved is not clear. Interestingly, Singer found that there was an increase in extracellular potassium accompanying the inhibition of the s.i.i. and that applying the potassium without any input had the same effect as cortical stimulation.

A very complicated model for a single cortical column has been constructed in equation 2.1. In the following subsection a method of simplifying the model is suggested, whereby a continuous approximation
replaces the connection matrices and the activity vectors.

2.1.4 Decomposing equation 2.1 into a hierarchical model

Equation 2.1 reflects data showing that the two interneuron populations within a column are strongly interactive. Of the nine within-column matrices listed in equation 2.1, all but one of them have one of the interneuron populations as the input. Any processing that the pyramidal cell or large basket cell populations do on a visual input must be done in a single step, since their output leaves the column immediately (or after one synapse in the large basket cell case). In addition, the main part of the visual input seems to reach the s.e.i., as indicated by the huge number of interneurons in layer 4 of area 17, compared to the non-sensory cortex (Shepherd 1974). All of this evidence points to the conclusion that the interneurons are the main processors of within-column information. Note that without dropping any terms equations 1A, 1a, 1B, 1b can be reduced to two equations in two unknowns if the extranuclear inputs are considered constant over a typical interneuron processing period.

If the interneurons are taking care of the within-column processing of visual information, the question as to what the other two populations are doing arises. Equation 2.1 indicates that the pyramidal cells are involved in between-column processing and the l.b.c are associated with pyramidal cells. Essentially all of the connections made by a pyramidal cell lie outside of its column, and these and other relay cells provide
the extra-nuclear input. Equation 2.1 reflects the lack of any direct influence on the interneuron network by the pyramidal cells of the same column. However, an influence is exerted indirectly, since the corticofugal fibers to the LGN have an effect on the strength of the sensory input reaching the interneuron system.

In comparison with the interneuron effects, the relay system effects are slow and their primary influence is on other nuclei, rather than the same column. The relative time scales for within system processing are 10-200 ms for the interneuron system, 100-5,000 ms for the relay system, and seconds on up for the reticular effects. The average distance that an axon connecting interneurons travels is about 100 microns (Scholl 1956), while the average distance between neurons connected by the main axon of a cortical or midbrain relay cells is on the order of a centimeter. The axon terminations of the s.e.i. are in a 30 micron cylinder, those of the relay cells are confined to a 300 micron column while those of the midbrain may contact a diffuse set of neurons in several nuclei.

With these concepts in mind let us look at a two population model of the cortex, bearing in mind that these populations correspond to the two interneuron populations. The interneuron communication times as well as the distances over which their processing takes place are assumed to be very small. This is reasonable, since the primary focus in this study is in perceptual events which last on the order of half a second (Kahneman 1973) and are probably tied to among column processing in the whole visual system. On the other hand, the processing times and the
distances involved with the midbrain projection systems would be considered large, if they were included in a neural model.

2.2- THE INTERNEURON NETWORK

2.2.1- Deriving the time coarse-grained model

The short distances and the fast processing times allow one to approximate the discrete interneuron model embedded in equation 2.1 with a continuous model. First, recall that the interneuron model extracted from equation 2.1 consists of 2.1A, 2.1B, and 2.1a, 2.1b. The activity vectors in equation 2.1 (for example $P(t)$) must be converted to an activity distribution, $P(x,y,t)$. In addition, the connectivity matrices components, such as $A_{ij}(kl)$ (where the stimulus goes from cell $i$ of cell type $k$ to cell $j$ of cell type $l$), must be converted into connectivity functions, $a_{kl} w_{kl}(x-x', y-y')$, that depend only on the distance. The amplitude of the connection strength between cell types is given by $a_{kl}$, while $w_{kl}$ is the distribution of connection strengths. The $ij$ subscript is dropped, since the specific connections between pairs of cells is no longer needed. The summation over the column, performed by the matrix multiplying the input vector, must be done by an integration over the area of the column, in the continuous case. Note that the $G$ function that converts potentials on a neuron into firing rates makes it possible to write the equation in either set of variables.
EQUATION 2.2: A CONTINUOUS APPROXIMATION FOR THE INTERNEURON NETWORK

\[
S(x,y,t) = S_o(x,y,I) = G_S\left[ \int_0^t \left[ \exp\left(\xi - t/\mu_{ex}\right) \right] \right.
\begin{align*}
&\sum a_{ee}(I)w_{ee}(x-x',y-y')S(x',y',\xi)dx'dy' \\
&- \exp(\xi - t/\mu_{in}) \sum a_{ei}(I)w_{ei}(x-x',y-y') \\
&\left. \right] dx'dy' + Z_S)d\xi \right], I
\end{align*}
\]

\[
B(x,y,t) = B_o(x,y,I) = G_B\left[ \int_0^t \left[ \exp(\xi - t/\mu_{ex}) \right] \right.
\begin{align*}
&\sum a_{ie}(I)w_{ie}(x-x',y-y')S(x',y',\xi)dx'dy' \\
&- \exp(\xi - t/\mu_{in}) \sum a_{ii}(I)w_{ii}(x-x',y-y') \\
&\left. \right] dx'dy' + Z_B)d\xi \right], I
\end{align*}
\]

The time delays were ignored in this equation under the assumption that
the sum of the time taken for axonal conduction, synaptic crossing, and
PSP rise are fairly small compared to the decay times of potentials.
This assumption, while often made, is not really justified, since as
noted above, the rise time of 20 ms for IPSP is about the same as the
time for the decay of EPSP. Since only the very gross behavior of this
model is of interest, it is hoped that leaving out the time delay will
leave the basic behavior intact. However, treating the interneuron
network with the \( \tau_1 \) time delay left in is an interesting project for
further modeling.

At this point a particular sigmoidal function is selected for use as
the \( G \) function. The logistic equation, zeroed for a zero input, is
taken as was done by Wilson and Cowan (1973) (see Figure 2.8). The
conversion function for potential at the axon hillock into an average
firing rate is:

\[ G_S[\alpha(x,y,t), I] = \left\{1 + \exp[v_s(I)[\alpha(x,y,t) - O_S(I)]]\right\}^{-1} \]

(2.3)

- \left\{1 - \exp[v_s(I)O_S(I)]\right\}^{-1}

The output is the actual firing rate minus the no sensory input firing rate, normalized so the maximum value of \( G \) is one. The maximum firing rate and the resting firing rate are functions of the midbrain input to area 17. (Changing the sensitivity of relay cells will be discussed in subsection 2.2 in connection with the model of between-column information processing.) The two parameters in the logistic equation have the following meaning: the \( O_j \) is the position of maximum slope of the logistic curve, and \( v_j/\mu \) is the slope of the curve at that position (Wilson & Cowan 1973). Both of these parameters are functions of the midbrain input.

Following a procedure given in Wilson and Cowan (1972) equation 2.2 can be time coarse-grained by a change of variables.

\[ \hat{S}(x, y, t) = 1/\mu_{ex} \int_{-t}^{t} S(x, y, \xi) \exp(\xi - t/\mu_{ex}) d\xi \]

\[ \hat{B}(x, y, t) = 1/\mu_{in} \int_{-t}^{t} B(x, y, \xi) \exp(\xi - t/\mu_{in}) d\xi \]

This combined with the equation \( S(x, y, t) = \mu \hat{S}(x, y, t)/\Delta t + \hat{S}(x, y, t) \) allows equation 2.2 to be written in terms of \( S \) and \( B \). These new variables comprise the influence that either an s.e.i or a s.i.i. exerts
FIGURE 2.8: THE SIGMOIDAL FUNCTION

The shape of the function $G_j$, where $j$ is $S$ or $B$, shows the change in the momentary firing rate as the axon hillock potential increases. The maximum slope of the function occurs at the axon hillock potential $O_j$, and takes the value $v_j/4$. 
on the cells it contacts at time $t$, due to its influence since the beginning of the neural event. The time coarse-grained equations are:

**EQUATION 2.4: THE TIME COARSE GRAINED INTERNEURON EQUATIONS**

\[
\frac{\partial \hat{S}(x,y,t)}{\partial t} - \hat{S}_o(I) = -\hat{S}(x,y,t) + G_S(I) \tag{2.4A}
\]

\[
\left[ \mu_{ex\,ee}(I) p_{ee}(x,y) \hat{S}(x,y,t) - \mu_{in\,ei}(x,y) \hat{B}(x,y,t) + Z_S \right]
\]

\[
\frac{\partial \hat{B}(x,y,t)}{\partial t} - \hat{B}_o(I) = -\hat{B}(x,y,t) + G_B(I) \tag{2.4B}
\]

\[
\left[ \mu_{ex\,ie}(I) p_{ie}(x,y) \hat{S}(x,y,t) - \mu_{in\,ii}(x,y) \hat{B}(x,y,t) + Z_B \right]
\]

where $w_{ee}(x,y) \hat{S}(x,y,t) = \int \int w_{ee}(x-x',y-y') \hat{S}(x',y',t) \, dx' \, dy'$

The similarity of equation 2.4 to a set of equations derived by Wilson and Cowan, which are given in equation 2.5, starting from different physiological assumptions is striking.

**EQUATION 2.5: THE WILSON AND COWAN EQUATIONS**

\[
\frac{\partial \hat{S}(x,t)}{\partial t} = -\hat{S}(x,t) \tag{2.5A}
\]

\[
\mu s[1-r_s \hat{S}(x,t)]
\]

\[
G_s[\alpha \mu \hat{S}(x,t) - \phi_{ee}(x) \hat{S}(x,t) - \phi_{ei}(x) \hat{B}(x,t) + Z_S(x,t)]
\]

\[
\frac{\partial \hat{B}(x,t)}{\partial t} = -\hat{B}(x,t) \tag{2.5B}
\]

\[
\mu s[1-r_s \hat{B}(x,t)]
\]

\[
G_B[\alpha \mu \hat{S}(x,t) - \phi_{ie}(x) \hat{S}(x,t) - \phi_{ii}(x) \hat{B}(x,t) + Z_B(x,t)]
\]

Note that their variables are proportions of a cell type at position $(x,y)$ becoming active per unit time at time $t$, which is essentially a normalized average firing rate. ($r_j, \alpha, \text{and} \phi_j$ are scale constants.)
Before turning to the dynamics of equation 2.4, some comments about its differences from the Wilson and Cowan coarse-grained set of equations are in order. Wilson included the \((1 - r_E)\) term to adjust for the refractory period, which could keep a neuron from firing in spite of sufficient stimulation. However, little is lost by absorbing this term into the \(G\) function, which is already attenuating the slope of the firing rate at high stimulation levels. A second difference is that Wilson assumes that \(\mu_\text{ex} = \mu_\text{in}\) although they differ by a factor of ten in much of the CNS. However, the difference that this produces between the two time coarse-grained equations can readily be compensated for by adjusting parameters. In particular, decreasing the slope of the inhibitory \(G\) function, \(v_1\), decreasing the position of the maximum slope, \(c_1\), or decreasing the connection strengths, \(a_\text{ie}\) and \(a_\text{ei}\), could compensate for the difference in the inhibitory decay length. The fact that they sum from minus infinity is equivalent to saying that the decay time of the PSPs is short compared to the length of the neural event. Thus, despite the differences, the qualitative response of the models in equations 2.4 and 2.5 should be almost identical. The agreement between the Wilson and Cowan model and the more complex model presented here reflects similar anatomic assumptions underlying their model and this model of the interneuron network.

2.2.2- Dynamics of the within-column interneuron network

2- An important factor in the generality of the interneuron model is that it does not have to be limited to a single column. If the inhibitory neurons on which the collaterals of the pyramidal cells terminate are left out of the s.i.i. population, then the relay system in that nucleus has no effect on the information processing in the interneuron network. If, in addition, the cortical afferents from other
The dynamics of the interneuron network can be studied as its two distinctly different types of extranuclear inputs are varied. If the strength of the LGN and the visual cortex inputs to area 17 are varied, the firing rate of localized groups of cells is changed. On the other hand, if the input from the midbrain and association cortex is altered, the sensitivity of a whole cell type may be changed. As a result, the parameters of the model must be allowed to vary. This could cause a totally different response mode, with or without changing its spatial distribution of electrical activity, in response to the same stimulus. In the elementary modeling of the visual system done here, knowing the mode of response is more useful than knowing the response to complex spatial stimuli. Thus, the LGN inputs examined initially are simple spots varied in size, duration, and intensity without considering the particular information carried.

In line with the data presented in subsection 2.1.3, the inputs generated in the midbrain and association cortex are assumed capable of altering four of the six parameters for each cell population in equation 2.4. The parameters of connection strength, $a_{sl}$, the potential at the maximum slope of $G$, $v_g$, the maximum slope of $G$, $v_s/4$, the no input firing rate, $S_0$, are functions of the input. The distance constant and the time constant of the neurons are structurally fixed. Changes in these four parameters can change the sensitivity of a cell type (uniformly, over columns or larger areas). The number of AP fired in parts of the cortex terminate on the pyramidal and large basket cells, then the interneurons do not experience the columns defined by these axons. When these assumptions are made, the within-column processing to be discussed can be extended to the entire nucleus.
response to a momentary axon hillock potential are affected by changes in any of the parameters, while the size of axon hillock potential elicited by an AP is controlled by the connection strength, $a_{ij}$.

Wilson and Cowan (1973) simulated the gross qualitative response of equation 2.5 as the LGN inputs and the parameters were varied. The similarity between the two equations allows their analysis to be applied to equation 2.4. Two things should be added to their discussion: a) a description of the events underlying the modes of behavior which they found, b) a discussion of how the response properties of interneurons in area 17, such as orientation and disparity, might arise from the model. To carry out this second objective some asymmetries in the axon and dendrite distributions in the different cell types must be incorporated into the neural model.

There are two types of activity that are of special interest in terms of explaining the functioning of area 17. The first is self-sustaining activity, or at least brief periods of positive feedback after the stimulus ends. Such activity is the prime candidate for preserving the initial analysis of a stimulus. One may consider many candidates for information transfer between two neurons. However, communication between two cell type populations must make use of the summed effects of AP from many different neurons. Variations in the temporal rate of individual neurons should be overpowered by other strongly firing neurons contacting the same targets.
One of the speculative assumptions that were made in writing down equation 2.1 was the existence of interneuron connections which have not yet been demonstrated anatomically. Without these connections any type of self-sustaining activity within a column is impossible. The only other candidate for the precise storage of information such as is seen in short term memory is sensitization and habituation of the neurons which have just received a stimulus. One type of behavior seen in the CNS that supports self-sustaining activity, rather than sensitization, is hysteresis. This type of behavior is seen to emerge naturally from the neural model, but it is hard to imagine how it might arise during sensitization.

The other type of processing in the interneuron network that one would wish to find is an explanation response variability. In other words, why should s.e.i. in the same perpendicular electrode path respond to different spatial frequencies or colors. The model of the interneuron network given in equation 2.4 is too general to answer this question, except perhaps in the case of orientation columns. This problem is examined in subsection 2.2.4.

2.2.3- Self-sustaining activity and related activity modes

A number of types of self-sustaining activity are possible for the interneuron network, and it is possible to move the model between these modes by simple parameter changes. Wilson and Cowan (1973) simulated equation 2.5 for several different sets of parameters and found several qualitatively different behavior modes. It is useful to add a
qualitative discussion of how these modes develop and how the parameters must be changed in going between modes.

The anatomical shapes of the interneuron axonal fields given in subsection 2.1.1 play a role in each of the modes of activity that are considered. Recall that the axonal fields of the s.e.i. are long and narrow compared to the s.i.i. axonal fields. Therefore, the inhibitory length constant is greater than the excitatory length constant \( \gamma > \lambda \), while the excitatory connection strengths are greater in the center of their fields that the inhibitory connection strengths (see Figure 2.9). The longer range of the s.i.i. axons is responsible for localizing the region responding to a spot of light. It is also responsible for edge enhancement, i.e. for a large circle of input the greatest activity is at the edges and the lowest in the center. Edge enhancement is reminiscent of visual system responses, such as Mach bands (see Figure 2.12). An excitatory neuron in the center of a spot of stimulation will receive significant effects from inhibitory neurons within a larger radius than from excitatory neurons. At the edge of the area stimulated by an input, there is no inhibitory contribution from the effective inhibitory cells that border the stimulated area. The net effect of the greater inhibitory summation in the center is that activity is depressed there.

The dynamics of equation 2.4 We wish to consider the dynamics of the model in response to simple inputs as the parameters are varied. Those that will be varied are the strength, \( a_{ij}; \) the level of input at which
FIGURE 2.9: THE AREA OF AXON INFLUENCE

The connections strength and their distribution around the cell body for the two types of interneurons. The amplitude of the of the s.e.i. input to population $j$, $a_{exj}$, is greater than the amplitude of the s.i.i. input to the same population, $a_{inj}$. The distribution of connection strength for the s.e.i. is assumed to be the negative exponential, $\exp(-|X-X'|/\sigma_{ex})$, while that of the s.i.i. is assumed to be, $\exp(-|X-X'|/\sigma_{in})$. The distance from the center of the cell body at which the connection strength of an s.i.i. falls to 1/e of the amplitude is greater than the 1/e distance for an s.e.i.
maximum slope of the firing rate occurs, $Q_e$, and the maximum slope, $v_e/4$.

Hysteresis is a phenomenon in which there is a sudden jump in the activity level at some point as the input is increased, and the elevated activity continues even if the input level is dropped below its level at the jump (see Figure 2.10). In hysteresis, the point at which the maximum rate of increase of firing as stimulation increases (i.e. the maximum slope of firing vs. stimulation) must occur at a greater value for the inhibitory cells than the excitatory cells. This allows a quicker climb of the firing rate of excitatory cells in response to a stimulus than of the inhibitory population. As the strength of the input to the region is increased, the firing rate of the excitatory neurons slowly rises till the self-excitation overcomes the decay and inhibition terms and reaches the positive feedback level. This causes the jump at $b$, with the upper value of the firing rate at $b$ determined by when the inhibitory stimulation becomes strong enough to overcome the positive feedback. As the stimulus level is decreased along the upper curve a high level of self-stimulation is maintained past point $b$, all the way down to point $a^2$, where the natural rate of decay plus the inhibition overcome the self-stimulation.

There is some recent physiological support for the idea of hysteresis in the interneuron network during sensitization. Singer (1979) found that simple cells in area 17, which are probably spiny stellate cells according to Kelly & Essen (1974), showed a hysteresis effect in
response to a midbrain cholinergic stimulus, that did not show without it. Cholinergic input is known to increase the strength of the response of cortical neurons (Singer 1979). So, hysteresis may depend on sufficiently high $a_{ee}$ and $a_{ex}$ sensitivity.

When an impulse, instead of a constant input, is applied to the hysteresis mode, it is seen that this mode is the same as what Wilson and Cowan call an active transient mode. An input greater than $b$ can cause a huge firing rate increase due to excitatory self-stimulation. Once within the positive feedback portion of the response the firing rate can continue to rise even if the stimulation is cut off, until it is stopped by rise of activity in the s.i.i. population. Figure 2.11 shows the response of the model to three intensities of input. In the lowest intensity the positive feedback threshold for the s.e.i. population is not reached, in the middle intensity it is reached after the stimulus ends, while in the highest intensity it is reached during the stimulation. Note that the threshold of positive feedback can be reached by varying the intensity (as in Figure 2.11), the duration, or the spatial extent of the input stimulus.

When the excitatory-excitatory connection strengths, $a_{ee}$, are increased, along with the inhibitory-inhibitory connection strengths, $a_{ii}$, the high response can be maintained without an input (see Figure 2.12). This is just the sort of effect that a strong cholinergic input, mentioned above, might cause. So, this sensitization could switch a interneuron network from the active transient to what Wilson and Cowan call the spatially inhomogeneous steady state. These parameter changes
FIGURE 2.10: HYSTERESIS
See discussion page 67
(From Wilson & Cowan 1973)

FIGURE 2.11: THE ACTIVE TRANSIENT MODE
See discussion page 68
(From Wilson & Cowan 1973)

FIGURE 2.12: A STEADY STATE SHOWING EDGE ENHANCEMENT
See discussion page 68
(From Wilson & Cowan 1973)

FIGURE 2.13: A LOCALIZED LIMIT CYCLE
See discussion page 70
(From Wilson & Cowan 1973)
allow the level of firing raised by an input at points in the region to be self-maintaining. In other words, in Figure 2.10 the value of the stimulus at which the drop occurs, $a$, has a negative value.

There is a variation of the hysteresis mode that occurs with constant input, called a localized limit cycle. The main parameter change needed to get to this mode from the hysteresis mode is an increased inhibitory to excitatory connection strength and a decreased inhibitory-inhibitory connection strength. In this mode, both interneuron populations oscillate with the same period and about 180 degrees out of phase (see Figure 2.13). Note that this is basically a sensitization of the s.i.i. If the inhibitory population rises more slowly, a wave of excitation in the s.e.i. population, traveling outward from the point of stimulation can form.

Before leaving the interneuron model, a comment is made about the visual experience that might accompany spatially inhomogeneous steady states in area 17. Emmentrout and Cowan (1979) found that activity in the shape of grating patterns would result as the connection strength of the s.e.i. to itself is raised to the positive feedback threshold, under certain assumptions of rigid motion invariance. Using the geometrical transformation of the retinal activity pattern in traveling to area 17 (Cowan 1979), it was shown that these cortical grating patterns corresponded to certain spiral, funnel, or lattice patterns in the retina. This is interesting because these are characteristic
hallucination figures seen by people using psychedelic drugs.

2.2.4- The formation of quality differences among cells of the same column

In subsection 2.1.1, it was stated that there are about three orders of magnitude more interneurons in area 17 than there are LGN inputs. It was suggested that one reason for this profusion of interneurons is the separation qualities. The LGN input to area 17 seems undifferentiated with respect to such input qualities as the orientation of lines, disparity, spatial frequency, direction and speed of motion. These qualities plus the ones to which LGN neurons seem to be sensitive, such as position, color, and ocular dominance, are jumbled together in each area 17 neuron. The individual qualities seem to be separated out of the visual input in areas 18 and 19, where Zeki (1974) has found retinotopic maps that respond mainly to movement, color, or disparity. It is proposed that these qualities are separated by the convergence of area 17 afferents on areas 18, 19 and the suprasylvian sulcus. This is not an interneuron based effect, so it is considered in the discussion of the relay system.

Looking at Figures 2.4 and 2.5, one can see that a variety of axon and dendrite distributions occur among the interneurons of area 17. In the neural model, it was assumed that all the s.i.i. had the same shape and stretched across all six layers of area 17. This allowed the use of a two dimensional model of the cortex. However, psychological data indicates that many of the s.i.i. are localized at a specific depth.
What happens when the model is expanded by restricting the s.i.i. axons and all of the interneuron dendrites to a small fraction of the cortical thickness? It is proposed that the emergence of selective response to qualities in area 17, that are not reacted to earlier in the visual system, is a result of asymmetric axon distributions, mostly in the s.i.i. population.

The first proposed example of the formation of a fine-tuned response to a quality in area 17 neurons due to interneuron processing is that of orientation columns. Without altering the model, consider the processing that might take place with the one asymmetric neuron included in the model, the large basket cell. Recall that less than ten LGN axons end on each cortical column in area 17, so a column would have about three axons across the diameter. The orientation columns change every 50 microns in crossing the column, and thus are finer-grain than the input. This implies that orientation information is not contained in the input of any single LGN input. However, the precise pattern of firing of the s.e.i. varies, as a thin line centered on the visual field of one column is rotated. Since different s.e.i. neurons in the column receive different combinations of LGN input, a particular orientation might initially trigger a particular set of these interneurons. The most strongly responding s.e.i. might then form a thin band through the column, and excite the pyramidal cells that they contact (see Figure 2.14). However, one must look to the within-column pattern of inhibition in order to explain the fading of response as the orientation is varied.
It is suggested that orientation sensitivity results from the anatomic configuration of large basket cells, combined with the spatial transform of a retinal distributions of activity in going to area 17. Schwartz (1977) pointed out that a line in the retina becomes a curve resembling a spiral in area 17. As a result, he proposed that there is a logarithmic transformation of a shape in traveling between these two structures. Cowan (1979) expanded on this idea and claimed the actual transformation was a partial log polar transformation. The large basket cell bodies have a roughly uniform, random distribution with respect to the surface of area 17. Their axon fields are arranged like a stack of chips, with the long direction pointing the same way. Therefore, there is a great overlap of the axonal fields, parallel to the long dimension of the axonal field and only minor overlap in the perpendicular direction (see Figure 2.14). Figure 2.14 also shows the region of strongest s.e.i. response to a line input. The figure shows that if the line of stimulation runs parallel to the long dimension of the large basket cells, the axonal fields of those large basket cells excited will overlap extensively. Any pyramidal cells lying in this zone will be strongly inhibited. On the other hand, if the strongly activated stimulation at a point is in the perpendicular direction, the axonal fields if the large basket cells will not overlap greatly and so the local pyramidal cells will not be strongly inhibited. In this case, the excitation from the interneuron network would overpower the inhibition from the few large basket cells and the local pyramidal cells would become excited. The polar nature of the transform would cause a change of orientation in the retina to be transformed into a small translation of the response pattern in area 17. Thus, it could account for the
Long dimension of l.b.c. axonal fields

Cell bodies of l.b.c.

Perpendicular locus of stimulation

Parallel locus of stimulation

Locus of maximum stimulation from a line

FIGURE 2.14: DIFFERENT LEVEL OF PYRAMIDAL CELL INHIBITION DUE TO THE ASYMMETRY OF LARGE BASKET CELLS

(The number of lines per box indicate the number of large basket cells(l.b.c.) firing strongly in a box.) See discussion page 73
steady progression of orientation columns in moving across area 17, as a line on the retina is rotated. Therefore, the shape of the large basket cells and the logarithmic nature of the transform lead to orientation sensitivity of pyramidal cells at certain points in area 17. In support of this idea is the observation that layer 4 interneurons do not seem to show orientation sensitivity, and probably do not receive large basket cell input (Hubel & Wiesel 1977).

In order to extend this proposal to include spatial-frequency response, that is constant in horizontal layers, would require adding a s.i.i., with a horizontal axon field, to the model. Other qualities also may be fine-tuned by particular subtypes of s.i.i.. If the quality tuning of area 17 neurons results from interneuron network processing, then the interaction of these qualities may lead to surprising results. If a cell responded to the color red and to a particular spatial frequency, then using a red grating for a stimulus, instead of a black one, might not increase its activity. This would depend on the interaction of the s.i.i. responsible for tuning these two qualities, according to the proposed model.

2.3- THE MODEL OF COLUMNS WITHIN A PATHWAY

In the literature there are few models with any relevance to the relay system. All of the models examined lack either between nuclei interactions or feedback loops, and therefore cannot really model the qualitative behavior of the relay system. However, these models do show some interesting capabilities for neural networks. Grossberg (1978)
presented a model which might be relevant to modeling columns in a single nucleus, which included lateral inhibition and self-stimulation. The model given by equation 2.1 has self-stimulation only in the interneuron network, but there should be qualitative similarities between the two models for early signal processing. Anderson, Silverstein, Ritz, and Jones (1978) looked at a small set of neurons with excitatory and inhibitory effects on a second small set of neurons through a matrix of connection strengths. They pointed out that by considering the input as a vector, orthogonal eigenvectors of the matrix would activate discriminable patterns in the same target set of neurons.

2.3.1- Relay cell axonal distributions

In examining the cell types in the previous subsection, it became apparent that the system of relay and large inhibitory cells operated in a very different fashion than the interneuron system. The hierarchical nature of the model given by equation 2.1 allowed the interneuron network to be isolated from the rest of the model and analyzed separately. The relay and large inhibitory cell equations cannot be separated as nicely from the rest of the model, but by assuming a particular mode of interneuron processing, a qualitative examination of the relay system can now be done.

The first thing of importance about this relay system is the concept of a column structure as defined by the cortical afferents to a nucleus. When the main axon of a cortical pyramidal cell goes to another cortical
area, it enters at layer 6 and forms a cylinder 300 microns in diameter that goes up to layer 1 and branches out (see Figure 2.6).\(^3\) Szentagothai (1978), and others using tracer techniques, have shown that cortical afferents from all over the brain will tend to end in-register on a column. Thus each area of the cortex is subdivided into a mosaic of 300 micron columns, with all the pyramidal cells in a column being contacted by the same cortical afferent inputs. This implies that all 800 pyramidal cells in a column can be thought of as comprising a single receptor for input coming from other columns.

The neocortex receives thalamic sensory input, in addition to cortical afferents. The thalamic axons from the LGN terminate largely as a very flat plexus in layer 4c of area 17, and unlike the cortical afferents they do not seem to be in-register (Szentagothai & Arbib 1975). These sensory inputs lead to a functional division of area 17 which have also been called columns, in spite of a slab-like appearance. The best known example of sensory response columns occurs in the visual system. Hubel and Wiesel (1961) found that the optimal retinal stimulus for area 17 cells of cat and monkey is a line. There is a regular progression of line orientations that elicit a response as an electrode is moved parallel to the surface. They also found periodic changes in the eye which had to be stimulated in order to get a response as an

---

\(^3\) There is some evidence that while the cortical afferent axons are spread uniformly with respect to the surface, they are not uniform with respect to depth in the column (Schibel & Schibel 1970). Since the apical dendrites of pyramidal cells run perpendicular to the surface, the non-uniformity of the axon spread in depth would not greatly alter the total number of connections between the incoming axon and the apical dendrite. However, the case might be different for basal dendrites, which run almost parallel to the surface.
electrode is moved across area 17. The neural groupings with these specialized responses were called orientation and ocular dominance columns respectively. Each orientation column is about 50 microns in diameter and differs by about 15 degrees from the previous one. In subsection 2.2.4 it was proposed that these columns may result from large basket cell asymmetry. The ocular dominance columns are about half a millimeter in diameter, and they overlap about half way into the visual field of the previous column so each eye covers the whole field (Hubel & Wiesel 1977). The LGN is divided into layers that respond only to one eye, so the ocular dominance columns seem to be determined by the anatomic control of LGN axon termination.

Mountcastle (1979) has catalogued other examples of sensory defined columns in auditory, touch, motor, and parietal association cortices. In the primary auditory cortex, there are columns for both pitch and interaural intensity difference. The motor cortex has columns for both body position and muscle groups that might participate in some movement. The parietal cortex was shown to have functional columns based on types of visually guided behavior and eye movements. These last two cortical zones, not being of the primary sensory type, might reflect the cylindrical nature of the cortical afferents a little more closely. Note that there is nothing to limit the number of information channels with spatial separation in a given cortical area. For example, in area 17 there have been claims for disparity columns (Nelson 1975) and color columns. Maffei (1979) claimed that area 17 neurons sensitive to a
single spatial frequency occurred at the same depth.

2.3.2- The basic connections of the relay system model

Equation 2.1C shows three distinct contributions to the activity of relay cells in a column: the interneuron network; other columns including subcortical input; and diffuse inputs such as that from the midbrain. The first contribution is from the interneuron network. A fairly intricate spatial pattern will form on this network in a single column, despite there being less than ten LGN inputs to the column. This is due to the processing of the interneuron network, as described by equation 2.4. An even more fine-grained response for the interneuron system is obtained when the asymmetries of the s.i.i. population are incorporated in the model to give quality fine tuning as discussed in subsection 2.2.4. The interneuron network is considered in the localized steady state mode in this section, so that its contribution to the relay system is constant over time. In the case of multiple synapses (such as those the spiny stellate cells make with the pyramidal cells) the firing of a single neuron can cause a large change in the firing rate of a pyramidal cell. Furthermore, the fine-grained patterns of activity that form within the interneuron system in a column are transferred to the pyramidal cells in that column, because the narrow apical dendrite of the pyramidal cell allows it to sample a small perpendicular piece of the column. In this preliminary look at the relay system, the large basket cells will be considered just another type of inhibitory interneuron. Their special properties are discussed in subsections 2.2.4 and 2.3.4.
The second contribution is from other columns, and it has two distinct components. There are the collateral axons which the pyramidal cells send to the surrounding columns that have a net inhibitory effect, and the main axons coming from pyramids in distant regions with an excitatory effect. In addition to the cortical afferents there is probably a small, fast input from the Y fibers of area 17 (see subsection 2.1.1). These, and other subcortical inputs ending directly on the pyramidal cells of visually responding cerebral cortex, will be included as columns in the model, even though they may not have the in-register property. (In fact, the degree to which in-register columns are needed is something that could be decided by calculations with the relay system model.) The cortical afferents to subcortical nuclei may define columns of a sort in these nuclei as well as in the rest of the cortex (Graybiel 1975). It is not clear from the available data whether anything resembling a columnar arrangement exists among the relay cells of subcortical nuclei that would allow treating the 800 pyramidal cells in a column as a single receptor unit. When the pyramidal input into a column is strong enough to get a response, all the pyramidal cells in this column would respond in unison, subtracting out the interneuron contribution. This does not imply a uniform response, due to different size and depth of pyramidal cells within a column. The pyramidal cells in different layers would have different length apical dendrites and therefore should make a different number of synapses with a given cortical afferent. This nonuniform response to the input from another column will be neglected in this survey.
Each of the pyramidal cells in the column has its own axon and sends its output to a particular nucleus. It is known that the neurons in layer 5 and 6 send much of their output to subcortical targets, such as the SC and LGN, while those in the upper laminas send their output to other cortical areas (Schiebel 1970). In addition, all layers send axons to their mirror image location in the opposite hemisphere (Szentagothai 1978). Thus, each of these functional columns reach about 500 of the 5 million columns in the cortex. Since each column sends out about 500 (as a minimal estimate) axons to other cortical areas, each also receives about 500. The thalamic input ending on the pyramidal cells will be lumped in with them, even though thalamic inputs are probably not in register with the columns defined by cortical afferents. Each of the five hundred main axons leaving a given column potentially has the five million columns in the cortex as targets. However, the model should acknowledge the limits placed on the connections by the organization of the projections between the nuclei. Essentially all of the nuclei reached by main axons from area 17 are retinotopically organized, and these axons end on columns that are sensitive to the same part of the visual field as they are (Singer 1977, Zeki 1974).

The third contribution to the columns comes from midbrain and limbic system inputs which seem to be almost uniform over an entire nucleus, rather than just a single column. The monoamine inputs appear to be ineffectual in elevating the firing rate of cortical pyramidal cells when presented alone (Cooper 1978). However, the second messenger
effects produced by these inputs may cause long lasting changes in the sensitivity of these neurons to sensory inputs.

Equation 2.6 combines the points from the above discussion into a model of the interaction of columns in the visual system via the connections provided by the relay cells. (Note that the feedback to the interneuron network from the relay system is ignored in the following model.)

\[ X^c(z,t) = X^c(z,I) = G_p(I) \]
\[ \bigg[ \sum_{n} X^0(z,\xi) A_{zz'}(cn)(I) \exp(\xi-t/\mu_{ex}) \bigg] d\xi + \text{IN}^c(x_z,y_z) + M^c \]

The activity of a column at position \( z \) in nucleus \( c \) and at time \( t \) is given by \( X^c(z,t) \). The contribution of the other columns in the visual system to activity in this column is given by the time integral on the right. The three factors in the integral are the activity in the other columns at time \( t \), multiplied by a connection strength \( A_{zz'}(cn) \) from column \( z' \) in nucleus \( n \), to column \( z \) in nucleus \( c \), multiplied by a rate of decay. The interneuron contribution is given by \( \text{IN}^c(x_z,y_z) \), where \( (x_z,y_z) \) is a position within the column at position \( z \). There is no time dependence, since the interneuron network is assumed to be in a steady state. Finally, the midbrain contribution to all the columns in nucleus \( c \) is assumed to be \( M^c \).
The diffuse and much slower acting (partly due to the long duration of the second messenger effects of the monoamines) influence of the brain stem projection systems on the relay processing may, in turn, be the basis of the next higher functional level. The next higher functional level. As suggested by Kilmer, McCulloch, and Blum (1969) in their model of the reticular formation, this system may play a role in determining behavior modes.

2.3.3- Data-based restrictions on relay system processing

A strong visual input will give rise to a localized steady state in the interneuron network of area 17, given the right type of sensitization, according to the model presented above. The localized steady state in each column should raise a subset of the pyramidal cells to an elevated firing rate. Therefore, for each column some subset of the 800 columns reached by its output will initially be stimulated due to the sensory input. In this way, the pattern of activity in the localized steady state in the columns of area 17 can lead to an initial firing pattern in the relay system involving many nuclei. After this, processing in the relay system itself takes over. Several basic processing modes for the relay system can be proposed by combining the following three anatomic properties of connections in the relay system model. 1) lateral inhibition of adjacent columns due to pyramidal collaterals, 2) convergence on many axons onto each column, 3) feedback loops among the nuclei.
The lateral inhibition between-columns supplied by the pyramidal cell collaterals forces the excitatory effects of one column on another to travel mainly in the between-nuclei direction. A strong stimulation at one spot leads to a depression of activity in the area surrounding the spot. Thus, only isolated columns in a single nucleus should be able to maintain high activity unless the input to adjacent columns is very strong. On the other hand, the effects of one column traveling over the main axons of relay cells on other columns should be mainly excitatory.

The need for convergence of axons onto a column may also constrain the processing of the relay system. The anatomy of the pyramidal cells raises a question as to the influence that one such neuron can exert on another. The tendency of cortical afferents to end high on the apical dendrite makes their effectiveness as a mediator of the firing rate questionable. The data indicate that the pyramidal cells receive a massive number of both excitatory and inhibitory synapses. Only the massive synapses of large inhibitory cells are likely to have a large individual effect on the axon hillock potential. Therefore, it is postulated that a number of strongly firing pyramidal cells from other columns must converge on a single column to change the firing rate of its pyramidal cells. The dynamics of the relay system depends on the exact fraction of the 500 or so relay cell inputs contacting pyramidal cells in a column that must be firing strongly in order to raise the average firing rate in the column. A rough estimate of the required column convergence based on the ability of small focal stimulation to activate portions of other nuclei is five to ten strongly firing relay cells.
feedback. What is actually required is a sequence of convergences (see Figure 2.15). A set of columns in which self-sustaining activity can exist is defined as an active circuit. Every column in an active circuit must receive some minimum number of axons from other columns in the active circuit which, should depend on the level of relay cell sensitivity (see discussion in subsection 2.3.4). For example, if there are seven columns in an active circuit and there must be at least four connections to other columns, then Figure 2.15 shows one possible active circuit. Due to lateral inhibition, an active circuit would normally be spread diffusely through many nuclei. Convergence places an additional constraint on the nature of the activity in the active circuit when it goes into a self-sustaining mode. The turn-on of self-sustaining activity in an active circuit requires a high degree of synchronization among the axons coming into each column in the active circuit. Basically, this constrains the columns of the active circuit to have the same temporal variations up to a time delay for traveling between nuclei. In the case where the active circuit is turned-on by a sensory input, the interneuron network would determine the particular active circuit turned-on through the set of columns originally activated. Due to the synchrony, the temporal variation would depend largely on the set of columns participating in the turned-on active circuit.

In a few special cases where there is probably a high level of sensitization of the relay cells in general, activity resembling what would be expected of an active circuit appears. The large components of evoked potentials (EP) appear several hundred milliseconds after the
The need for the convergence of a number of axons into a column to change its average firing rate significantly may be the basis of separating particular visual input qualities that are jumbled together in area 17 response. Suppose a number of area 17 pyramidal cell axons with the same disparity response, but random with respect to other qualities, converged on the same area 18 column. That column would respond to disparity, but would ignore the other qualities like noise. Thus convergence may allow selection of particular qualities that are formed by the processing of the area 17 interneuron network. In the inferotemporal cortex (IT), convergence could result in columns that select for highly specific combinations of qualities. In each case, the columnar response to other columns is modulated by the fine-grained response of the interneuron network within that column, whose main input may come from the posterior thalamus.

The third anatomic property that is important to the relay system model is the existence of closed loops of relay cells involving two to five synapses. The inference that these loops exist is based on data showing reciprocal connections between most cortical areas and the thalamic nucleus that supplies their input (Singer 1979). There are also reciprocal connections between most cortical areas (Zeki 1974).

When these three anatomic properties are incorporated into the relay system model the possibility of an interesting mode of relay system activity arises. The need for convergence makes the existence of simple closed pools insufficient for generating self-sustaining positive
FIGURE 2.15: AN ACTIVE CIRCUIT COMPOSED OF SEVEN COLUMNS AND REQUIRING FOUR CONNECTIONS PER COLUMN

The number of strongly firing columns from the same active circuit needed to achieve self-sustaining activity are shown. Although each column has about 800 connections, only the contributing ones are shown.
input (Donchin 1979). This may be due to the time required to bring a large number of columns into synchrony. Over such a circuit, the temporal response would be almost in synchrony in all the columns. That such states occur is indicated by the studies of Livanov (1977) with Aslanov which showed that the cross correlation of EEG electrodes on the scalp increased dramatically during periods of concentration. Further support comes from the observation that similar waveforms appeared all over the brain in response to a conditioned stimulus (Thatcher & John 1977).

An additional physiological finding that might be related to the formation of the active circuits hypothesized above is the contingent negative variation (CNV) (Walter 1960). A CNV is a long lasting depolarization, seen in the EEG, during the period between a conditioned stimulus and its unconditioned stimulus. If there is some way to prevent the output from pyramidal cells from leaving a column, then the result of input from a turned-on active circuit onto this column might resemble a CNV. One mechanism that might be able to achieve this blockage is a class of area 17 s.i.i. ending on the initial segment of the pyramidal cells (Szentagothai 1978). It seems reasonable that these s.i.i. themselves may be inhibited by a strong sensory input entering the column from the thalamus, and thus releasing the pyramidal output from the column.
2.3.4 Changing the sensitivity of relay and large inhibitory cells

In the interneuron network model, changes in the basic mode of processing result when the sensitivity of either type of interneuron population is effected. The basic effect of sensitizing the s.i.i. is to increase the amount of filtering done by the interneuron network on the sensory input. By decreasing the sensitivity of one particular type of s.i.i. it might be possible to make a particular quality more likely to determine the momentary pattern of interneuron activity. The sensitization of s.e.i. in some conditions can lead to the preservation of an input pattern after the input has ended through the formation of active transients, local steady states, or limit cycles.

What is likely to be the result if the relay cells are sensitized? The main effect would be to reduce the number of cortical afferents which must converge on a column to raise its average firing level. If this occurred the number of columns involved in an activated circuit would increase as well as the firing level of those columns already involved. The number of columns actually involved in an active circuit would tend to remain a small fraction due to the circle of lateral inhibition around each active column. The number of columns in a single nucleus involved in such an active circuit can be increased by increasing relay cell sensitivity, thus decreasing the need for convergence. As a result, the number of synchronized columns in a small region would increase and begin to show up in EEG measurements.
Singer (1979) produced some data that indicates that the pyramidal cells in area 17 are sensitized by acetylcholine along with the s.e.i. The response of output cells in layer 5 to visual stimulus is greatly increased by midbrain stimulation or direct application of acetylcholine. Deutsch (1973) found evidence that cholinergic pathways were important in memory formation. There is some evidence for a diffuse spread of norepinephrine fibers from the locus coeruleus to the limbic system and the cortex. Cooper, Bloom, and Roth (1978) have found evidence that the relay cells are the principle target. Norepinephrine may play a role in learning which will be described in the next subsection.

Suppose the strength of the reaction of large inhibitory cells (l.i.o.c.), in terms of the size of the elevation in the firing rate and the length of time the elevated firing lasted, was temporarily increased. Note that large basket cells are the type of l.i.o.c. found in area 17. What would be the effect on the response to a stimulus in a typical column? It is proposed that this sensitization would lead to periodic oscillations of the firing rates of both the relay cells and the large inhibitory cells. A qualitative description of the sequence of events that drives the oscillation could be the following: (calculating values for equations 2.1C and 2.1D could also give these results)

1) Suppose that a stimulus has reached the sensitized large inhibitory cells in some nucleus. The effect on the large inhibitory cells in a single column would be long fast firing, which would inhibit relay cells over a number of adjacent columns as well as their own.
2) Normally, the relay cells have a high spontaneous firing rate (Steriades, et al 1979), which plays a role in maintaining the activity of the inhibitory cells that they contact. (This resting rate may be altered by midbrain inputs, specifically norepinephrine can lower it (Cooper, et al 1978) and make self-sustaining oscillations impossible.) The IPSP on the relay cells, in response to large inhibitory cell burst, will keep them inhibited for 100-200 ms plus the length of time of the elevated firing rate of the inhibitory cells. The effects of any EPSP on the large inhibitory cells caused by the same stimulus that caused them to burst will fade within 30 ms after the stimulus ends.

3) By the end of the period of relay suppression both the relay and large inhibitory cells should have almost stopped firing at all. As a result, after the IPSPs on the relay cells fade, they will begin firing with almost no inhibition at all, and thus at a higher than normal rate. The long rise time of the IPSP would allow the relay cells to climb to a high level of firing before the large inhibitory can exert much effect.

4) If their sensitivity is raised high enough, the higher than normal firing rate of relay cells might trigger a burst in the large inhibitory cells, even without input. In this fashion, sustained oscillations could be maintained as long as the large inhibitory cells were sensitized.

It is important to note that due to the size of the axonal field of the large inhibitory cells, these periodic oscillations would involve many synchronized columns. Elul (1974) calculated that EEG intensity
from a large region of the cortex should increase linearly with the number of synchronized columns in its field, but only as the square root of the number of unsynchronized columns. Purpura (1974) proposed a similar sequence of events as the basis of the EEG oscillations seen in the thalamus with the onset of sleep.

The most important example of the sensitization of large inhibitory cells occurs in the thalamus. There seems to be a projection of serotonin-releasing axons going from the Raphe nucleus (where most of the serotonin in the CNS is produced) to the thalamus (Jouvet 1974). The reticular nucleus is a thin layer of neurons that surrounds most of the thalamus and projects a diffuse network of axons into the other thalamic nuclei. Singer (1979) found that these projections are largely inhibitory. The relay cells of the LGN, like most of the thalamic nuclei, send collaterals back to the reticular nucleus and stimulate the large inhibitory cells there. These serotonin fibers may act in two ways to increase the sensitivity of a group of large inhibitory cells in the reticular nucleus of the thalamus. First, they may reduce a mesencephalic RF acetylcholine input to these cells that is known to quiet them (Singer 1977). It may also directly sensitize these cells or affect one of their input populations. If the serotonin actually does this, the configuration of neurons is anatomically well structured to support the periodic oscillations described above. Some support for this concept is given by the fact that serotonin is critical to the initiation of slow wave sleep in which there are firing rate oscillations of the thalamic relay cells (Jouvet 1974). Further, one would expect the smallest of these oscillations to have a period of
about the length of the IPSP, or 100 ms (the alpha period), and to get longer and stronger as bursting period of the l.i.c. is increased.

What are the effects of a localized stimulus reaching the LGN under conditions of large inhibitory sensitization? A local maximum of relay cell firing would lead to a larger circle of inhibition around it. The oscillations in the stimulated spot lead to a circular pattern surrounding the oscillations that weaken with distance from the spot. The shape of the outer edge of the spreading ripple would depend on the shape of the axonal field of the large inhibitory cells at its center. The cycling of firing rate changes seen by Verzeano (1973) during states of high amplitude EEG might partially be explained by this effect. What is the effect of large inhibitory cell sensitization on the processing of retinal input after oscillations have started? Its main effect is to attenuate strong signals more than weak ones. This is due to the fact that strong signals would tend to increase the rebound suppression of activity by the inhibitory cells of the reticular nucleus. Weak signals would be less attenuated by the rebound from the large inhibitory cells. A condition under which this predicted equalization of experienced stimuli becomes particularly evident is drowsiness, when comparatively large amplitude EEG waves (with respect to the normal waking alpha or beta waves) are seen. This effect implies that the turn-on of an active circuit would be difficult during periods of high large inhibitory cell sensitivity.
A change in sensitivity of these same large inhibitory cells of the reticular nucleus, in the direction, of desensitization may also occur. Stimulation of some parts of the midbrain seem to release cholinergic transmitters that momentarily block or inhibit these inhibitory cells during saccadic eye movement or any fast movement of the visual field across the retina (Singer 1977). As a result of this, by the end of a saccad the thalamic relay cells are almost completely unblocked and ready for a new input. This effect seems to be localized to the LGN, so the whole reticular nucleus is not involved.

The sensitization of large inhibitory cells, due to their large axonal fields, clearly has effects that transcend the processing of input within a single column. The following discussion will examine how large inhibitory cell sensitization effects the interaction of columns within a single nucleus.

The periodic oscillation of one nucleus, due to the sensitization of its l.i.c., can effect a target nucleus. What would happen to the oscillations generated in one region on reaching another region that is not sensitized? Suppose the LGN were oscillating, then area 17 would receive a strong stimulation to the s.e.i. followed by a lull. The high firing rate of the s.e.i. would increase the firing of both relay and large basket cells in a multicolour area. This would result in a 15 ms stimulation of the large basket cells and a 100 ms inhibition of the relay cells which would overlap with the lull in LGN input. Thus, we would expect an echoing oscillation from an area that is attached to an
already oscillating area. This should be strong if the large inhibitory cells are sensitized, and weaker if they are not. Andersen (1968) found that the generators for alpha and delta waves appear to lie in the thalamus, which suggest that it is the major site of this type of sensitization. However, the isolated cortex has the ability to produce high amplitude brain waves (Purpura 1974). This implies that some way of sensitizing the large inhibitory cells of the cortex exists which involves neither thalamic input nor brain stem monoamines. The eventual appearance of slow wave sleep in animals with their serotonin source removed may be due to this effect. Possibly, under conditions of isolation or lack of brain stem input, inhibitory interneurons that terminate on the large basket cells may become habituated, and thus mimic sensitization of the large basket cells.

2.3.5- Modifiable synapses

The author's basic hypotheses about the nature of a neural recording mechanism can be traced to the modifiable synapses postulated by Hebb (1949) to be the basis for learning. Synaptic modification can either act by strengthening the size of the PSP produced by existing synapses, or by creating new synapses. Indirect support for the creation of synapses comes from the finding that the size of the dendritic tree in areas involved in learning increases (Greenough 1976). This suggests that new synapses are being formed, along with the new dendrites. The strengthening of existing synapses received experimental support when synapses in the cerebellum were found whose responses were permanently
changed after repeated exposure to a strong stimulus (Eccles, Ito, Szentagothai 1967).

There are several restrictions that the behavior of the nervous system places on this candidate for neural recording. These involve the number of neurons and synapses which may participate in recording any scene, and the time at which recording is likely to occur. The number of columns in a nucleus simultaneously participating in turned-on active circuits is restricted to a small sparse set by lateral inhibition. The neurons in any nucleus seem to change their firing rate at least several times per minute. This means that the cells on which either type of synaptic modification is taking place cannot be out of action very long. However, the synaptic modification that is associated with learning seems to require hours or even days to complete (Deutsch 1973). It is therefore necessary for the cells whose synapses are being modified to remain reasonably functional during the modification process.

Arguments can also be made in favor of only a very few synapses per neuron participating in any single learning process. Synapses monitored over hours show no evidence of continually varying their response while the subject is in the same behavioral state. It seems reasonable that synapses undergoing modification are not functional, in which case only a few synapses should be undergoing modification at any time. This restriction is met by an active circuit model where only a few of the thousands of synapses on any relay cell would need to be modified to strengthen the connection between pairs of columns.
What type of signal can cause synaptic modification and where does it occur? The constant barrage of visual input also implies that most neural activity causes no permanent modification of synapses. This concept can be built into the active circuit model by requiring a special type of input from the midbrain be present before any synaptic modification can begin. In accordance with Eccles' findings, the author postulates that the strength of connections between the most strongly firing columns in the active circuit would be reinforced by synaptic modification. The relay cells in thalamic columns send input directly to the interneuron network of cortical columns. This could lead to a turned-on active circuit with thalamic columns altering the input from the interneuron network within its own columns. The fine-grained response of the interneuron network could reinforce the firing of particular pyramidal cells more than others in the same column. This mechanism could preferentially strengthen the key relay neurons in the active circuit, i.e. the ones that go to other columns in the active circuit.

Over a period of time, a much used set of columns could become reinforced to the point where they could even be turned-on by a sensory input without special sensitization. During this process only the most active columns would get much reinforcement, thus the set of columns that could be turned-on without sensitization would be much smaller than the original active circuit. Thus, the amount of convergence within this set of columns required for elevated firing is reduced. This set of reinforced columns will be called an active circuit core.
Construction of an active circuit core probably entails protein synthesis, shown, in numerous studies to be important in memory formation (Hyden 1978). This would require a midbrain transmitter, in addition to the proposed relay cell sensitizer (ACh), that could initiate metabolic changes. The most likely candidate is norepinephrin, which activates a cyclic AMP sequence that could lead to protein synthesis (Cooper et al 1978), and which in some situations was found essential for learning (McIlwain 1979).

It is postulated that the conditions, necessary for triggering the protein synthesis needed to make a new active circuit or active circuit core, will occur only if something in the stimulus is recognized as demanding attention. For example, the recognition of novelty may act as a signal to the nuclei that support recording, indicating that they should allow their synapses to be modified by the incoming signal. Once the recording is formed, the stimulus must stop initiating the recording mode. It is presumed that with repeated stimulations the midbrain trigger for recording becomes habituated. Supporting evidence is found in the reticular formation (RF) response of mice, which fades over repeated presentations of a click (Groves 1976). Thus after a number of presentations and turn-ons of an active circuit, the level of sensitization per presentation would drop as would the amount of change in connection strength. However, by this time an active circuit core would have been formed. Most of the columns originally involved would drop out of the circuit, in the sense that they would not have their connections strengthened. Since the driving stimulus would no longer
lead to sensitization, these unreinforced columns would not participate in the turn-on of the active circuit after the formation of the active circuit core. During normal alert conditions, the strengthened portion of the original active circuit could be turned-on by a simple sensory input. However, since the original connections still exist, the complete active circuit could still be turned-on by relay cell sensitization. Note that the number of synapses affected on any one relay cell would be a few out of thousands. This would allow a single column to participate in hundreds or thousands of these active circuits.

2.4 NEURAL WAVES

Thus far, a general model of the nervous system has been constructed; one that accounts for some of its physiologically observed activity modes. The rest of the paper will concern connecting this model to the recognition process. In particular, the relay system model and its active circuit mode are related to information storage and the process of selecting among the stored information during various behaviors. Several assumptions about the relay system are made throughout the rest of this dissertation:

a) An active circuit is the basic mode of operation for this system.

b) The extent of an active circuit depends on the level of sensitization of the relay cells.

c) Only a special type of active circuit, called an active circuit core, which has been reinforced through repeated exposures can be turned-on by sensory input without sensitization.

d) Active circuits are an information store.
e) Most of the information in an active circuit, turned-on by a stimulus, is drawn from storage.

f) A model of recognition can be based on the process of selecting an active circuit after presentation of a visual input. An active circuit model of recognition raises a number of questions which must be addressed. Note that the concept of an active circuit and the related notion of a neural wave provide a link between neural models and psychological processes.

If recognition involves using information stored in active circuits, then the question of how this information is distributed must be posed. One way to approach this question is to ask if the activity on each column in a particular active circuit can be assigned an information content. The information in a column seems to have two components: a contribution from its role in maintaining the active circuit, and a contribution from the generation of the associated behavior. An example of the behavioral contribution of one column might be the generation of a muscle contraction triggered by a neuron in the motor cortex. Many subsets of columns in an active circuit are large enough to specify the full set of columns involved (assuming connection strengths and the level of sensitivity are known). Just the fact that a column is in an active circuit gives little information about either the current active circuit or behavior, since the column may participate in hundreds of active circuits.
The concept of synchrony requires that a specific temporal variation, or waveform, exist in order to maintain self-sustaining activity in an active circuit. The knowledge that a column is part of an active circuit, and that it has a particular waveform may be sufficient to completely specify which columns are participating in an active circuit. (The vast number of active circuits possible may mean that finding the particular active circuit with the above information is possible only in theory.) Thus, while complete information about the active circuit may be contained in the waveform on any participating column, information about the behavior generated when the active circuit is turned-on requires knowledge about the role of individual nuclei.

There is a practical question, relevant to active circuits, about the nature of the available data. The connections among columns are unknown, and are likely to remain so, while measurements of the waveform and the distribution of activity across a nucleus may eventually be accessible to EEG methods. Instead of asking which other columns are involved in an active circuit, given one column and its waveform, it may become feasible to find the distribution of columns in an active circuit given the distribution and waveform in a single nucleus. Knowing this would still leave the problem of isolating the contribution of the activity in a particular column of an active circuit to an associated higher cognitive processes. Something about this question is known from brain lesion and stimulation data.
The transformation of electric field distributions in moving between locations is routinely calculated by physicists handling problems in optics. If a similar type of transformation were available for the distribution of columns in a nucleus, it would yield a much simpler method for studying the processing within an active circuit than would be involved in solving sets of integral equations. In addition, setting up integral equations, such as equation 2.6, requires unobtainable data about the connections between columns. The objective of this chapter is to construct a rigorous mathematical definition of a neural wave. It is hoped that this will permit the mathematical tools of optics, such as transformations, to be applied to the study of active circuits.

Neural waves play a critical role in constructing an analogy between neural recording and holographic recording (see chapter 4). Although the concept of neural waves arose in connection with holography (Pribram 1969), a neural wave itself cannot be reasonably defined without incorporating the properties of an active circuit in it. The following properties, drawn from the model of the relay system, will be used to characterize a neural wave:

1) The original selection of an active circuit is determined by the interneuron input to the pyramidal cells in area 17, when it is intact. (Otherwise, inputs coming from the SC through the posterior thalamus may initiate an active circuit visually.)

2) Due to the need for convergence, the columns sending input to a given column must be synchronized, in order to raise a column to the turned-on state. Since this holds for all the columns, most of the active circuits should be synchronized.
3) The temporal variation, or waveform, of a turned-on active circuit is determined by the need for synchronization in the set of columns involved. Thus, the same waveform should be initiated when a particular active circuit is turned-on, independent of whether the input signal is externally or internally generated.

4) The small fraction of columns participating in an active circuit control the information processing in each nucleus involved. Input to columns outside of the active circuit would not converge in sufficient number to raise the firing level of these columns significantly. The activity in columns outside of the active circuit can therefore be considered as noise, in discussions of global activity.

For comparative purposes, the properties of optical waves are discussed next.

2.4.1- Properties of optic waves

An optical wave is an information carrying disturbance of the electric field traveling through space. As such, it must obey the electromagnetic wave equation derived from Maxwell's equations (see discussion in Born & Wolf 1975 p.10). Thus the electric field associated with a wave of monochromatic light at any position and time is given by the real part of the solution to the wave equation, for some fixed initial conditions. The real electric field vector at time $t$ and position $(x,y,z)$ is

$$w(x,y,z,t) = \text{RE}\{a(x,y,z)e^{i(p(x,y,z))}e^{i(\omega t)}\} \quad (2.7*)$$

$$= a(x,y,z)\cos[\omega t+p(x,y,z)]$$
Where $a$ is the amplitude of the wave or the maximum value of the electric field over the time domain at some position, $f$ is the frequency of the light, and $p$ is the phase at some position. The term $\exp(ip)$ is called the complex amplitude. The right hand side of equation 2.7 has two factors, called the waveform and the amplitude. The waveform, $\cos(wt+p)$, is periodically variable at any point where the wave is defined, and over the time domain $t$. This waveform is: 1) given by the cosine function, 2) independent of the source of the wave, and 3) is the same at every position up to the phase $p$. Note that the terms $a$ and $p$ (and possibly $w$) contain complete information about any optic wave. At a fixed time there is a sequence of surfaces of maximum amplitude, called wavefronts, periodically spaced and perpendicular to the direction of wave propagation, for any wave (see Figure 4.3A).

There are three important restrictions on the domain over which equation 2.7 is valid. First, equation 2.7 is valid only between a distinct beginning and end point in space at each moment. The distance between the end points measured in the direction of propagation is called the coherence length (Collier, et al 1971 p.26). A particular coherence length, divided by the velocity, gives the time it takes a wave to pass a fixed point in space, this time is called the temporal coherence. A second restriction on the extent of an optical wave is needed to insure that a wave does not interfere with itself. This is prohibited if the extent of the wave at both the source plane and the target plane is much smaller than the distance, $z$, between the planes. Spatial coherence is a measure of the ratio of wave diameter to the
distance it travels between source and target. An optic wave with long
temporal (or spatial) coherence is called temporally (or spatially)
coherent. A final restriction is that the medium through which the wave
travels must be homogeneous. (It is also assumed isotropic for
simplicity.)

For current purposes, the object of interest is the optical wave in a
plane of constant \( z \) (where \( z \) is the direction of propagation).
Therefore, it is useful to work with the equation of a wave in a plans,
given by

\[
\psi(x,y,t) = a(x,y)\cos(\omega t + \phi(x,y))
\]  \hspace{1cm} (2.7)

An optical wave in a plane will be defined as any electric disturbance
described by equation 2.7. It must be monochromatic, temporally and
spatially coherent in the domain considered, and must travel in a
homogeneous medium. The wave equation also implies that the wave is
reproducible by repetitions of the same event that originally caused it.
Note that acoustical waves also obey the wave equation and the
restrictions given above, thereby qualifying as wave phenomena.

Most electrical disturbances are not waves in this sense. For
example, in an electric circuit, the electric field at one point (given
the electric field at a second point) is calculated from the total
resistance, capacitance, inductance, and potential generated along the
lines connecting the two points. Because the wires are surrounded by a
different medium, refraction at the boundary of the two media prevents
electromagnetic waves from traveling in a straight line. Also, the
closed circuit allows feedback, resulting in a constant or a non-
sinosoidal time variation of the electric field. Therefore, equation
2.7 is nowhere valid in an electric circuit.

2.4.2- The properties of neural waves

A brief review of how information spreads through the visual system
will be useful in the study of neural waves. While optical information
is spread via disturbances of the electric field, neural information is
spread through disturbances of the nerve membrane potential on the cells
making up the visual system, as described in subsection 2.1.1. Neural
information travels through the visual system in a series of conversions
from action potentials (AP) to post synaptic potentials (PSP) and back
to AP. The dispersion of neural information in passing from one nucleus
to another results from the profuse branching of axons when they reach
their target nuclei.

A block diagram of the visual system with the component central
nervous system (CNS) nuclei in their relative spatial positions is
presented in Figure 4.4. Visual information entering through the retina
can follow various neural pathways through the lateral geniculate
nucleus (LGN), the superior colliculus (SC), and the pretectum before
going on to the deeper brain regions. Since humans are capable of
recall, information about past experience must also be internally
generated somewhere along the visual pathway.
A number of experimentalists have pointed out that elevated firing rates spread through the nervous system like wavefronts (Popper and Eccles 1977). Pribram (1969) has suggested that the spread of neural information represents a type of wave phenomenon.

"One can imagine that when nerve impulses arrive at synapses, they produce electrical events on the other side of the synapse that take the form of momentary standing wavefronts. Typically, the junctions made by a nerve fiber number in the dozens, if not hundreds. The pattern set up by arriving nerve impulses presumably form a microstructure of waveforms that can interact with similar microstructures arising in overlapping junctional contacts."

It is argued that the initial rush of depolarization due to a stimulus is not the important factor in the resulting behavior, but rather the pattern of self-sustained activity that results from this input. The support of the neural wave is the active circuit, i.e. activity in the rest of the system is considered noise. The firing pattern in a turned-on active circuit has two of the three properties listed for optical waves. First, due to the synchronization of firing times, an active circuit has the same waveform on all of its columns, although the phase may differ. Second, the waveform of the neural wave is independent of the particular input responsible for turning-on the active circuit, by property three of active circuits listed above. However, the waveform of a turned-on active circuit should vary from one active circuit to another, according to the superficial analysis of the relay system model. Therefore, the third property of optical waves,

---

4- The 'momentary standing wavefronts' are not caused by the reflection of a wave back on itself like the standing waves on a string. They are simply chemically induced PSP's that last up to three orders of magnitude longer than the effects of an AP on the membrane at a point, and thus are comparatively stable.
i.e. the waveform is always the same cosine function, does not hold for neural waves. Physiological evidence against periodic variation is given by the lack of periodicity in evoked potentials, which should be associated with the formation of active circuits. (Spekreijse 1978).

The elevated firing rate associated with the turn-on of an active circuit is restricted to a certain time domain, as the validity of equation 2.7 is restricted in the optical case. The beginning is the turn-on of the active circuit, while the end coincides with the breakdown of self-sustaining activity in the active circuit. This breakdown may be the result of interference from saccadic eye movements which disinhibit the LGN and area 17 (Singer 1978), other inputs, or habituation effects. Thus neural waves are restricted by a concept analogous to temporal coherence. Physiological evidence puts a rough limit on the length of time that an active circuit can last. Because any transfer of information between neurons causes a PSP on the target neuron, it would be unreasonable to consider any temporal domain shorter than the several hundred milliseconds which an inhibitory PSP lasts (Thompson 1967 p.181). There is reason to believe several hundred milliseconds is long enough for the shortest recognition event to occur in some parts of the visual system. This is indicated by the fact that saccadic eye movements of up to 30 degrees occur about every 200-400 ms in the free examination of a scene. This movement strongly inhibits firing of the LGN and visual cortex neurons (Brooks 1977, Singer 1977). A natural upper limit to the duration of the turned-on active circuit in the visual cortex is thus formed. However, structures further along the
visual pathway are not so sensitive to saccades, and neural events there may last much longer (Gross 1974, Pribram & McGuiness 1975).

Interference of signals within the active circuit are prevented by the requirement for synchronized activity. Thus, a property analogous to spatial coherence for optical waves is insured for neural waves whenever an active circuit is turned-on.

In the model of the relay system, a certain kind of homogeneity of the medium is supplied by the equivalence of each column in the active circuit. The difference between columns lies in the information entering each column through their inputs. However, the directional nature of the connections between nuclei eliminates free spread between the columns. Nonetheless, the question of whether the spread of neural waves through an active circuit may follow the same rules that govern the spread of optical waves can still be considered. Since the active circuit constitutes a sparse set of of columns spread through many nuclei, there can be no claim that the turn-on of an active circuit leads to the synchronization of whole nuclei.

Another interesting similarity between optical waves and neural activity in an active circuit concerns the formation of time varying, traveling activity in response to a constant, localized input. Heating a physical object leads to black body radiation emanating from it. A localized steady state in the interneuron network can lead to variable activity traveling all over the visual system on an active circuit. The
importance of the interneuron input is exemplified by an effect of barbiturates. This drug affects the interneurons much more than the larger relay cells and can cause gross changes in the spread of a response to a stimulus (Talbot & Marshall 1941). There is some physiological evidence supporting the concept that the waveform of an active circuit in any column is independent of the input and that the waveform is the same in all the columns. Thatcher and John (1977) presented a series of EEG experiments in which a phenomenon called generalization occurred. A stimulus which is distinguishable (i.e. has a different early response) from the stimulus that test animals have been trained for elicits the same waveform and behavioral response as the original. John et al (1974) also pointed out that animals waiting in a test box show the same EEG waveform as occurs in response to the conditioned stimulus. Such waveforms do not occur when the animal is in other places. A supposition about what occurs when two diverse objects such as a cat and a house are presented is that two very different active circuits are turned-on. However, if the visual presentation of a house should somehow turn-on an active circuit appropriate for a cat, then a cat would be perceived. Thus visual experiences can be the same independent of how they are initiated.

The same set of experiments by John (1977 p.161) indicate that, at least in conditioning experiments, a particular waveform appears over a wide area in response to an input. In these experiments, small macroelectrodes were used to show that a waveform lasting on the order of half a second appears in numerous nuclei. The starting times of these waveforms are different in the various nuclei. For repeated
trials with the same conditioned stimulus, the waveform seen throughout the visual system is quite stable. While an animal is waiting in the test environment it spontaneously produces the conditioned waveform, but not elsewhere. Pribram (1969) also performed a set of experiments that agree with the general facts cited from John's studies. Because Pribram used smaller electrodes than John, he was able to examine the waveform at various sites within each nucleus. He found that only certain areas in a nucleus produced the waveform. They were randomly distributed, but consistent over repeated stimulation, as expected for an active circuit. The visibility of synchronized activity in the EEG record during classical conditioning is unusual. However, conditioning is probably a state of high relay cell sensitivity, a time when much larger grouping of columns would participate in active circuits than during less sensitized conditions. This situation would lead to much larger groupings of columns participating in active circuits than average.

It is now possible to propose a description of a neural wave in a given nucleus associated with a particular active circuit. The equation in the nucleus would be of the form:

\[ w(x,y,t) = a(x,y)F[t,p(x,y)] \]  

(2.8)

Here, \( p(x,y) \) gives the phase of the neural wave at time zero, and \( a(x,y) \) is the amplitude of the potential change at the axon hillock of a neuron at position \((x,y)\). \( F(.) \) is a complicated function that for a given wave depends on position in a nucleus only through the phase, as did the cosine function in equation 2.7. Unlike the optical wave where only \( a \) and \( p \) vary with the particular wave, the form of \( F(.) \) is
specific to the particular active circuit. Note that there is no guarantee that an active circuit is turned-on each time a scene is presented. However, neural waves will be restricted to those situations where active circuits do occur.

Chapter 5.4 presents a model of attention and relates it to the idea of an active circuit. It is argued there that while an active circuit is turned on, the average potential in certain columns is raised, making them more responsive to sensory stimuli. Thus, an active circuit can act as an attention channel. The difference between active circuits, neural waves, and attention channels in this model becomes the point of view from which they are considered. An active circuit is a neural model concept, a neural wave relates to the activity across whole nuclei, while an attention channel is a psychological concept.

The main purpose in introducing neural waves is the possibility that well understood optical techniques can be used in the analysis of the relay system. The two techniques tested are transforms (in the next section), and holography (in chapter 4).

2.4.3- The transformation of optical waves

There are two basic properties of optical waves, (defined by equation 2.7) on which the transformation of optical waves depend:
1- The spread of light from any point on an optical wave is described by Huygen's principle, i.e. each point on an advancing wave acts as a source of a circular wave.

2- The interaction of one wave with another is governed by the superposition principle, i.e. when two waves intersect they add at each point.

(Note that any electric disturbance which obeys the electromagnetic wave equation must also have wave properties 1 and 2 (Born & Wolf 1975 p.370))

The formula for the transformation of waves used in holography can either be derived from Huygen's principle and the superposition principle (as done by Fresnel, 1818; see Born and Wolf, 1975 p.370), or from the wave equation (as done by Kirchhoff, 1882; see Goodman, 1968 p.58). The formula is given by:

\[ W(x,y) = \exp(ikz)/i\lambda z \int W(x_o,y_o)\exp(ik/2z((x-x_o)^2+(y-y_o)^2))dx_ody_o \]  (2.9*)

The subscript o refers to the object plane, unsubscripted values of x and y refer to the hologram plane, and the subscript i refers to the image plane (see Figure 4.3). Here \( W(x,y) \) is the complex amplitude at a point in the hologram plane, while \( W(x_o,y_o) \) is the complex amplitude of the same wave in the object plane; \( k=2\pi/\lambda; \) and \( z \) is the distance between these two planes. Referring to Figure 4.3, the transformation of an optical wave in going from the object plane to the hologram plane is described in this subsection. The transformation between these planes is used in section 4.2, but the transformation itself applies to any pair of planes. Note that this equation is the convolution integral of \( W(x_o,y_o) \) with \( \exp(ikz)/i\lambda z\exp(ik/2z(x_o^2+y_o^2)) \).
In the case where \( z \gg x_o^2 + y_o^2 \) equation 2.9* becomes:
\[
W(x, y) = \exp(ikz) \exp(ik/2z(x^2 + y^2)/i z)
\]
\[
W(x_o, y_o) \exp(-ik/2z(x_o x + y_o y)) dx_o dy_o
\] (2.9)
which is proportional to a Fourier transform.

2.4.4- The transformation of neural waves

In analogy with optical waves, the two properties of neural waves which are examined in order to find a between-nuclei transformation for them are the spread and the interaction between neural waves. The dispersion of neural information in an active circuit is controlled mainly by the between-nuclei axons. It also depends on the extent of the active circuit. The midbrain neurotransmitters seem to control the extent of the LGN response to a stimulus (Singer 1977). In general, it may sensitize relay neurons and increase the size of the active circuit, without changing its basic waveform. Other influences on the dispersion of information arise through feedback relationships with areas outside the active circuit.

A single column sends axons to many nuclei, and many axons to some nuclei. The rule for much of the visual system is that retinotopically equivalent areas are connected. Thus, for between-nuclei connections a column is more like a hole, with non-diffuse, coherent light passing through it, rather than a point source. By the time the IT is reached, the retinotopic layout is largely lost (Gross 1974). Columns sending inputs to the IT, hippocampus, and other nuclei where cells react to
visual stimuli over a large fraction of the visual field should be considered holes with diffuse coherent light passing through them. Note that Huygen's principle is relevant to points, not holes.

It seems that an optical model for a nucleus should be a sheet with a number of holes in it, corresponding to the columns in an active circuit. A number of light beams should pass through the sheet at different angles of incidence and should proceed to different target nuclei. Some beams might pass through ground glass diffusers, while others would not. In optics, light passing through a hole forms an Airy pattern, as in Figure 2.16, on a distant screen (Collier, et al 1971). This sort of pattern seems impossible in the nervous system. However, the question of interest is whether points where more than a certain number of these patterns intersect could correspond to the position of columns in the active circuit. A different active circuit would be modeled by holes in different positions in the sheet representing a nucleus. Global transformation of the retinotopic layout, such as the partial log polar transform in going from the retina to area 17, should also be accounted for in the model. Perhaps this could be done by introducing an anisotropic medium between the nuclei. However, there is no experimental method capable of locating the columns in an active circuit.

The interaction of neural waves— The neural wave concept being considered is based on the assumption that each synapse on a neuron contributes an independent effect at the axon hillock. Therefore, the
FIGURE 2.16: AN AIRY PATTERN

This is the pattern that is seen when a beam of monochromatic, coherent light passes through a small hole. It would also give the positions of the active circuit in one nucleus, if the optical transformation is applicable to the nervous system.
effects of neural waves converging on the same nuclei would be linearly additive. Each active circuit is composed of a set of columns, which are points of intersection for various information carrying pathways within the circuit. A signal coming from one nucleus could interact with a signal coming from another nucleus in the same active circuit. These two signals, having undergone different processing, would have the same waveform since they were both part of the same turned-on active circuit. Their additivity constitutes a superposition principle for neural waves.

The contribution of particular nuclei to behavior—There are a number of different types of response characteristics found among the retinal ganglion cells. Their response, therefore, no longer simply indicates illumination in some portion of the visual field. As neural signals travel away from the retina, the maximal stimulus becomes more and more complicated. Therefore, one must distinguish between activity on neurons and the information content of this activity.

The role of the LGN in information handling seems mostly concerned with reorganizing and filtering the visual input. Filtering could be done by the periodic oscillations that have been proposed to accompany large basket cell sensitization, or by increasing the extent of the response to a stimulus due to frontal or midbrain stimulation. The superior colliculus (SC) and pretectum may also be involved in filtering, especially in connection with movement. They may play a role in transforming an input distribution from one coordinate frame to another. A transformation must be performed on the retinal input to
compensate for the position of the eyes in the head, so that combined eye and head movements can be made. Similarly, the position of the head with respect to the rest of the body must be compensated for, so that visually guided movements can be made (Shebilske 1979).

Area 17 has several functions, one of which is the formation of sensitivity to particular qualities by processing in the interneuron network. Marr (1976) has proposed that response to slits of various sizes and orientations in area 17 could be combined to extract the outlines of figures. It is proposed in subsection 2.2.4 that such a combination could be based on the convergence required in the formation of an active circuit. The extraction of objects from a scene which are perceived as figure, as opposed to background, is a key step in visual processing. The emergence of a figure is postulated to correspond to the turn-on of an active circuit core (see section 5.6).

There is physiological evidence that the functions performed by the LGN, SC, and area 17 are not critical for maintaining the active circuits whose turn-on leads to the experience of visual imagery. It is found that damage to these areas results in loss of part of the visual field, but not in a loss of the ability to recognize (Pribram 1971). The areas that have neurons with large enough receptive fields, and where damage interferes with recognition include the IT, PRF, PUL, LP, areas 18 and 19 of the visual cortex, and the SSS. The hippocampus and the frontal cortex also play a major role in visual processes, as well as other sensory processes.
In section 3.4, evidence is cited that the visual system makes available information that is independent of viewer position, i.e. non-egocentric information. It is also proposed that non-egocentric information plays a key role in the formation of long term memory. The first step in the transformation that leads to the formation of this type of information takes place in the visual cortex. The receptive fields of neurons in the visual cortex outside of area 17 are much larger. Zeki (1974) has found at least five loosely retinotopic representations of the visual field in areas 18 and 19 of monkeys. Each one seems to specialize in a different type of visual information such as color, movement, or binocular disparity. Palmer, Rosenquist, and Tusa (1978) have found six to ten more representations of the visual field in the suprasylvian sulcus (SSS). These areas may each specialize in a quality that area 17 cells respond to, such as contrast sensitivity, the size or shape of an object, or spatial frequency\(^3\). These are the first areas along the visual pathway that seem critical to perception rather than just the extent of the visual field. People with blindness due to damage in these cortical areas can no longer recall visual imagery, unlike those with damage in LGN or area 17 (Konorski 1967). Thus a role is indicated for these areas in the storage of

---

3- In spite of the enthusiasm created by the discovery of spatial frequency sensitivity in area 17 cell, it is only one of many visual channels (Campbell 1974). Lately, the notion that spatial frequency sensitivity indicates that the visual system is performing a Fourier transform of the retinal input (Pribram 1974, Pollen, Lee, & Taylor 1971) has fallen in disrepute. At best the visual system performs a Fourier transform on one kind of input, gratings. A simple disc is not transformed into an Airy pattern as it should be. In addition, Wilson (1976) has argued that the sensitivity to gratings in humans is not finely tuned enough to support a Fourier transform. Julesz (1979) also objects to the notion of a Fourier transform based on the way certain visual patterns are recognized.
visual information, which may be connected to the information that the
signals which arise there carry.

It is possible that the ability to initiate internally generated or
reference neural waves rests in the subcortical areas connected to the
visual cortex, such as lateral posterior nucleus (LP), the pulvinar, and
the RF. The idea that LP, PUL, and PRF are storage sites is supported
by the important role they have been found to play in short term memory
(Kesner 1977). The active circuit cores turned-on in short term memory
may include these nuclei. This idea supports a role for subcortical
input in selecting which relay cells in a column are strengthened during
synaptic modification. These subcortical areas may be the main sites
from which internally generated signals can turn-on an active circuit in
the visual system. Note that the use of visual information after the
removal of the visual cortex makes the cortex itself a questionable site
to propose for the total store of visual information (Weiskrantz 1974).

The IT seems to form a critical intersection for visual signals.
Many of the paths in a visually triggered active circuit, including ones
from the SC, converge on the inferotemporal cortex (IT). The IT is
probably the first nucleus along any visual pathway in which individual
cells react to non-egocentric forms. Gross (1974) found a monkey neuron
that reacted to a monkey's hand over most of the visual field. Most IT
neurons react to complex retinal patterns presented to either eye over
at least the foveal region. Data also indicates that IT activity
reduces the reactivity of early visual regions. Pribram (1969) found
that IT stimulation reduced the size of the LGN response to a stimulus.
Luria (1973) reported a condition known as simultaneous agnosia, in which the patient could only see one object no matter how large. The condition could be partly relieved by injecting a stimulant into the IT.

The following role for the IT is proposed during exploration of a new object or free examination of a scene in the absence of strong expectations. The reduction in the sensitivity of relay cells caused by IT stimulation could reduce the possibility of an active circuit turning on. On the other hand, this effect of the IT should serve to increase the number of active circuit cores that can simultaneously be turned-on without interference between them. Simultaneous agnosia indicates that the IT would also participate in reassembling meaningful scenes within the brain. Thus, the proposed role of the IT is in incorporating a number of active circuit cores into a large active circuit.

There are severe deficiencies in the ability to recognize scenes when the IT is damaged. Because of its position between visual input and the limbic system the IT seems to occupy a pivotal role in the turn-on of any large active circuit. In cases of strong expectation, an active circuit involving the IT may already be turned-on when the visual input arrives. Recognition may occur when the visual input compliments or easily fits the active circuit already in the IT. However, when the active circuit core initiated by the visual input is in conflict with the active circuit in the IT, the breakdown of the active circuit in the IT may result in a general recognition of novelty.
The role of the IT as an integrator of information is not compatible with certain data about the nature of stored information. Triesman (1977) pointed out that it is quite common to confuse the color or shape of an object being recalled. The stored information about different qualities may lie in different places. The particular active circuit turned-on, would have to incorporate the right columns in order to retrieve the correct information. The inability to form visual imagery when the IT is damaged, according to this model, would result from inability to integrate active circuit cores into a whole. So, without the IT animals should be restricted to turning on a number of unrelated active circuit cores in the visual cortex and the posterior thalamus.

The neural wave model is grounded on the electrophysiological behavior of whole nuclei. This in turn depends on the synchronized activity of a small number of columns composing active circuits within the nuclei. Some of the best studied data in neurophysiology are those connecting the type of global cellular response considered here to the global behavior of the organism. The study of brain waves (e.g. alpha, beta, and delta waves), and their relation to sleep and waking (Thompson 1967 p.211) make this connection. Other examples are John's studies of the EEG accompanying conditioned response (Thatcher & John 1977), and studies by Livanov and his associates of distal synchrony of EEG during rest and mental arithmetic, in schizophrenics and normals (Livanov 1977). One could also put the studies of performance after brain lesion in this category. Although electrophysiological measurements of lesioned animals have only occasionally been taken, the comparison of behavioral deficits and EEG abnormalities may be illuminating. EEG data
seem to have a more direct connection with higher brain functions than single neuron measurements. It appears that the best way to use such data in a neural model entails a wave description of neural activity.

This concludes the discussion of the relay system. The concepts which have been developed are related to the inputs of the recognition process, i.e. stored experience and attention, in the rest of the paper. The following chapters will discuss the various modes of memory, and their effects on subsequent behavior.
3- MEMORY

3.1- INTRODUCTION

Memory implies the ability to both store and retrieve information; the following chapter will primarily examine the different stages of information storage. The input from stored past experience, used in the recognition process, is examined from a psychological point of view, i.e. via experiments that test behavioral responses. In attempting to deduce the information stored in the brain from behavioral responses, researchers encounter two types of problems. Some of the stored information may be lost in a very short time, or may no longer exist in its original form. In addition, the stored information may not be available at the conscious level. For instance, the sequence of eye movements used to examine a picture is the same from trial to trial, even though the subject may be completely unaware that he is moving his eyes at all. Therefore, experiments testing conscious memory do not have access to the complete set of stored information.

Although researchers in the field of memory are not in agreement on how memory actually works, three stages are generally agreed upon: iconic, or sensory memory; short term memory (STM); and long term memory (LTM) (Atkinson and Shiffrin 1968, Lindsey and Norman 1972). These three stages differ in duration and capacity, in the type of information stored, and in the effects on future processing of sensory information. Connections between the different stages of memory exist; continued
processing of iconic memory causes information storage in STM, and further processing can lead to LTM. Following a discussion of each of these stages of memory, a neurological model of each one is presented.

The processing properties of the interneuron network and the relay system, covered in sections 2.2 and 2.3, were used to construct a model of the various stages of memory. Modeling memory with a four cell type model allows more versatility than previous one or two cell type models, such as van der Malzberg's model of orientation learning. The connections between the interneuron network and the relay system that are built into equation 2.1 lead to some predictions about the interactions between the stages of memory that are, in fact, supported by psychological data. Further refinements of the general neural models developed in sections 2.2 and 2.3 are suggested in order to better fit the psychological properties of memory. The sequence of events and areas involved in the modification of synapses were not treated in the previous discussion of the permanent recording of visual information. This short coming in the model of LTM will be partially remedied (section 4.2) with the presentation of a neural wavefront model, based on an analogy with holography. The holographic model is able to take into account the role that context plays in information storage.

3.2- ICONIC MEMORY

The first step in the visual perception process is the formation of a sharp image on the retina. In the first 200-400 milliseconds after
stimulation the pupils adjust to the brightness of the stimulus and the eyes accommodate (Gregory 1966). Accommodation is a process which prepares the eyes to deal with input from a particular region of the visual field. For a near point, the eyes converge so that the fovea of each eye is directed towards the point, both lenses become more spherical; and the pupils contract (possibly to reduce spherical aberration) to form the best image on the retina. If the chosen point is not straight ahead the eyes move in order to center it in the field of vision. In addition to focusing the image on the retina, accommodation allows binocular fusion of the scene in each eye into a single entity. Once a good image has been formed it is possible to speak of storing visual memory.

An afterimage is a typical example of iconic visual memory. Lasting on the order of one second, an afterimage is seen when the eyes are closed after viewing a bright object. The information seems to be stored in its original, unprocessed form. The capacity to store visual information in this manner is quite large, compared with what can be consciously attended, and depends on the length of exposure to the bright object.

Sperling (1969) conducted an experiment wherein human subjects were presented for 50 milliseconds with a 3 x 4 matrix of symbols, followed by one of three tones. The tone indicated to the subject which of the three rows of symbols he or she was to reproduce. It was found that subjects were much more successful with the tone than when they were
verbally instructed which row to reproduce, although most of the benefit of the tone was lost if the time was delayed for just one second. This implies the existence of a large capacity, short endurance store that fades in a time period on the order of a second.

When a strong stimulus is presented for several minutes the result can be a long lasting aftereffect, such as the McCollough effect. To produce the McCollough effect, subjects stare alternately at a black and orange vertical grating, and at a black and blue horizontal grating for a four minute period. Thereafter, for the next twenty-four hours subjects report perceiving black and white vertical gratings as black and blue, while black and white horizontal gratings are seen as black and orange.

Another experiment that seems to demonstrate iconic storage is cited by Julesz (1971), and involves stereographic images of random dot patterns. In the normal viewing of a stereogram, each eye is presented with one half of the stereo pair. In this experiment, a region of dots on just one of the stereo images in the pair was shifted. The subject then sees the shifted region as standing out, in depth, from the rest of the picture. Julesz cites the case of an eiditiker who was exposed to the left stereo image of a 100 x 100 dot figure for several minutes one day, and then the right figure the following day. When shown the right half of the stereo image, she was able to correctly identify the shape of the shifted region of dots, demonstrating that it was possible for her to store a 10,000 unit figure in iconic memory—for a period of twenty four hours.
It may be argued that two distinct neural mechanisms are responsible for iconic memory. For short term effects (lasting less than one second) the iconic storage maybe due to the visual pigment in the retina. During the recovery period following stimulation, the affected cones would fire at a reduced rate, causing the formation of a negative afterimage. In contrast, the McCollough effect may be caused by a habituation effect on those cells that respond to combined color and orientation information. McCollough has postulated that the habituation of cells in area 17 is responsible, in much the same way that mollusc cells become habituated after extended periods of stimulation. Presumably, these habituated cells would lie in the interneuron network of area 17 where the author proposes that perceptual qualities, such as line orientation, are discerned.

3.3- SHORT TERM MEMORY (STM)

The chief difference between iconic and short term memory is the increased amount of processing performed on the input. An example of the use of STM would be when a subject tries to remember a string of numbers, letters, or words, soon after they have been presented. Several types of processing must be carried out in the brief interval before the information is recalled, in order for the information to be remembered. Using a visually presented string of numbers as an example, the scene must be separated into distinct units, each of which is used to trigger a verbal representation. Note that a visual presentation of numbers is a very complex stimulus that would need to be examined for
several minutes in order to cause the habituation required to preserve it in the extended iconic store. However, by breaking it up into units identified by labels, or by small sets of qualities, storage can be quickly accomplished. In addition, the information to be recalled must be continuously kept in the focus of attention through the process of rehearsal. When rehearsal is disrupted, either through neglect or by interference from another task, STM degrades in a matter of seconds. A picture of seven numbers presented for half a second can be remembered for minutes, or longer, if rehearsal is allowed. Subjects, told to remember a set of three letters made essentially no mistakes in reporting them at any time interval tested. But when an interfering task disrupted their rehearsal, only 80% of the responses were correct after 3 seconds, and just 10% were correct after 18 seconds (Peterson and Peterson 1959). Thus, it seems that the duration of STM varies greatly, depending of both activity and mental state.

It is a difficult task to remember a set of arbitrary objects; seven objects are about the limit of a human's capacity for this task (Norman 1969). The simpler and more familiar the objects to be remembered, the greater the capacity for storage. In order to use STM to remember a set of arbitrary objects, the input must be 'kept alive' through constant rehearsal, and concentration maintained during the rehearsal process. When the maximum number of objects are stored in STM, attention is focused almost exclusively on rehearsal, by blocking out sensory input and other types of cognitive processing.
The following characteristics are necessary in order for objects to be stored in STM:

1) They (the objects) must be familiar enough to be divided into figure and ground.

2) They must be identifiable by a few specific qualities, e.g. a verbal label in the case of a verbal rehearsal.

3) There may not be a meaning, or context, built into the order in which the objects appear. A telephone number, for example, has meaning but there is no reason for a particular number being in any specific position in the string.

4) They may be composites of simple units constructed through a process called chunking, which requires time allowed for reorganization of the data. It is possible to remember twelve numbers, as opposed to seven, by organizing them into four groups of three (Kahneman 1973).

Locating a particular piece in a jigsaw puzzle is an example of how STM can be used to store nonverbal objects. Here, the desired piece of the puzzle is specified by its size and shape. A mental concept of the piece being sought is put into STM and used to search among the remaining pieces for the correct one. Rehearsal, in this case, seems to be purely visual. In addition to sensory units, movements and plans of action can also be stored in STM, although they are normally converted to verbal descriptions for storage.

Short term memory storage also affects the future processing of information through interference and by altering the amount of arousal.
STM interferes with the amount of attention that is allocated to additional sensory input. Adding a zero to the front of a string of verbally or visually presented digits interferes with the number of digits (excluding the zero) that can be remembered, even if the subject is told in advance that the zero will be there. Surprisingly, adding three zeros to the front of the string has a less detrimental effect on the number of the following digits that can be recalled (Kahneman 1973). The three zeros apparently form a better 'group' and can thus be separated from the rest of the digits more easily.

In addition, STM affects the processing of future input through changes in the amount of arousal to certain stimuli. This is demonstrated by experiments which show that the response to the second in a pair of stimuli is faster than to the first. The amount of information that can be obtained from the second stimulus is also changed. Arousal seems to reduce reaction time, but also to reduce the amount of detail that can be retained from sensory input. A related arousal effect is the change in a subject's intensity threshold to a picture of two words, that results when an auditory signal (noise) is either present or absent. The threshold for a lightly written word, in the presence of noise, is higher than for a dark word, compared to the noise-absent situation. In this experiment, the amount of arousal is increased by noise, as it was increased in the previous experiment by
the first in the pair of stimuli.

3.4- A MODEL OF SHORT TERM MEMORY

In the discussions of the interneuron and relay system models, neurological mechanisms for maintaining self-sustaining activity were presented. A localized, steady state (active transients) was proposed as the mechanism in the interneuron model, and the active circuit concept was proposed for the relay system. It is now further proposed that proper combinations of localized steady states would lead to the turn-on of an active circuit core for the boundary of the familiar object. This would lead to the recognition of the object associated with that active circuit core. Included in the active circuit core would be neurons whose firings give access to the proper verbal label.

It was proposed earlier that area 17 may be responsible for splitting a scene into 'homogeneous' areas, or 'good groupings'. In terms of the neural model of the interneuron network, this could be due to the formation of a localized steady state composed of neurons reacting to boundary patterns. Specifically, the edge of an object might trigger a localized steady state in a particular orientation slab within a column. All the cells in the slab need not be involved in the localized steady state; perhaps only those cells responding to a particular color and spatial frequency would develop self-sustaining activity. This localized steady state would maintain a precise representation of certain characteristics of the visual input.
The interesting finding that inhibiting the Na/K ATP-ase blocks short term memory in chicks (Mark 1979) supports the claim that the interneuron system is critical to STM. When the Na/K ATP-ase is inhibited, the sodium pump stops functioning and the membrane's resting potential drops as the neuron continues firing. Such an inhibition would quickly interfere with actively-firing small cells, and interrupt a localized steady state composed of interneurons. Because long term memory does still form under this situation, it must have a different mechanism, if not a different site.

Interference limits the capacity and duration of short term memory. The capacity of STM, in this model, would be determined by the number of active circuit cores that could be activated simultaneously without interfering with one another. During rehearsal, attention is shifted from one active circuit core to another. The interneuron model predicts that localized steady states would be subject to interference by future inputs, causing the breakdown of self-sustaining activity.

The large amount of mental effort required to maintain short term memory could be due to the difficulty in maintaining a sensitizing midbrain input to the interneuron network. Strong mental effort also has been postulated (section 2.3) as a prerequisite to the turn-on of active circuits. Part of the mental effort may be used in turning on an active circuit which supports the rehearsal program. In fact, it is probably part of this circuit that maintains activity in the midbrain neurons that sensitize excitatory interneurons in area 17. Thus, the
model predicts that the active circuit associated with rehearsal is required for maintaining localized steady states. Indeed, rehearsal is required for maintaining STM. The ongoing rehearsal active circuit must rapidly suppress any active circuit core turned on by sensory input. The circuit would also have to sequentially reinforce each active circuit core involved in STM in turn. Thus, STM may depend on the interaction of the activity in the relay and interneuron systems.

One datum in support of an important role for the midbrain in maintaining short term memory is that a protein synthesis inhibitor caused the greatest deficit in STM when applied to the midbrain. The neural model proposed here predicts this effect as a result of the interference that the inhibition of protein synthesis would cause in the maintenance of localized steady states.

3.5- LONG TERM MEMORY

The duration of long term memory is on the order of hours to years; there is some question, in fact, as to whether any information once entered in LTM can ever actually be lost. Perhaps the inability to retrieve information simply results from the inability to locate it. Forgetting seems to increase with the amount of time since the subject was last reviewed (Krech 1974). Thus, maintaining the ability to retrieve LTM seems to depend on bringing the stored information to awareness occasionally, as was true of STM. However, in LTM, these recalls can be hours, days, or even years apart, in contrast to STM
where the frequency of information recall is on the order of seconds apart. The retrieval of information stored in LTM can be triggered in many ways, indicating that LTM has multiple paths of access. The total capacity of LTM has no clear limit. However, there does seem to be a rough upper limit on the rate at which new information can be stored.

The hypothesis that LTM and STM are two distinct processes is supported by the existence of a dysfunction termed conduction aphasia. Patients who suffer from this condition are deficient only in tasks requiring STM. In the normal use STM, objects to be memorized are 'stripped down' and remembered without any particular context, as in the case of memorizing a list of words. If the word 'cat' were on a list of words to be memorized, one would only remember the word 'cat', and not any of the interesting things one might ordinarily associate with cats. Conduction aphasics are unable to perform this sort of task.

Memory for exact phrases—as in the ability to exactly recall sentence syntax—also has the qualities of STM. But when a string of words can be connected to form a coherent thought, the meaning of the thought, rather than the phrasing, is remembered. This type of information seems to be stored in a different fashion. This is further evidenced by the inability of people with conduction aphasia to recall the exact words used in a sentence, although they may still be able to paraphrase its meaning.
Items in LTM interact, as demonstrated by the existence of proactive and retroactive inhibition. In proactive inhibition items learned prior to the items of interest are intermingled with the current list during recall, while in retroactive inhibition items learned after the current list are intermingled with it. Interference between the two lists tends to be more severe in both cases when the items on the list have greater similarity. Interestingly, the interference between similar lists is less severe if the lists are learned in different rooms. In addition to items on a list, meanings can become intermingled. There is a strong tendency to store a meaning for a story that fits one's life experience rather than its exact meaning. This tendency is illustrated nicely by the gross modifications made by people when they are asked to interpret folk tales from another culture (Bartlett 1932).

The greater interference among similar items in LTM also illustrates an important point about the way things are stored in LTM; namely, that they are organized into units of meaning. Although items of the same type tend to be remembered in the order they are presented, similar meaning can overcome this tendency. A random list of words chosen from four categories tends to be remembered by categories independent of the order of presentation (Bousfield 1944). In fact, the interaction among similar items going into LTM may set a limit on the rate at which items can be stored in LTM.

The interaction of similar items in LTM is probably tied to the ability of higher mammals to put movements or concepts together like
pieces in a puzzle. In contrast, lower animals often have movement sequences hard-wired into their CNS. Etiologiststs have called some of the longer sequences fixed action patterns (Fentress 1977). Higher mammals also have some fairly stable response patterns, but almost all of these can be changed given a long enough training period.

The difference between flexible and stable behaviors in primates suggests that there are two types of items stored in LTM. The stable type of storage is associated with routes. For example, a route could be to walk one mile; turn left, go a hundred yards, etc. Another type of route is the reciting of words in a list, or the ideas in a message. A route is defined as any stable sequence of actions or concepts where each one flows into the next.

The flexible type of storage in LTM is associated with maps. A person's memory of his or her bedroom is an example of an internally stored map. Tolman (1932) called these stores cognitive maps, and O'Keefe and Nadel (1978) have recently renewed the interest in such things. The key properties of this type of store are the ability to insert items in specified positions and the ability to know the relative positions of any two points in the environment, even if there is no known route between them. In using a cognitive map of a bedroom, for instance, many views become available that were never observed by the subject. This implies information independent of the subjects position is being generated. The formation of viewer-independent or non-egocentric objects is the key to understanding cognitive maps.
The behavior during which cognitive maps are formed implies a method for achieving this viewer-independent property of objects. It is proposed that the ability to do this in the visual scene analysis system is associated with the character of the information extracted in the search that takes place after novelty is recognized. Behavioral studies show that animals are sensitive to differences in an observed scene compared to the past experience of the scene. This response in the case of a well known scene can be considered the recognition of novelty. It is demonstrated by the fact that a novel component in a scene will receive more attention than it does in later presentations (Kahneman 1973). Another behavioral response to novelty appears when rats are placed in a new environment. They sniff around the unfamiliar objects and move through every part of the environment (O'Keefe & Nadel 1978 p.240). In higher animals, the exploratory behavior that is initiated is dominated by eye movements. This exploration could give an animal a good enough understanding of an object to include it in a cognitive map. The non-egocentric information extracted in exploration can readily be combined to give the full set of properties of the cognitive map. This implies that the animal is synthesizing the cognitive map by incorporating more elementary maps, which will be called templates, into a consistent structure. For this reason, the author will call the map-like LTM stores synthesized scenes for static objects and synthesized events for dynamic processes.

Evidence that non-egocentric maps are used by various animals in navigation and other activities has been presented by O'Keefe and Nadel
(1978 p.63). Another example of the use of a non-egocentric map occurs when a person mentally rotates a geometric object.

Support for the hypotheses of novelty recognition and non-egocentric information can also be found at the cellular level. Hippocampal cells seem to respond earliest and most consistently to novel things and places. Thompson (1976) found that certain hippocampal cells are the first to respond to the combined appearance of a conditioned and unconditioned stimulus in rats. The strongest evidence that the hippocampus registers novelty is the existence of hippocampal mismatch cells (O'Keefe & Nadel 1978 p.197). These neurons seem to respond maximally when a new object appears in a familiar scene, or when an old object in the scene is missing or has been moved. An interesting point about these neurons is that their response is independent of the subject's position relative to the scene. This indicates that the subjects are reacting to non-egocentric information.

Hippocampally damaged animals show a severe memory deficit, especially in the formation of new memories. This deficit may be tied to the lower distractibility to novel stimuli and the disturbance of exploratory behavior in the lesioned animals (O'Keefe & Nadel 1978). The formation of new synthesized scenes should be severely limited by this type of deficit. Repeated presentations of the identical task to hippocampally damaged primates results in a small improvement in performance. However, most of this improvement can be attributed to more precise motor control rather than to the learning of the task (Pribram & McGuiness 1975). For example, in repetitions of a maze
problem, a hippocampally damaged subject will make as many mistakes in subsequent trials as the first time, but he will navigate the maze more skillfully. O'Keefe and Nadel's (1978) basic claim is that the effect of hippocampal damage is to impair the use of old cognitive maps and prevent the formation of new ones. Possibly, this occurs through the disruption of the behaviors needed in the formation of these maps.

In concluding the discussion about the two types of objects in LTM, i.e. templates and synthesized events, it is noted that there is a large amount of interaction between them. Take the mnemonic device as an example (Norman 1969). The subject starts out with a synthesized scene of a synthesized event and selects a room, from which he creates a route through the room. Following the route through the room each item to be remembered is placed on a familiar object. This creates a new synthesized scene of the room with the items to be remembered included. When the subject wishes to recover the remembered items, he simply follows the same route through the room and picks up the desired items when he reaches the place in the room where he put them. Thus, it seems that routes can lead to changes in synthesized scenes and synthesized scenes can be used to form synthesized events from which routes can be selected. Another interesting example of the interaction of maps and routes occurs during the proof of mathematical theorems. Pask (1975) found that there are two basic types of learning strategies used, serial or holistic. The serial learner prefers to learn a route through the theorem first and later (if ever) understand why it works. On the other hand, the wholist learner prefers to understand the point of each step
in the overall strategy as it is learned, i.e. place it into a synthesized event. In either case the end result is a synthesized event of the theorem that allows it to be applied.

3.7- A NEURAL MODEL OF LTM

In this section a neural model of LTM is proposed, centered around the relay system model presented in chapter 2.3. In the model, the synthesized scenes and events stored in LTM would be associated with active circuits, while the routes are associated with active circuit cores. An active circuit is structurally defined by a set of connections between columns in the CNS, assuming the amount of convergence required is fixed at a particular sensitization level. The exact set of columns in the active circuit is determined through the process of synaptic modification, which alters the connection strength among columns. When several sets of neurons responding to different qualities of the input are commonly activated together, synaptic modification can strengthen the paths between them. These paths may then become part of an active circuit that will be turned on in response to a stimulus with a variety of subsets of the qualities.

A full active circuit is quite extensive and it requires mental effort to turn it on completely. An unanticipated stimulus would normally turn-on an active circuit core. Active circuit cores involve a much smaller set of columns, with stronger connections among them, than the connections among active circuits, which makes them much easier to turn on.
The properties of LTM that must be accounted for are its: duration in spite of continued learning, multiple access, loss of specific detail, capacity, organization by meaning of the synthesized scenes, and flexibility. The active circuit or active circuit core has the following properties.

1) It has long duration because it is structurally defined.

2) It is resistant to new learning because of the redundancy required by convergence and the sparseness of the columns involved.

3) It has multiple access since it can be turned on by numerous combinations of stimulus qualities. In addition, active circuit cores can usually be turned on by a single word, while active circuits might require a number of verbal cues to be triggered.

4) It cannot retain exact information about an input for the same reason it has multiple access. If a variety of signals can turn on the same active circuit, a particular circuit being turned on would only indicate that one of this set has been seen or thought about. The exact information stored in STM must therefore be associated with the interneuron network as claimed above.

5) It has a large capacity due to the number of active circuits possible. Suppose that each column has 500 afferents coming from other columns, and that 10 of these must be firing strongly at once in order to bring the column into an active circuit. In this case, each column could participate in 500!/10!x490! = 10^{21} by Sterling's approximation, active circuits if exactly 10 were required. Multiplying this by the five million columns in the CNS gives more than 10^{26} possible active circuits. Factors that would tend to increase the estimate of the number columns are that there may be
more than 10 columns involved in some active circuits and that any 10 columns could participate in a number of active circuits. A factor that would tend to make this estimate too high is that all the possible combinations might not lead to feasible active circuits.

The organization by meaning and the flexibility of synthesized scenes and events are not properties of active circuits in general, but depend rather, on the properties of specific nuclei in the CNS.

Using the convergence property of active circuits, and the response properties of neurons in different parts of the CNS, leads to an argument supporting organization by meaning for information stored in active circuits. This property can be related to the columns shared by different active circuits or active circuit cores. A subset of cells in any column responds to a specific quality such as red color. Neurons that respond to this property in various parts of the visual field converge onto some column, in area 18, for instance which will respond primarily to color. Other regions in the extrastriate cortex respond to other qualities due to the convergence of area 17 neurons responding to multiple qualities. This sequence is based on the assumption that reactions to qualities are randomly mixed. In turn, a number of extrastriate neurons and posterior thalamic neurons responding to special qualities over a fair portion of the visual field appear to converge on an IT column. The IT column may respond to specific shapes and colors in a non-egocentric way. The IT output may go to an area that also gets input from an auditory pathway concerning verbal labels.
A final convergence onto an area receiving labels from the same category seems likely. All the neurons mentioned above may be stimulated by a visual input, and an active circuit core incorporating the columns they lie in could be turned on. If two successive pictures were in the same category the turn on of its active circuit would be made easier because some of the columns would already be involved in self-sustaining activity. Trying to learn a second list in the same category would be made more difficult, since turning on the category columns would tend to activate the recently learned items in the original list.

In the next two chapters an approach for modeling the flexibility of maps is presented. This involves accounting for the use of context in the formation of LTM. The method of linking active circuit cores will be discussed when attention is considered.

3.6- THE EFFECTS OF SENSORY DEPRIVATION

A number of experiments have indicated that synaptic modification, involving a major fraction of the synapses under area 17, may occur in special circumstances. Deprivation of specific sensory qualities during early youth are examples of such experiments. Hubel and Wiesel (1977) have shown that suturing closed one of a kitten's eyes for as little as a week, during a certain critical period, can totally change its visual perception. Such a kitten becomes blind in the sutured eye after the sutures are removed. The effect on the visual cortex is that the ocular dominance columns disappear, and all cells are driven by the unsutured eye, none by the sutured one. Greenough (1976) found that the suturing
lead to a sizable reduction in the length of relay cell dendrites as well as a reduction in their dendritic spine density. Interestingly, the effect of an enriched environment is an increase in dendritic length and spine density equal in magnitude to the deprivation effect. This may be why Movshon (1976) found that reversing the sutured eye after a week lead to the whole visual cortex responding to the originally sutured (and blind) eye, but not the other.

Hirsch and Spinelli (1971) showed that equally dramatic effects could be obtained on the orientation columns in area 17. They placed goggles over a kitten's eyes that presented vertical stripes to one of its eyes and horizontal stripes to the other eye. After the goggles were removed, the eye that saw vertical stripes responded only to vertical lines while the other eye responded only to horizontal lines. In addition, there were almost no binocularly driven neurons in area 17, whereas these are usually in the majority. Although this result has been demonstrated many times, when Maffei (1979) tried this experiment with a sinusoidal grating he found almost opposite results. He found that other orientations of gratings still gave a response, while the response to the shown grating was weakened.

Considering the vertical arrangement of axons and of the orientation (or ocular dominance) columns, it is possible that strong activity in the orientation columns could build up strong cycling activity and gradually modify the surrounding neurons. On the other hand, the sinusoidal grating seems to activate cells lying in a horizontal band
Maffei (1979). The configuration of responding neurons, in this case, should not reinforce each other, due to the limited horizontal spread of the axons. The repeated firing of a neuron, without its becoming involved in cycling activity, may lead to its habituation.

In a somewhat different experiment, Metzler and Spinelli (1977) presented kitten's with stripes whose orientation was fixed with respect to gravity. This experiment also disrupted the kitten's' orientation columns although many orientations were seen. A majority of the cells became sensitive to a fixed tilt, although normally the fraction is less than five percent. Again, cycling activity may have caused sensitization, however, this time, to the vestibular input.

The discussion in the iconic memory section suggests that the usual response to a prolonged exposure to a particular quality is a long lasting habituation. Older animals do not exhibit these physiological changes to deprivation and may become habituated to the goggles. The kitten's are still in a period of neuronal growth during this period. It is plausible that as was proposed for LTM, the place where the greatest synaptic growth occurs is in the neurons with the greatest firing rate.
4.1.1- Introduction

The following summarizes the series of events which the author has proposed including in a model of long term memory (LTM):

1- The incoming signal is separated into figure and ground on the basis of qualities and boundaries extracted from the scene. This is accomplished by the interneuron processing in area 17, leading to the turn-on of a small number of active circuit cores.

2- Visual input may turn-on active circuit cores, whose integration into an active circuit results in recognition and/or behavior. Specific aspects of behavior are mediated by different portions of an active circuit, lying in different nuclei.

3- The creation of new active circuits is the basis of LTM storage. This creation is proposed to result from synaptic modification at sites of high activity, under special circumstances. New active circuits are modifications, or additions, to existing active circuits, rather than new entities. This is illustrated by the discussion of creative forgetting in section 3.4.

There are several things that need to be added to the model of LTM. First would be a discussion of the sequence of events in various nuclei that lead up to placing the visual system in the recording mode.
Another is a discussion of where do past experience and sensory input interact, and how would this cause synaptic modification to start and end.

In order to answer these questions, this section relates the part of the visual system that supports complex scene analysis, called the neural scene analysis (NSA) system, to an optical holographic system. Interest in the similarities of these two systems began in the early 1960's with a series of papers published by various authors. Each stated his belief that a neural analogy to optical or acoustical holography could serve as a model for information storage and retrieval in the brain. The similarities between memory and holography were first mentioned by van Heerden in a 1963 paper about the holographic storage of information in solids. While van Heerden was primarily interested in using holograms as a memory for an artificial intelligence device, he was also aware of the biological implications. He noted that a neural model presented by Beurle (1956) could serve as a basis for an analog of a hologram in the nervous system. The list of researchers who contributed to building the analogy between brain function and holography include Julesz and Pennington (1965), Longuet-Higgins (1968), Gabor (1969), Westlake (1970), Baron (1970), and Cavanagh (1972). Details of a physiological basis for 'neural holography' were presented in a series of articles by Pribram (1966, 1969, 1971, and Pribram, Nuwer & Baron 1974).
The use of an analogy is a different type of modeling than those discussed so far. Function from structure modeling approaches generally start with various simplified models and combine systems supporting simple behaviors into one capable of more complete ones.

The analogy is an alternative to this step by step build up as a prelude to more rigorous study of the higher brain functions. The characterization of the systems to be considered depends on the desired information. When the desired information about a system is identification of the input-output (I/O) pairs, a perfect model is a second system with identical I/O pairs. An instance of two systems that are quite different structurally, but have the same input and output pairing is found in algebraically equivalent systems (see Figure 4.1). Padulo & Arbib (1975) give an extensive discussion of the variety of systems that can be specified by the same input-output (I/O) pairs.

The nature of the relationship between the neural scene analysis (NSA) and holographic systems remains unsettled. A general relationship connecting any two reasonable models is what Zeigler (1978) calls an I/O function observation morphism. This type of morphism is illustrated in Figure 4.2*.

There are two separate systems represented in Figure 4.2*. The upper part of the figure is concerned with the holographic system. Its components in this model are an input space with elements \( u \), an output space with elements \( y \), and a function \( f \) which defines a map from
**FIGURE 4.1A**

A TIME VARYING LINEAR SYSTEM

**FIGURE 4.1B**

A CONSTANT LINEAR SYSTEM THAT IS ALGEBRAICALLY EQUIVALENT TO THE SYSTEM IN FIGURE 1A

**FIGURE 4.2**

I/O FUNCTION OBSERVATION MORPHISM
**FIGURE 4.2A**

I/O FUNCTION OBSERVATION MORPHISM OF THE RECORDING-SUBSYSTEM

**FIGURE 4.2B**

I/O FUNCTION OBSERVATION MORPHISM OF THE RECONSTRUCTION-SUBSYSTEM

In all figures, —— stands for transforms, while ——— stands for mappings.
elements of the input space into elements of the output space. The lower part of the diagram represents the visual scene analysis system. In this case the input space has typical element \( u' \), the output space has elements \( y' \), and the function \( f' \) transforms input elements into output elements. The I/O function observation morphism consists of two mappings between these systems and an equation that they satisfy. The map \( g:u' \rightarrow u \) takes elements of the input space of visual scene analysis into elements of the input space of holography. While the map \( k:y \rightarrow y' \) takes elements of the output space of holography onto elements of the output space of visual scene analysis. An I/O function observation morphism is said to exist when the diagram in Figure 4.2* commutes, that is when any input element \( u' \) is taken into the same output element by any allowable path. Mathematically stated, the morphism exists when \( k(f(g(u')))=f'(u') \) for all \( u' \). Note that the algebraic equivalence mentioned above is also an I/O function observation morphism, with the \( g \) and \( k \) maps equal to identity.

There are two ways to characterize the elements in the input space of holography. One may either place an arbitrary label on every optical wave or one may give the spatial distribution of the electric field vector in a plane. Similarly, the elements in the output space may be characterized either by a label, or by a spatial distribution of reflectance. The distribution contains more information than the label in both the input and output spaces. The holographic transition function, \( f \), may either associate labels or be a transformation between spatial distributions.
In this section, a morphism based on labels for the elements in the input-output (I/O) spaces of holography and NSA is demonstrated. The utility of the morphism approach being considered is that the existence of the morphism implies \( f' = k \cdot f \cdot g \). There is a discussion of what is known about the spatial distributions to be associated with the labels for input or output space elements. This knowledge is complete in the holographic case, but not in the NSA case. The specific distributions of neural records and of neural waves are unknown. However, certain comparisons can be made between the distributions of elements related by the label morphism.

In the next section, the holographic transfer function, \( f \), is derived based on properties of optical waves. This requires introducing state variables in addition to the I/O variables used here. There, an attempt is made to locate the substrate for processes composing the NSA transfer function, \( f' \), which are analogous to those composing \( f \).

The plan for finding the desired I/O function observation morphism consists of three steps. First is the characterization of elements in the input and output spaces of both systems. Second, simple mappings for \( g \) and \( k \) will be proposed. Lastly, it will be argued that the diagram in Figure 4.2* commutes with these choices. The mappings \( g \) and \( k \) will be based on the work of previous investigators of the hologram-brain function analogy.
4.1.2- The basic I/O pairs in a holographic system

This subsection will identify the input and output space of the holographic system, while the next subsection will do the same for the visual scene analysis system. The stage will then be set for finding mappings between the input and output spaces of each system.

Holography consists of two distinct stages, recording and reconstruction. Each stage can be treated as a separate system with its own input-output pairings. Figure 4.3 shows one possible physical arrangement for both subsystems in a holographic system. 4

The recording subsystem of holography—The inputs to the recording subsystem of holography as shown in Figure 4.3A are:

a) an object wave (which is an optic wave as discussed in section 2.4) of coherent monochromatic light coming from the object of interest;

b) a reference wave, also coherent and monochromatic, with the same frequency as the object wave.

Both of these waves must simultaneously converge on the recording medium or hologram, as shown in Figure 4.3A.

---

4- In this paper, examples of holographic systems will all come from optical holography, although essentially all the statements hold for acoustic (Hildebrandt & Brenden 1974) or computer generated holography as well (Collier, Burckhardt & Lin 1971).
FIGURE 4.3A
A HOLOGRAPHIC RECORDING SUBSYSTEM

FIGURE 4.3B
A HOLOGRAPHIC RECONSTRUCTION SUBSYSTEM

In Figures 3A and 3B, lines with arrows represent traveling wavefronts within an optic wave and dashed lines represent the edge of optic waves.
The output of this subsystem is the recording, made on the hologram, of the interference of these two input waves. The recording medium is a photographic plate or any thin, light sensitive material. The developed plate is used in the reconstruction stage (see Figure 4.3B).

The reconstruction subsystem of holography—The input to the reconstruction subsystem of holography is a reconstruction wave, generally identical to the reference wave, used in the recording stage. This input is sent through a diffraction screen, made from what was recorded on the hologram in the preceding stage, as shown in Figure 4.3B.

The output of this subsystem is the part of the diffracted reconstruction wave that is identical to the object wave when it reaches the image plane. Mathematical details of both stages of holography are given in section 4.2.

4.1.3—The I/O spaces of the neural scene analysis system

The NSA system, like the holographic system, consists of two subsystems, corresponding to two stages (see figure 4.4). An experience of a scene is stored in one stage, and recalled in a second. The components of the inputs to the recording subsystem of NSA are the flow of information from the retina and the flow of internally-generated information. It is assumed that both of these flows turn-on active circuit cores. In recognition of this result, the flow of information
FIGURE 4.4: A BLOCK DIAGRAM OF THE VISUAL SYSTEM

The nuclei in the diagram connected by the heavy solid lines form the pathway of the retinal neural wavefront. The nuclei connected by the light solid lines belong to the eye movement effector pathway. Those connected by the heavy dashed lines are involved with the storage of information, with the LP, PUL, SSS, areas 18 and 19, and possibly IT as the sites of storage. These latter nuclei should also be the source of reference neural wavefront. The hippocampus, PRF, and the entorhinal cortex participate in the activation of the recording mode. The main difference between the recording and reconstruction stages is the processing that takes place in the latter pathway. Those connected by the light dashed lines probably belong to a pathway that controls the cognitive programming of visual activity. The following abbreviations are used: IT=inferotemporal cortex, LGN=lateral geniculate nucleus, LP=lateral posterior nucleus, PUL=pulvinar, SC=superior colliculus, OMN=oculomotor nuclei, PRF=pontine reticular formation, SSS=Supersylvian sulcus. (Mainly adopted from Graybiel 1974, Gross 1974)
from the retina will be called the **object neural wave**, while the flow of internally-generated information will be called the **reference neural wave**. The existence of neural waves generated by sensory input was shown to have considerable physiological support in section 2.4. These waves are analogous to the object wave in the holographic system. A more controversial question concerns the existence of internally-generated neural waves analogous to the reference waves of the holographic system.

There is some physiological evidence supporting the existence of such internally-generated waves. John (Thatcher & John 1977 p.234) has shown that if an expected stimulus does not arrive when expected, an evoked potential of the size usually elicited by that stimulus appears in the same nuclei. So, a neural wave, basically equivalent to the expected one, is internally generated somewhere. Perhaps the best direct evidence of internally-generated waves is the circulating neural activity found by Verzeano (1973) in the thalamus and cortex of cats in slow wave sleep, waking, and paradoxical sleep states. These circulating activities are modified by interaction with sensory input. It seems the internally-generated neural waves exist and neural effects of a stimulus are modified by past experience somehow. But it is not yet known whether the information from past experience that affects the sensory neural wave is in fact delivered as a neural wave. In section 4.2, a way in which part of the active circuit can supply reference information is proposed.
The output space of the neural recording subsystem—The neural recording is a set of synaptic modifications that leads to the creation of new active circuits and active circuit cores, as described in subsection 2.3.5. These are hypothesized to contain the new information from the recorded scene.

The input and output spaces of the recall subsystem—The input space of the reconstruction, or recall, subsystem of the visual scene analysis system has elements that are identical to the internally-generated neural waves used by the recording subsystem. These inputs can be considered to be a neural reconstruction wave analogous to the reconstruction wave in optical holography.

Neural reconstruction waves travel to the neural hologram plane in the same way the identical wave did in the recording subsystem. The firings triggered off in the hologram plane by these waves will lead to visually induced behavior in the animal. It is this behavior, including eye movements, that will be taken as the output of the recall or reconstruction subsystem of NSA.

4.1.4—The morphisms

The recording subsystems—The nature of the elements in the two-dimensional input space of the recording subsystem of holography were described in subsection 4.1.2. In subsection 4.1.3, a description of the two-dimensional inputs to the recording subsystem of visual scene
analysis were described. Therefore, the map $g$ between them can now be specified (see Figure 4.2A) Define $g$ as the two dimensional map that takes each ordered pair of retinal and internally-generated neural waves to a specified ordered pair of coherent optical waves. Actually, a set of coherent optical waves is needed to produce a colored scene (Collier, Burchhardt & Lin 1971 p.494). Also note, that the scene must be shorter than the coherence length of the optic wave in the direction of propagation. The first optical wave of the pair could be any arbitrary one of those capable of giving rise to the retinal neural wave. The other optical wave should be one that led to the original recording of the internally-generated neural wave in some set of nuclei.

The map $k$ takes the elements of the output space of the recording subsystem of holography to the output space of the recording subsystem of visual scene analysis, and can be considered a map between two records of the same scene. Define $k$ as the map which sends the record of the interference of two optical waves in holography into the neural recording which originated with the same two optical waves. The existence of a retinal and an internally-generated neural wave resulting from any optical wave presented, and the possibility of recording them was postulated in the previous subsection. It is also known that mammals can distinguish a large class of scenes. Combining these ideas implies that different neural waves are generated and can be recorded by each optical scene that can be recognized. This statement is equivalent to claiming that both the $g$ and $k$ maps must exist.
What can be known about the distribution of NSA system elements is summarized as follows. The spatial distribution of a retinal response is quite similar to the spatial distribution of the maximum amplitude of the initiating optical wave. Recall that the object neural wave is not defined until an active circuit is turned-on. The distribution of active circuit columns in some nucleus cannot yet be recorded. The spatial distribution of the internally-generated neural wave presents similar problems in measurement. In addition, separating these neural waves requires knowing something about the function of the nuclei that support them. The distribution least likely to be found by measurements is the specific set of synapses modified in the formation of a LTM.

The reconstruction subsystems—A model of the neural reconstruction system is shown in Figure 4.2B. A visual input may turn-on a neural wave which can act as a reconstruction neural wave. Similarly, an internally-generated input can turn-on a neural wave. Note that the same neural wave which acts as an object, or reference neural wave, can also act as a reconstruction neural wave. Defining the $k_1$ map on the output spaces of the reconstruction subsystems presents a problem. The output of the neural system is a behavior rather than a wave. The output mapping for the reconstruction subsystems must take an object wave in the image plane into an appropriate behavior for that same object wave. For the $k_1$ map to exist every recordable scene must elicit a distinct behavior when 'illuminated' with a recall neural wave.
4.1.5- Comments about the spatial distributions associated with the $g$ and $k$ maps

There is some neurobiological data that limits the set of functions that might transform the distribution of elements in the holographic recording output space into distributions in the NSA recording output space. It is these findings, together with the interaction of the retinal input with past experience, that originally attracted people to the holographic analogy. Pribram's articles on the brain function-holography analogy contain a collection of data from his experiments and those of others. In them, he notes similarities between the nature of the output space of the neural recording subsystem and that of the holographic recording subsystem. An expansion of his list includes:

1) **There is a wide distribution of information coming from each point of the object plane.** In holography, light from each point on the object reaches the entire hologram. The wide distribution in the nervous system is seen in the similarity of EEG waveforms in different nuclei to each other in response to a visual input (Thatcher & John 1977 p.158).

2) **Distorted objects can still be recognized.** Note that not all aspects of the model proposed depend upon the holographic analogy. This response includes the production of the same EEG waveform in the visual system nuclei as early as the LGN (Thatcher & John 1977 p. 242).
3) Removal of a piece of the developed recording medium results in a loss of resolution rather than the loss of part of the visual scene in the reconstruction stage. In holography, when a large piece of the developed hologram is cut away and a reconstruction wave is sent through what is left, the complete scene appears in the image plane, although somewhat distorted. In the nervous system, lesions of the inferotemporal cortex (IT) result in recognition ability deficits, although an object is still known to be in the field of vision (Gross 1974).

Note that this statement does not apply to areas involved in early visual processing like the retina, LGN, and area 17. Damage to these areas should be considered removing part of the reference source and using what is left as part of the reconstruction wave source, rather than removing part of the 'neural hologram'. Damage to areas in the 'holographic' visual scene analysis system should lead to specific problems, or syndromes of the higher brain functions (Luria 1973).

4) Information about the viewer position, relative to the scene seems to be lost in the transformation resulting from traveling to the recording plane. In holography, this result is due to the shift property of the fourier transform which the retinal wave undergoes on its way to the hologram plane. A displacement of the scene in the object plane is transformed into a phase change in the hologram plane. For the nervous system, this is based upon the use of non-egocentric information in constructing cognitive maps.
These data tell us something about the distribution of information on the records made in the NSA system. They are widely distributed, in some sense redundant and resistant to damage, and independent of shifts in viewer position. Each of these facts is also true of optical waves, though probably for different reasons. One difference between neural and holographic recording is that instead of changing the response across the recording medium, the neural input selects the set of columns which will act as the recording medium. Also, the neural record may partition a scene, based on different qualities, and store the separate qualities in different locations of the same active circuit.

4.1.6- Discussion and conclusions

The model of NSA presented was based on a number of assumptions:

a) The neural wave concept must be valid.

b) The system can be split into recording and recall subsystems, which can be treated independently. A possible objection to this assumption is that due to the length of time required to make a recording, the area in which the recording is taking place must be simultaneously used for recall. However, the active circuit concept implies only a small number of synapses need to be changed to make a recording. Thus, simultaneous recording and reconstruction, involving the same column or neuron, should be possible.

c) A neural record is made of the interactions of a neural wave from the retina with an internally-generated neural wave. Several things are implied here. First is that the neural information used in the NSA
system can be characterized by the neural wave equation. The validity of this was discussed above. Second is the existence of internally-generated waves. It was argued that this is in agreement with observation. Also implied, is that the internal neural wave must interact with the retinal neural wave and play a role in the recording process. A hypothesis for the role that this interaction plays in neural recording is elaborated in the next section.

d) Recognition of a scene and the appropriate response that follow are due to the appearance of a particular waveform, in a particular set of neurons, i.e. the turn-on of a particular active circuit.

The principle conclusion is that we are able describe an I/O function observation morphism for the recording subsystems of holography and NSA, as well as one for the recall or reconstruction subsystems in subsection 4.1.3. The value of having established an I/O function observation morphism between our model of the NSA system and holography is limited by the difficulty of finding functions that perform the same element matchup as the mappings, and by lack of knowledge of the distribution of the internal neural waves and of the distribution of the neural record. Similar problems arise in connection with the details of the behavior that a reconstruction neural wave will initiate.
4.2- THE STATE TRANSITION FUNCTIONS OF THE HOLOGRAPHIC AND NEURAL SCENE ANALYSIS SYSTEMS

4.2.1- Introduction

This section examines the processes from which the holographic transition function, $f$, and the neural scene analysis (NSA) transition function, $g$, can be derived. In both systems, the recording subsystem performs two distinct processes. First is the spatial transformations of both the object and reference waves in traveling from their respective sources to the hologram plane (see Figure 4.3). This process was discussed in section 2.4. Second, is the recording of information contained in the interference pattern due to the interaction of the two input waves on the hologram plane. As in the recording subsystem, two processes are carried out by the reconstruction subsystem. Again, there are the spatial transformations of an optical wave, in this case of the reconstruction wave, as it travels between planes. In addition, there is the modification of the reconstruction wave as it crosses the transmittance in the hologram plane. In the NSA system, the transition functions for each subsystem can be decomposed into the same pairs of processes as in the holographic system. However, these processes are considerably more complicated than in the holographic case. It is reasonable to hope that the way these processes are performed in holography will tell us something about the properties to look for in the neural system.
Since the full holographic and neural models will include details about the internal conditions of the systems, a state space is needed for each. The \( f' \) and \( f'' \) transforms, which map from input to output, will be expressed as a pair of functions involving the state variables. In general, a model of a system with state variables represented by the vector \( \mathbf{x} \) is governed by two functions:

a) the state transition function \( \mathbf{d} \) defined by \( x(t+1) = \mathbf{d}[x(t),u(t)] \)

b) the output function \( \mathbf{q} \) defined by \( y(t) = \mathbf{q}[x(t)] \).

Note that combining \( \mathbf{d} \) and \( \mathbf{q} \) would allow \( f \) to be calculated.

Figure 4.5 illustrates an I/O systems morphism for the recording subsystems. This morphism involves three mappings.

1- The map \( \mathbf{g}:u'_i(t)\rightarrow u(t) \) takes pairs of neural waves in the input space of neural recording subsystem into pairs of optical waves in the input space of the recording holographic subsystem.

2- The map \( \mathbf{k}:y\rightarrow y' \) takes records of the interference of optical waves in the output space of the recording holographic subsystem onto records in the output space of the neural recording subsystem.

3- The map \( \mathbf{h}:x(t)\rightarrow x'_i(t) \) takes as yet unspecified elements of a subspace of the state space of the recording holographic subsystem onto elements of the recording neural subsystem state space.

Note that the maps \( \mathbf{g} \) and \( \mathbf{k} \) are the same as those defined in section 4.1. The I/O system morphism exists when the diagram in Figure 4.5A commutes.

In other words the following two equations must hold:

\[
\mathbf{h}(\mathbf{d}(x,\mathbf{g}(u')))) = \mathbf{d}'(\mathbf{h}(x),u')
\]

\[
\mathbf{k}(\mathbf{q}(x)) = \mathbf{q}'(\mathbf{h}(x))
\]
FIGURE 4.5A
THE I/O SYSTEM MORPHISM RELATING THE RECORDING SUBSYSTEM

FIGURE 4.5B
THE I/O SYSTEM MORPHISM RELATING THE RECONSTRUCTION SUBSYSTEMS

SEE SECTION ONE FOR THE DEFINITION OF THE VARIABLES AND THE LOCATIONS WHERE THEY ARE MEASURED.
Figure 4.5B illustrates the I/O systems morphism for the reconstruction subsystems. The $g_1$ and $k_1$ maps are the same as defined in paper section 4.1, and the map $h: x_1(t) \rightarrow 1'(t)$ is a map between the state spaces of these systems.

The mathematical description of the processes in the subsystems leads to a natural specification of the system state variables in both the recording and the reconstruction subsystems. The approach advocated in this section for deriving these descriptions involves studying the consequences of a set of optical wave properties. From these, mathematical descriptions of the processes of the recording and the reconstruction subsystems of the holgraphic model can be derived. These descriptions specify which input is transformed into which output in each subsystem, and thus they determine the functions $d$, $g$, $d_1$, $g_1$. Next, the analogous properties for the NSA system will be studied in the light of the understanding of the holographic system. The elements included in the state space are those required to support these properties, and the governing functions, $d'$ and $g'$, describe the dynamics that the properties lead to in the neural model of recording. The last step involves considering the relationship of these expanded models of holography and NSA, namely the validity of the I/O systems morphism shown in Figure 4.5A. The same general treatment will be given to the reconstruction subsystem (see Figure 4.5B).
4.2.2- The state transition and output functions of holography

Rules which govern a well validated behavioral model of the holographic system can be derived from the properties of optical waves. Here, an optical wave in a plane is defined as an electrical disturbance whose amplitude obeys the following equation over a specified space and time domain:

\[ w(x,y,t) = a(x,y)\cos[ft + p(x,y)] \]  \hspace{1cm} (2.7)

There are four basic properties of optical waves, described by equation 2.7, on which all the processes in holography depend: (the first two have already been stated in section 2.4.)

1- The spread of light from any point on an optical wave is described by Huygen's principle, i.e. each point on an advancing wave acts as a source of a circular wave.

2- The interaction of one wave with another is governed by the superposition principle, i.e. when two waves intersect they add at each point.

3- The intensity of interaction of a wave with a recording medium is given by the square of its amplitude averaged over time. This follows from the facts that the recording medium darkens in a linear fashion with the energy delivered, and that the energy in a wave is proportional to the time average of the square of its amplitude, i.e.

\[ E(x,y) = s(x,y,t)^2 dt \] (Born & Wolf 1975 p.10).
The interaction of a wave with a transmittance is governed by the Kirchhoff boundary conditions; i.e. a) a wave passing through the transparent part of a transmittance is the same as it would be if the rest of the transmittance were not there; b) there is no disruption of the electric field, due to the wave directly behind the opaque parts of the transmittance.

Next we will look at how these properties can be used to find the functions \( d \) and \( q \).

The transformation of optical waves—The transformation of optical waves in traveling between planes used here is given by equation 2.9.

\[
W(x,y) = \exp(ikz)\exp(ik/2z(x^2+y^2)/i\lambda z) \\
\iint W(x',y') \exp(-ik/z(x-x',y-y')) dx'dy' \quad (2.9)
\]

which is proportional to a Fourier transform. Figure 4.3A shows that in holography both the object and reference waves must travel to the hologram plane, and thus undergo transformations that can be described by equation 2.9. Since this is an important intermediate step in the recording process, our model will use the reference and object waves in the hologram plane as two state variables. Figure 4.3A also shows that these two waves overlap in the hologram plane, and thus by the superposition principle, the total wave is the point by point addition of the two state variable waves. This total wave in the hologram plane will be taken as a third state variable. For these state variables, equation 2.9 gives a method for calculating the state transition function \( d \) (see hologram model in Figure 4.5A).
As shown in figure 4.3B, the reconstruction subsystem also has two between-plane transformations. Equation 2.9 thus allows the reconstruction wave in the image plane to be found from the reconstruction wave behind the developed hologram, as well as allowing the reconstruction wave in the hologram plane to be found. Therefore, the two state variables chosen in reconstruction are the reconstruction wave just in front of the hologram plane and just behind it. Thus the output function \( q_1 \), plus one of the two variables in the state transition function \( d_1 \), can also be found using equation 2.9 (see hologram model in Figure 4.3B). Since Huygen's principle is only required in the transformation process, it will not be essential to the feasibility of a neural hologram. However, the particular form of the transformation is critical to the spatial distribution of information over the hologram during the recording stage of holography.

**Recording in the hologram plane**— Suppose that the distributions of two waves on reaching the hologram plane is known. Then the pattern recorded by the recording medium can be derived from the superposition principle and the square law \(^1\)

\[
E = \int_0^T (x_r + x_o)^2 dt = T \int_0^T \{ a_o \cos(ft + p_o) + a_r \cos(ft + p_r) \}^2 dt = T/2 \{ a_o^2 + a_r^2 + a_o a_r \cos(p_o - p_r) \}
\]

\(^1\) In an acoustic hologram it is possible to record the amplitude directly and almost instantaneously, so a hologram can be formed by multiplying the amplitude times a periodic electronic reference (Hildebrandt & Brenden 1974). Thus the time-averaged square of the amplitude need not be recorded for the recording to be of the form in equation 4.2.
when \( T \) is many cycles long. (The subscript on the \( w, a, \) or \( p \) will always indicate the wave with which it is associated: 1 = object, 2 = reference, 3 = reconstruction.) Although it is convenient to work with complex amplitudes, equation 4.1 is derived using the real amplitudes (Hildebrandt & Brenden 1974 p.13).

To see the distinct advantage that holographic recordings have over ordinary photographs consider the energy delivered to a point by exposure to a single wave.

\[
E = \int_0^T w^2 dt = \int_0^T a^2 \cos^2(\omega t + \theta) dt = a^2 \int_0^T (1 + 2 \cos(2(\omega t + \theta))) dt
\]

This is approximately equal to \( a^2 T \) if \( T \) is many cycles long.

The maximum amplitude information is preserved, but the phase information is lost in the averaging process. The problem is unavoidable in ordinary photography, since even the fastest optical recorders need many cycles before an image can be formed. Equation 4.1 shows that a hologram can preserve the amplitude, phase, and sinusoidal waveform of the object wave, which contains all of its information. (Actually, the point by point phase difference between the object and the reference waves are stored, but usually the phase distribution of the reference waves is known.) Given that the spatial distribution of the summed object and reference waves in the hologram plane, \((x_r + x_o)(t)\), is known from equation 2.9, Equation 4.1 gives a way to calculate the pattern that will be recorded in the hologram plane, i.e. \( y \). Thus, it defines the output function \( q \).
Modification of the reconstruction wave in the hologram plane—Suppose that a wave has the following amplitude distribution directly in front of the hologram $x_1(t) = a_1 \cos(\omega t + p_1)$. In addition, assume that the percentage of light allowed through the developed hologram is proportional to the energy to which it was exposed in the recording process (see Figure 4.3B). Thus the transmittance of the developed hologram, say $M(x,y)$, is proportional to $E$ in equation 4.1. Under the condition that $x_1(t) = x_r(t)$ the reconstruction wave just behind the hologram is, by the Kirchhoff boundary conditions, of the form $Mx_r(t)$, which can be expanded into the following three terms using equation 4.1.

$$Mx_r(t) = [a_o^2 + a_r^2]x_r(t)$$

(4.2A)

$$+ 1/2[a_r]x_o(t)$$

(4.2B)

$$+ 1/[a_r]a_o \cos(\omega t - p_o + 2p_r)$$

(4.2C)

Under these conditions it is seen that term (4.2A) is proportional to the reference wave $x_r$ in the hologram plane; (4.2B) is proportional to the object wave when it reaches the hologram plane; while equation 4.2C can be thought of as a distorted real image of the object. It is important to note that the reconstruction wave has given rise to a wave that is identical (within a factor $(a_r(x,y))^2$ which is constant in time) to the actual wave that would be coming from the object. Equation 4.2 is basically the diffraction transform in Figure 4.5B. It allows the second state variable, $x_{12}(t)$, of the reconstruction subsystem to be calculated from the other state variable or the input, assuming the between-plane transform is already known. Therefore, the state
transition function $d_1$ is defined as well as $d, q$, and $q_1$. Equation 4.2 gives a way to calculate the pattern that will be recorded in the hologram plane, i.e. $y$.

4.2.3- The state transition and output functions for the visual scene analysis system

In subsection 4.1.3, an argument was presented supporting the claim that the neural recording mechanism makes use of a sequence of non-egocentric (i.e. independent of viewer position) waves. It was also argued that a nonspecific recognition of novelty is used to initiate search and exploratory behavior. These behaviors would entail some extra steps, thus, requiring extra state variables, in the processes of the recording subsystem of NSA, as compared to holography. For example, the initial response to visual input that leads to the turn-on of active circuits on which the neural waves are defined also entails filtering and channel separation (see figure 4.4). The initiation of recording by recognition of novelty and the averaging over a sequence of neural events also makes the recording process more complicated than its optical counterpart. In the recall subsystem, complexity is added by the fact that a single recall seems to build up in a cascade-type effect that continues over many neural waves.

The neural wave transformations-- A discussion of possible neural wave transforms in going between various nuclei was given in section 2.4. Although a specific transformation was not found, the discussion in the
last subsection indicates that it should still be possible to study the neural record in active circuits using transform methods.

Several aspects of the NSA model that should probably be included in the set of state variables of both subsystems have been mentioned. One such variable is the retinal or object neural wave in area 17, \( x_{oI}'(t) \). Also, the reference neural wave at the site from which it leaves to be compared with the object neural wave, \( x_{rI}'(t) \), (which we assume does happen) should be a state variable.

**Recording a neural hologram**—Synaptic modification was proposed as the mechanism for recording neural waves. Here, the type of information that is stored and the characteristics of the neural wave to which the recording tissue reacts are the main concern.

As can be seen in equation 2.7, the information in an optical wave maybe factored into two parts: the amplitude information and the dynamic information. Both of these factors depend on position and are specific to a given scene, but in the dynamic factor these dependencies only appear through the phase.

The model of neural information dispersion represented by equation 2.8 also has two factors. The amplitude of the neural wave depends on the particular event and position, as does the optical wave's maximum amplitude. However, the situation is different for the dynamic factor of the neural wave. Here, both the form of the function \( F(.) \) and the phase vary with the particular scene presented. Therefore, complete
information about a particular neural wave must include the form of F(·) as well as the phase and amplitude distribution in some nuclei. Note that this may be equivalent to knowing the set of columns participating in the active circuit. The ability to recall a scene, and in fact to internally reproduce the EEG in many nuclei that was present during viewing the scene (John 1974), leads us to expect that complete information about neural waves is being stored in the visual system.

A qualitative model, consisting of seven steps for the storage of information from a scene, are presented below. The first two steps characterize the transformations of neural waves in traveling to the hippocampus, where presumably the storage process is initiated. These steps summarize the discussion in subsection 2.4.4. The last five steps occur during neural storage of the information in a scene. These steps include new assumptions, based on physiological data presented throughout this chapter.

1- Visual information from the retina is filtered and reorganized in traveling to area 17. There, the scene is decomposed into a number of 'homogeneous' regions. Various categories of visual information are extracted from each of these regions in the scene in the rest of the visual cortex.

2- The IT receives input from each of these information channels, plus the original scene decomposition from area 17, and synthesizes them into a non-egocentric map. The output of the IT can lead either to the recognition of the scene when an active circuit is turned on, or to the recognition of novelty in the scene when it is not.
3- The hippocampus would react to the overall differences between the object wave, coming from area 17 through the IT, and the reference wave generated within nuclei of the categorical storage system, discussed in subsection 2.4.4. These combined waves could be averaged over the whole event by the mismatch cells to decide if the form of the $F(.)$ functions for the two neural waves match. The model predicts that the hippocampus would not initiate recording if a full active circuit formed. Thus, the inability to turn-on an active circuit combining the retinal and internal input must trigger a reaction in the hippocampus. Attention must be focused on the retinal input, and the expectation about the retinal input carried by the internal wave must be compatible, in order for an active circuit to form. The model requires the two input waves to converge in several nuclei, including the IT, and one with direct inputs to the hippocampal mismatch cells. The entorhinal cortex receives input from the IT and some thalamic nuclei, and it projects to the hippocampus (O'Keefe & Nadel 1978 p.125). It is therefore a plausible site for the interaction of the neural input waves.

4- From the role of the hippocampus in using and forming synthesized scenes arises a proposal for its effect on neural activity. The hippocampus would appear to be the key to bringing about the increased level of concentration required to allow the turn-on of a full active circuit. Most of the time the turn-on of the full active circuit would involve using recording to extend a synthesized scene represented by an active circuit. So, the hippocampus may be responsible for initiating behavior leading to new active circuit
formation and triggering synaptic modification, possibly through the midbrain.

When the hippocampus recognizes novelty, it sends out signals to initiate search for novelty and then explore it, as well as a signal to start recording. The signal to record probably goes through the PRF, while the search is probably directed by the frontal cortex. Damage to the hippocampus leads to deficits in exploratory behavior in both animals and humans (O'Keefe & Nadel 1978). The frontal cortex seems to play a crucial role in monitoring sequences of behavior (Luria 1973 p.187).

5- These processes lead to recording complete information about the novel objects. The first step is the search to locate these objects. Once this is accomplished, exploration continues until a non- egocentric map of the object is built. This map ties together the visual information from various channels plus the verbal categories into which it fits. As this occurs, the spatial distributions of a and p are stored and the final form of F(.) for the object develops as the new active circuit evolves. Thus, the final internal map of the object stores complete information about it, including the form of F(.). At the same time that exploration and non-egocentric map building are going on, behavioral responses to the novel object are being tested.

6- Eventually, the new active circuit is complete and any new active circuit cores needed are formed, and the input will no longer initiate recording.
7- Any neural recording entails changing only a few synapses on a sparse set of cells in any nucleus, so recording a single neural event should not interfere with any other neural hologram. However, if a neural hologram is not reinforced through use, the continuing recording of new holograms may eventually distort the hologram. (Interestingly, many scenes can also be recorded on an optical hologram, although reconstruction cannot start until the recording stage is finished.)

The list of state variables should now have a number of things added to it. These include: the combined object and reference neural wave in the entorhinal cortex, \((x_{r1}+x_{o1})'(t)\); the time averaged difference of the two waveforms in the hippocampus, \(z_1'\); whether the record mode in the PRF is activated, \(z_2'\); and the sequence of neural waves in the storage sites as exploration goes on, \(x_{e1}'(t)\).

The process outlined here can be put in better perspective by looking at a well studied example of the treatment of novelty, the conditioned response. Changes in the evoked potential (EP) response to the conditioned stimulus in various nuclei follow a general pattern over the course of conditioning (Thatcher & John 1977 p.158). There is an increased response to the conditioned stimulus in many nuclei, including those in other sensory modalities, and an increase in the degree of similarity in the shape of the EP in the affected nuclei. This should mean the conditioned stimulus is triggering both relay cell sensitization and the recording mode. The similar waveform seen in many nuclei is taken to indicate a turned on active circuit is present.
These electrophysiological changes accompany a gradual improvement of the probability of success in the given task. After the conditioned response is well established, further exposures to the conditioned stimulus cause the EP to shrink back to almost the original size. The EP in all but the specific sensory nuclei become very small. However, even as this reduction of the EP response is taking place, the performance of the task remains almost perfect. Apparently, once the animal recognizes that there is something unusual about the conditioned stimulus, it begins to draw more attention. The build up of the EP response would correlate with the exploration and the construction of new active circuit cores and a new active circuit. The gradual shrinking of EP extent seems associated with the slow construction of an active circuit core and its incorporation into the active circuit for the scene. Once an active circuit core is constructed, further sensory input leads to an appropriate response without additional concentration.

The reconstruction stage of neural holography—The way in which a neural wave is affected by passing through a nucleus (transmittance) should determine the output of this subsystem. It is interesting to note that a neural equivalent of the Kirchhoff boundary conditions determines the interaction of a neural wave with a nucleus through which it passes. This is due to the lack of interaction between the weak electric fields associated with the potentials of two adjacent neurons in a nucleus.
In the previous subsection it was argued that appropriate behavior develops in parallel with recognition. Therefore, the path that the reconstruction wave will travel after it leaves the hologram plane is part of the newly formed active circuit. This is not very surprising since neural waves do not travel freely through space.

The state variables in the reconstruction subsystem of visual scene analysis would be essentially the same as those in the holographic reconstruction subsystem. The optical reconstruction wave in front of the hologram plane would correspond to the neural reconstruction wave in the visual cortex, which is denoted by $x_{11}'(t)$. The optical reconstruction wave, after passing through the hologram, would correspond to the neural reconstruction wave after it has passed through the sites where modification took place in the first stage, which is denoted by $x_{12}'(t)$.

A preliminary hypothesis about the neural activity related to recognition is implicit in the holographic model. It is postulated that recognition occurs when the object and reference neural waves forming in response to a scene are compatible. This condition leads to the turn on of an active circuit of columns, centered in some combination of the
visual cortex, IT, LP, or pulvinar. The specific sites involved in the active circuit determine the experience.

4.2.4— Conclusions

The objective in undertaking a rigorous comparison of holography and visual scene analysis is to model an inaccessible, hard to measure system with one that is far more tractable. For holography to be a direct model, the black box function of holography, \( f \), and that of NSA, \( f' \), would have to be the same spatial transformation. The functions \( f^1 \) and \( f'^1 \) would also have to be essentially the same (see figure 1). Whether the internal states that occur in the black box for each system are the same is not important to this objective. However, finding equations which describe the changes that the internal states undergo may be the only way to specify the black box functions.

To begin, the results from section 4.1 about the holographic analogy to NSA are summarized. The I/O spaces of the two systems have the following similarities. First, each consists of two independent stages of activity, recording and reconstruction. Second, in the recording subsystem of holography or visual scene analysis, a permanent recording of the information in an optic wave is the output. The output of either reconstruction subsystem is the response in the image plane that would be expected if the object information, instead of the reconstruction information, was contained in the input. Third, the input to the recording subsystems of each must consist of two signals, one an object and the other a reference signal.
Beyond these superficial similarities, what must be shown in order to use holographic techniques on the neural system is that the neural input signals have certain critical properties, and that it makes sense to transform these signals. Introducing the concept of a neural wave makes it possible to define each of the maps shown in Figure 4.2. These maps allowed a demonstration that an I/O function observation morphism exists between each set of subsystems. In other words, the mapping identities \( f' = k \cdot f \cdot g \) and \( f'' = k_1 \cdot f' \cdot g_1 \) are valid. Because of the inaccessibility of the CNS, it is not yet possible to measure neural waves in internal nuclei or neural recordings. This precludes finding functions that perform the same matchup as the \( g \) and \( k \) mappings. Therefore, this morphism cannot be used directly to calculate the function \( f \) or \( f' \).

The hypothesis that \( f \) and \( f' \) (and \( f_1 \) and \( f''_1 \)) are essentially the same was checked in this subsection. To do this required an understanding of the internal processes underlying all four black box functions. It was argued that the same general processes leading to holographic recording (i.e. the transformation of waves traveling between planes, the interaction of waves, and the permanent changes caused on the recording medium by the combined waves) are appropriate descriptions for the processes leading to recording the complete information contained in neural waves during NSA. The comparison of the recording subsystem breaks down when the equations governing these three processes are considered. This is due to the numerous subprocesses which occur in neural, but not in optical recording. In these subprocesses the reference neural waves are not simply used to make an interference pattern with the object wave for recording. In fact, the
first thing the interaction of neural waves does in recording is to cause a recognition of novelty. Since the source of the reference neural wave must also be the place where the neural recording is done, the exploration process triggered by novelty leads to a progressive modification of the reference neural wave elicited by the object neural wave.

In addition to the differences between the equations controlling holography and those suitable for controlling visual scene analysis, the map \( h \) (see Figure 4.5A) between the two state spaces of the recording subsystems is hard to specify. The explicit mapping requires complicated comparisons between the state in a single holographic event, and the changing state in a sequence of neural events that lead to recording. Thus the I/O system morphisms shown in the diagrams in figures 4.5A and 4.5B could not be shown to exist, because appropriate maps \( h \) and \( h_1 \) could not be found.

The principle result is the proposed holographically motivated model of visual scene analysis described in subsection 4.2.3. It arose during the search for processes capable of duplicating the known processes of holography. If such processes take place, it was argued, they must occur at particular locations. For example, there must be a site where retinal input triggers a reference neural wave against which a transformed object neural wave is matched. A trigger for the recording mode must exist, a place where non-egocentric neural waves are formed must also exist. Locations for such activity are proposed in the model.
This analysis has shown that the black box functions $f$ and $f'$ (also $f_1$ and $f_1'$) arise in very different ways. Thus, the direct use of holography to model neural processing is not possible. However, the successful use the holographic analogy to guide the construction of the general neural wave model in subsection 4.2.3, indicates that the analogy has heuristic value in the construction of neural models of higher brain function. Note that a system theoretic model at the nuclear level has been used with success for the simpler oculomotor response system (Robinson 1975). While the system supporting NSA is larger and has more complicated spatial distributions of activity in its nuclei, such models might be workable.

In conclusion, the model suggests that specific rules, comparable to those of holography, govern the internuclear response to a scene, and indicates the type of information needed to formulate these rules. Rules of this type can provide constraints for models of lower level neural activity, including neural network models. They in turn are constrained by the particular behavior mode in which the organism is operating. Operating in a particular behavior mode entails controlling attention. The next chapter discusses the way attention relates to the four cell type model which has been proposed.
5- THE ROLE OF ATTENTION IN ACHIEVING RECOGNITION

Attention regulates the recognition process by controlling the selection of the input configurations that are perceived as figures, distinct from the rest of the scene. It also controls which quality (such as color) characterizes the figures selected. Attention has been compared to a beam of light which illuminates one portion of a scene (Kahneman 1973). Thus, the neural control of attention plays a role analogous to the holographic photographer in choosing the portion of a scene which is to be analyzed. In addition to regulating incoming information, attention also regulates the information coming from permanently stored past experiences. In chapter 3.4 it was proposed that there are two types of objects stored in long term memory, templates and synthesized scenes. Attention determines which experiences will reach awareness by activating particular templates to be incorporated into synthesized scenes. So, attention is the critical factor in choosing the correct context for interpreting a particular visual input.

The attributes of attention are considered before constructing the actual neural model of attention. The most obvious attribute of attention, and one that applies to all situations, is its limited capacity. A second attribute of attention is the division of its limited capacity among the qualities of a scene. Two other important attributes are the interference of attended tasks, and the number of attentional shifts required to complete a task.
To characterize the variability of the limited capacity, and its specific allocation to visual qualities between attentional shifts, the idea of dimensions of attention is introduced. These include: the intensity, the tightness of focus on each of the numerous bands in the attention channel (section 5.5), and the dwell time. The intensity determines how strongly the attended target stands out; it varies with the amount of arousal. The width of attention restricts the portion of each quality that contributes strongly to the emergence of figures. The dwell time, in vision, can be related to the intersaccadic interval. An additional dimension of attention is the amount of interference that attention to one signal causes on the simultaneous or subsequent allocation of attention to other signals.

The primary determinant of capacity is the amount of arousal (Kahneman 1973). This, in turn, depends on the demands of the current task. The portion of the attention capacity allocated, or the mental effort exerted for the task, does not increase linearly with the demand. (A possible reason for the association of mental effort and attention are suggested by the model of attention presented in chapter 5.3.) For an easy task, it seems impossible to be aroused enough to supply more than the attention needed for a particular task. Thus, one does not really become very aroused during an easy task, such as adding two single digit numbers, even under circumstances where doing this simple task is very important (Kahneman 1973 p. 15). The attention allocated for a difficult task may fail to keep pace with the demand for attention. This failure may be explained by experiments indicating that
FIGURE 5.1: SUPPLY AND DEMAND OF ATTENTION

See discussion page 188.
(From Kahneman 1973)
the capacity of attention and arousal are related by an inverted U-shaped function. Attention capacity increases with arousal, up to a maximum, and then declines as arousal continues to increase. Thus, for difficult tasks the arousal level may be so high that attention capacity is reduced.

The above discussion indicates that the operation of attention varies in different circumstances. The next chapter suggests a way to separate cases of attention allocation.

5.1- CASES OF PERCEPTUAL PROCESSING

The following discussion refers to the division of perceptual processing cases listed in Table 5.1. The six cases are formed from the intersection of two types of prior expectations, specific and no intentions; and three types of visual objects, complex, simple, and very familiar. The type of processing tends to move the value of each dimension of attention in a particular direction. The differences among the cases of processing can be illustrated by examples.

5.1.1- Simple, complex, and very familiar objects.

It is necessary to give more precise meaning to the types of objects that have been defined. First, consider a real scene, such as portrait of a group of men. As Norman (1969) points out, only four to seven items can be extracted from a single presentation of a scene, depending
<table>
<thead>
<tr>
<th>TYPES OF SCENES</th>
<th>SPECIFIC EXPECTATIONS</th>
<th>NO EXPECTATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMPLEX</td>
<td>Many steps before recognition (short)</td>
<td>Many steps before recognition, if it occurs at all (long)</td>
</tr>
<tr>
<td></td>
<td>Focusing required, based on initial instruction and information from the previous steps</td>
<td>Focusing required, based on previous steps</td>
</tr>
<tr>
<td></td>
<td>(Ex. Photograph shown after a question, such as 'what is the man doing?')</td>
<td>(Ex. Photograph)</td>
</tr>
<tr>
<td>WELL KNOWN</td>
<td>One step, focused on a non-spatial quality</td>
<td>One step, weakly focused attention</td>
</tr>
<tr>
<td></td>
<td>(Ex. Locating a red word, as instructed.)</td>
<td>(Ex. After a red word has been found, what is it?)</td>
</tr>
<tr>
<td></td>
<td>Several steps, weakly focused attention, unfocused spatial attention</td>
<td>Several steps, focused attention on a higher order template</td>
</tr>
<tr>
<td></td>
<td>(Ex. Find the word 'five' on a page)</td>
<td>(Ex. Reading)</td>
</tr>
<tr>
<td>SIMPLE</td>
<td>One step (shortest), focused and unfocused attention are the same</td>
<td>One step (short), focused and unfocused attention are the same</td>
</tr>
<tr>
<td></td>
<td>(Ex. Instructed to describe a simple geometric shape)</td>
<td>(Ex. Looking at a simple geometric shape)</td>
</tr>
</tbody>
</table>
on its complexity. Therefore, one of the figures initially selected from a scene may be further broken down after attention is shifted and focused. For example, the outline of a man is a very familiar object, and should activate a template in a single step. While a first glance might reveal only a group of men, a second look may discern that one of those men was missing a finger. In the discussion of iconic memory, it was noted that all the information in a scene is stored in almost photographic detail. The missing finger apparently can register as anomolous before recognition, and can both activate a template and select the target of attention in the subsequent attention shifts. The vast number of possible figures that are present in a scene are reduced to the four to seven that can be extracted by the process that leads to the emergence of figures. Thus, this scene is actually complex, but at each step of the way it is broken down into a number of objects, most of which are very familiar. A pure complex scene, such as an electron microscope photograph of nervous tissue would have no familiar objects to draw attention for the average person and figures could not be extracted without a great deal of processing. A simple object is one, such as an isolated geometric figure, that would draw attention whether it is familiar or not.

5.1.2- Cases with specific expectations vs. no specific expectations

The two types of expectations are actually two points in a spectrum ranging from expecting a specific item, through expecting classes of items, through vague concepts about distance and size, to no
preconceptions at all. Searching describes the use of specific intentions to guide the selection of figures and to focus the attention. Searching, in this paper, is defined as the act of consciously trying to locate a set of known qualities in a scene, and describes the use of specific expectations to guide the selection of figures and to focus the attention. Exploration, on the other hand, describes the selection of figures without intention; it is the process of locating and cataloging the unknown qualities in a new scene.

In the no expectation cases, i.e. exploration, attention must either be unfocused, or focused on very general qualities. The difference between the processing of simple and complex figures is discussed by Beck (1972). He had subjects examine three arrays of letters, most of which were 'T's, with the rest either 'O's, 'L's, or tilted 'T's. The first array based on the round quality of the 'O's was treated as a simple object with the 'O' region as the figure and the 'T' region as the background. Note that when an entire array is viewed, the attention must be unfocused. This experiment indicates that in one unfocused step visual processing leads to the same separation of regions that focusing the attention on the shape of each letter would have produced.

Beck also found that tilted 'T's are more discriminable than 'L's when the rest of the array is composed of upright 'T's. However, if there is only one letter, or if there is sufficient time to center the fovea on the array, tilted 'T's and 'L's are equally discriminable. When there is single letter, unfocused processing gives the same result
as focused processing. The focusing of attention after centering the fovea is based on general prior knowledge from the first unfocused viewing of the scene. If an extended period is allowed, then a synthesized scene type of long term memory can be formed from a sequence of focusings, each obtaining non-egocentric information about the objects in the scene. When there is a single letter, unfocused processing leads to the same figure emergence as focused processing.

There are indications that with very well known patterns, such as familiar words, unfocused discriminations may make use of operations as complex as semantic decoding. For example, real words tend to stand out when incorporated among strings of letters. Other evidence comes from an experiment where a pair of sounds was presented either to the same or different ears, and subjects were instructed to pick out one of the sounds (Treisman 1970). The ability to separate the two sounds was equally poor when nonsense syllables were used. However, when real words were used, the subjects were able to identify the sounds by focusing on only one ear. But subjects were not able to identify two real words when they were presented simultaneously to the same ear. The difference is presumably due to the difficulty of constructing a synthesized scene, compared to activating a template.

Turning from exploration to search, the attention directed at the same scene is found to differ. When searching for a well known object in a complex scene, it is not necessary to examine the entire scene item by item. In this case, increasing the number of objects in the scene
has little effect on search time. But when the target object varies from trial to trial each object in the scene must be examined individually. When the attention must be focused on each object in order to decide whether it actually is a target, the search time increases linearly with the size of the set (Schneider and Schiffrin 1977). Thus, the search for a well known target resembles the exploration of a scene with very familiar items in it, while the search for variable targets resembles exploring a complex scene. However, the difference is that the items, which emerge as figures in the case of search, may not be the ones which would have emerged if specific intentions had not been used. The values assumed by attention in the dimensions differ in search. The intensity of attention given to emerging figures is greater, the initial width of attention is narrower, and the dwell time is shorter; thus, less information is extracted per shift.

Beck (1972) found that under certain circumstances movement, proximity, texture, color, tilt and size allow a subject to locate a target object before being conscious of it. Similarly, Neisser (1967) found that subjects searching for several targets at once realized that they had found one before they knew which one it was. In other words, the visual system is made sensitive to particular qualities in the scene by intentions, so that unfocused attention can pick out targets in one step. Search strategies can cause moderately unimportant innate factors to become very important. Thus, past construction of templates plus momentary expectation cause moment to moment fluctuations in the process leading to figure emergence. An example of the case in which two
different intentions are used to direct the search of a scene was performed by Erickson and Collins (1969). They presented subjects with a rapid fire sequence of digits, and had them decide whether any of the numbers between 1 and 9 were missing. Each number in the sequence had to match a newly activated template exactly. When the subjects were told which digit might be missing, they were able to correctly identify the missing number, even when the data were presented at the rate of one digit per 50 milliseconds. However, without prior information, the subjects required at least 200 milliseconds between the digits. When the subjects were warned which digit might be missing, each number in the sequence could be compared to the same template to check for differences. Clearly, it is much easier to compare items for any difference rather than to check for identity when the items are quite different, and easier to maintain the same template rather than to switch to a new one after each signal. These differences could account for the greater recognition speed, when the warning was given. This example also supports the notion that greater expectations increase the amount of focusing.

The examples presented so far either had one step or many steps in the analysis of a scene. The next example gives a situation in which two steps are involved. Suppose that a subject is asked to find the capitalized or red word on a page, which involves two tasks and therefore two steps. First, the subject must search the field for a simple quality, then the area with the quality must be focused on and explored, hopefully leading to the emergence of a well known object in one step.
Adjacency errors provide one demonstration of the need for several hundred milliseconds in order to process information through the recursive use of quality templates. When a subject is asked to read the word shown in capital letters in a sequence of briefly shown words, the word that follows the capitalized one is often erroneously selected (Lawrence 1971). This indicates that the quality of capitalization is recognized before the word is analyzed. By the time the capitalization can cause attention focusing on an area, the original word is gone and the next word already commands attention. In general, cues that focus the attention take 100-200 milliseconds to become effective. Averbach and Coriell (1961) presented test subjects with a tone indicating where to look prior to the presentation of a complex tachistoscopic display. The cue was only effective after 100 milliseconds, or more, had elapsed between the cue and the display. The focusing of attention should require both time and effort.

There are some further differences in the nature of attention between the cases that are not distinguished in the table. In search triggered by a cue, the cue arrives a specific amount of time before the focusing of attention must occur. On the other hand, exploration of a complex scene after the initial presentation may be examined by the subject whenever he or she is ready. The search triggered by the cue must have a more complicated focusing program of these two cases, since it must account for the exact time delay needed before focusing attention. However, the complex scene may contain more overall information.
5.2- PREVIOUS MODELS OF ATTENTION IN THE EXTRACTION OF SENSORY INFORMATION

Previous models of attention have tried to incorporate some of the attributes of attention listed above. The limit on the amount of detail that can be attended at once is the characteristic which most early models attempted to explain. One type of model has assumed that certain operations carried out in localized regions of the brain limit the information that can be handled. Broadbent (1957) described a model of perception with three steps: sensory registration, perceptual analysis and response selection. He proposed that a bottleneck prevents all the information in sensory registration from being processed in the perceptual analysis step. There is a selective filter that allows a limited amount of information from iconic memory into a limited capacity channel, which passes information to structures which carry out perceptual analysis. Deutsch and Deutsch (1963) suggested that the bottleneck in Broadbent's three-step model of perception occurs before response selection. Neither of these models works in all situations and the position of the bottleneck may differ in different circumstances. One alternative to the bottleneck theory is the capacity theory (Moray 1967) which states that there is a general limit on the attention, or mental effort, available. According to the capacity theory, the distribution of available attention to simultaneous activities should be independent of their nature. On the other hand, the bottleneck theory holds that when the information from several different activities stored
in the sensory register must be processed by the same neural structure, there is interference among them.

5.2.1 Kahneman's model

Kahneman presents a model of attention during perception that deals with many of the key properties of attention, and also presents a comprehensive review of attention experiments in the process. Figure 5.2 is a model of the temporal sequence by which a new scene is analyzed, and behavior or recognition generated. The multiple arrows in this diagram indicate parallel processing of units, while the dark arrows indicate units allocated attention. A similar scheme would hold for any of the sensory modalities, such as hearing, taste, smell, or touch. The squares in the diagram represent a sequence of steps leading to perception. (Note that the recognition experience is an output of this selection of interpretations stage in Kahneman's model.) Attention and past experience are represented through the circles off to the side. In his model, attention affects two stages, figural emphasis and response selection. Internally generated experience enters the selection of interpretation stage in the form of perceptual readiness and the response selection stage through response readiness.

Kahneman separates the following stages:

A- The image is registered in iconic memory.

B- Unit formation is the fragmentation of a scene into homogeneous regions, or 'good groups'. It is a pre-attentive, or unfocused, process which relies on the physical properties of a stimulus. For
FIGURE 5.2: KAHNEMAN'S MODEL OF PERCEPTION AND ATTENTION

See discussion subsection 5.2.1.
(From Kahneman 1973)
example, Gestalt psychologists found visual qualities that lead to the emergence of objects. Three of the most important are: common rate (whole thing lies on same trajectory); good continuation (a closed outer contour can be formed by adding a smooth or curved line); and proximity (Bower 1974). Other contributing factors are warm color (red and yellow versus blue and green), complexity, texture and isolated contours.

The rules governing the emergence of objects in scenes are quite different for the portion striking the foveal part of the retina, as opposed to the peripheral part. The central 5 degrees of the foveal area has by far the highest visual acuity in the retina. The exertion of mental effort to focus attention causes the the foveal area to be processed more fully at the expense of the periphery (Cornsweet 1970).

It has been argued, that in a model of attention, cases in which focused attention is used must be separated from those that do not. Although Kahneman presents examples demonstrating this point, and acknowledges there is a problem, he tries to incorporate all the cases in a single sequence of stages. Table 5.1 organizes examples of attention in a way that indicates very different types of processing may go on in different situations. Kahneman's model runs into problems, because it tries to account for perception during a search with the same stages as are used to account for free examination of a scene. Kahneman's formulation of unit formation may be applicable to exploration, but ignores expectations and prior experience. His figural
emphasis stage is relevant in the process of searching and in the latter part of exploring a complex scene. Much of the incentive for introducing the model given in chapter 5.4 comes from the need to handle different cases.

C- Figural emphasis, deals with the formation of figure and ground. Some of the characteristics of a figure are that it tends to appear closer, and more impressive than the background and that it has closed, bounding contours that appear distinct from the background. In some cases, the Gestalt rules governing attention are all that are used in figure selection. In this case, figural emphasis is not really a separate stage from unit formation. However, in many cases collative factors such as novelty, incongruity and significance (due to expectation or selective intention based on prior instructions) can affect figure-ground emergence. In addition, unit formation can often provides several different 'good groupings' within a single scene, for example, viewing either the letters or the words on a printed page. Usually the subject's intentions resulting from prior intentions determine which grouping will become the figural part of the scene. In some cases, it is impossible to view a desired part of a scene as a figure. For example, when a mixed array of digits and letters are briefly presented, it is not possible to read only the letters (Kahneman 1973 p. 79). This type of array causes unfocused attention to be used in the scene analysis and focused attention is needed to separate digits from letters.

D- Recognition units are a type of internal information store that can be activated by special combinations of qualities, which include size,
shape, color and movement. The use of recognition units is not really a separate stage, but something that could be part of figural emphasis. Kahneman was lead to introduce the idea of recognition units in order to explain why a variety of figures with very different qualities are grouped together, on the basis of a few similar qualities. Posner (1969) found evidence that there are recognition units for identical letters (a-a), letters with the same name (a-A) and letters versus numerals, but not for consonants versus vowels. Posner's results are supported by both reaction time and search time experiments. The evolution of these recognition units is a very slow process. Prolonged practice seems to lead to a recognition unit for almost any arbitrary combination of qualities.

E- The selection of interpretation stage is required to settle the ambiguities that might arise among the recognition units activated. An interpretation which is acceptable to the subject is influenced by perceptual readiness, or by what is already known about the scene.

F- Response selection out of the set of available responses.

Kahnman's notion of a recognition unit seems closely related to the concept of templates, introduced in the discussion of LTM, chapter 3.5. There, it was proposed that templates are flexible, elementary units which can be integrated into synthesized scenes or events. This implies that the extraction of information is only a part of the recognition process, and that a complimentary synthesis must take place before a complex scene is recognized. Further, it implies that the role of attention is to change the probability of any template being activated. This view of attention is closely related to the analysis by synthesis
model introduced by Neisser (1967), which proposes that perception occurs through a constructive process. Attention selects the input components that go into the construction. A limit on what can be attended is achieved automatically, since unattended stimuli are not subjected to analysis by synthesis.

It is suggested that there are three types of templates that should be distinguished in a model of attention: those connected with sensory experience, emotions, and programs of action (such as a plan to search for a particular sensory object, or a plan for getting from one location to another). This suggestion is based largely on the way attention to sensory experience, emotions, and programs interact, discussed in subsection 5.3.3. It seems that either sensory input or internally-generated signals can activate the sensory templates, while emotions and programs of action are internally activated. The author's selection of input sources to which attention can be allocated was also influenced by concepts found in a number of eastern philosophies. The internal landscape described by various mystics consists of four planes among which attention or awareness is divided. Twitchell (1969) describes placing awareness on the astral (emotional), causal (memory), and mental planes which correspond precisely to the input sources given above. Evidence indicates that there are interactions between the types of templates, but that attention is usually centered on a single type. Once a template is activated, it can be integrated into a synthesized scene or event, which, when fully activated can include templates of all types.
5.3 A NEW MODEL OF ATTENTION COMBINING INTERACTION AND SYNTHESIS

5.3.1 Basic components of the model

The model suggested here goes beyond previous models in claiming that extraction and synthesis form an iterative sequence leading to recognition. The synthesized scene guides the extraction of information, and the extracted information then modifies the synthesized scene.

Besides controlling the extraction of information, the synthesized scene may activate new templates. There are many examples of iterative interaction of extraction and synthesis where most of the information is internally generated. Kahneman describes numerous experiments in which the subject is instructed to perform a complex task, such as to count backwards by 7's, and then is left without further sensory input to complete the task on his own. Although no further sensory input is required for the ongoing performance of such a task, Kahneman claims that these activities can demand large allocations of attention and mental effort.

Mental effort and allocation of attention are not equivalent, but are often associated. The question arises as to why this association exists. It is proposed that constructing a synthesized event or scene is the act that requires mental effort. Focusing attention, or
increasing the probability of selecting a certain template, is accomplished through the use of a program. A program can focus attention directly on stored templates, leading to the effort demanding task of building a synthesized event for holding items in short term memory. Since any motor behavior requires a program, this proposal on the nature of mental effort predicts that any action will require a fair amount of mental effort. Indeed, making any response, even a free one, requires a large amount of mental effort and a large allocation of attention (Kahneman 1973).

A schematic diagram of the proposed model of perception and attention is given in Figure 5.3. The following sections explore the characteristics of the objects in this model and the model operation in particular examples.

5.3.2- Templates and attention channels

It is suggested that programs select templates by sensitizing for particular qualities. For example, the visual quality of color could be considered a band with red and green being points on the band. The activation of a template is achieved by sufficiently strong signal in a specific portion of a number of different quality bands. For each of the three types of templates there is an attention channel, that contains a set of quality bands. The attention model included three attention channels corresponding to the three types of templates. The sensory, emotional, or program attention channels act as limited
Each attention channel is given momentary emphasis as well as a set of allowable templates by the program of action. (Multiple arrows indicate parallel processing.)
capacity filters for the quality bands over which the respective information comes.

An attention channel setting specifies the total sensitization allocated to each quality band, relative to the others. The allocation is centered on a particular portion of each quality band. For the sensory source the most basic bands are modalities, with different types of receptors and different places of entry into the cerebral cortex. The sensory attention channel may select for the visual signal over the auditory signal, at any given moment. At a hierarchical level below modalities, there are bands within a modality that have a set of visual cortex neurons that respond to different portions of the band. These bands correspond to the combinations of qualities that could be habituated by long exposures, such as the orientation and color band demonstrated by the McCollough effect. Other qualities with this property are movement, disparity, spatial frequency, and left or right eye. The simplest kind of templates can be activated by a portion of a single quality band with visual cortex response. For often observed objects rather complex combinations of bands, and ranges within each band, can be emphasized by the attention channel. The setting of the sensory attention channel to a complex combination of quality band portions corresponds to the slow construction of templates, or the recognition units mentioned by Kahneman. The following two experiments may be examples of situations in which new templates or allowable configurations within the attention channel are arising. Pribram (1976) observed that after a number of training sessions the initial EEG
response of human subjects to geometric objects moved from the IT to area 17. Metzler and Spinelli (1977) showed that the percent of area 17 in oats responding to bars continuously parallel to gravity increased from 5% to 15% during a long exposure to such bars. Thus, there appears to be a close correspondence between combinations of visual qualities that promote figure emergence and those to which there are a responding subset in area 17.

The attention channel at any given moment can be set by the currently activated program to sensitize for nothing, a portion of a quality band, or a specific object template. There is a continuous spectrum of verbal instructions that influences the nature of an attention channel ranging from the concrete to the abstract. For example, one may locate a particular chair in a scene, or one may look for an object that can be moved to a table and sat on. In the first case, one specific template in memory corresponding to a past visual experience could be triggered. In the second, a specific template is not called for, but rather a flexible set of quality templates for a particular height, strength, and weight. An example of setting the attention channel for a simple quality, rather than a specific item, occurs with the verbal instruction: "What is the red word?". The program generated can only select for the quality which can then be used to focus on the location of the figure.

In exploration, the visual input itself must activate a set of templates, some of which must be integrated into a synthesized scene
that will be used to control the nature of the attention channel in the
next moment. Note that if the original fragmentation of the scene can
be done quickly enough, the attention channel can be directed at the
fading iconic memory and a particular template can be selected.
Usually, more information is desired and an eye movement places the area
of interest in the center of the fovea. The attention channel selects a
small portion of the visual field, and uses the fragmentation of this
input to activate a particular template.

Most of the time, organisms deal with dynamic events rather than
static scenes. The moment to moment changes in a dynamic scene would
make a sequence of fragmentations incompatible. Fortunately, movement
or changes tend to be the most powerful attention drawing qualities, and
overpower even strong momentary expectations in determining which
quality bands draw sufficient attention to activate templates (Bower
1974). The process of dividing a movement into basic units is here
defined as the segmentation of an event into routes. In event
segmentation, one, or at most, a few, objects can be followed.
Attention stays focused on those quality bands of the object that are
changing. When a subject's attention is not tightly focused on the
changing band, he is only aware that the system has gone from state A to
some other state B. The next time the subject views the system in state
A, he can focus his attention on the band where the change occurred
previously and pick out some intermediate states. To aid in obtaining
all the information in an event when the changing band is movement, the
dominant type of eye movement becomes smooth tracking, rather than
saccads (Robinson, 1975).
5.3.3 Synthesized scenes (events) and the interaction of attention channels.

A sensory driven synthesis can lead to either a synthesized event composed of route templates or a synthesized scene composed of object templates. The building of a synthesized event or a synthesized scene must be directed by a synthesized event for a program. The rules that these programs follow are mostly learned in the first year of life, but can be modified for special situations thereafter (Bower 1974).

The other type of synthesis that a human can perform is a program driven synthesis. A good example of constructing a program is the preparation for making a new movement. The subject making the movement starts with a vague image of moving to the desired final position. The process by which the synthesized event is constructed involves a sequence of attention focusing directed by a program. A vague synthesized event for going between the initial and final positions is activated by templates for these positions. It is vague in the sense that there are not specific, well-known routes leading between the initial and final states. The best defined intermediate route template in this synthesized event then draws focused attention, which activates a number of stored movement templates going from the initial state to the intermediate and from the intermediate to the final state. Attention to the memory source then selects one set of these and this
strengthened template is fitted into the synthesized event, if possible. If fitting is impossible, a different one may be tried. Once a template is fitted to the synthesized event, it remains available even though attention is taken away from it. Progressively, routes are incorporated into the synthesized event for the movement program until it can be performed. Again, a program guides this synthesis by a set of rules established early in life.

An example of a program driven synthesis occurs during mental arithmetic. There is an initial sensory input of the two numbers and the operation to be carried out on them (37 times 3). A program to store both of the numbers in STM is initiated. Once the STM program is well established, attention is shifted to a program for building a synthesized event to carry out the multiplication. This synthesized event must remain activated throughout the calculations and control the attention channels for all four sources. Its creation involves the activation and linking of a number of program route templates. These route templates can be stored in something akin to the short term memory for sensory input. Note that each route template can act as the center of a complete synthesized event when it receives attention. When one route is completed, its ending must return attention to the program or to the next route in the program. The program contains a sequence of route templates for focusing attention on pairs of numbers held in STM and a route template for replacing them with a new number, one for addition and one for multiplication. The proper replacements must be a template activated by the pair of numbers in STM that are being attended. The synthesized event for the program also specifies a
particular positioning for the new numbers and an order for carrying out the replacements. In addition to the replacement programs, considerable attention must be allocated to a program for maintaining these intermediates in STM. 3x7 is replaced by 21 and stored, 3x30 is replaced by 90 and stored, then 21+90 is replaced by 111. Before each replacement, the focus of attention for input coming from LTM is made very narrow. The sensory and emotional inputs are largely suppressed.

The synthesized event for multiplication calls for a high level of arousal. It requires large allocations of attention throughout, due especially to the large demands on STM. In fact, for longer multiplications, chunking of the intermediate numbers is required to hold all the information. Chunking is also an attention focusing process that requires mental effort. Therefore, any other program initiated before the multiplication is complete is likely to disrupt the process.

During the synthesis of templates, the quality bands sensitized within the same attention channel interact, as well as the band simultaneously activated in different attention channels. Stated differently, attention is allocated to the three types of channels through the amount of sensitization given to the quality bands in each type. Sometimes the input coming through the different attention channels activates compatible templates which fit together in a way analogous to the way different parts of a mosaic supplement each other to form a coherent picture. Sometimes their interaction lacks
compatibility and confusion results. An example of a compatible interaction might be drawn from seeing a pet cat. The cat would visually trigger a template. This could activate a program to increase arousal, and construct a synthesized event or scene. The activated visual templates could be used to internally generate activity in the emotional or program quality bands, or new activity in the sensory bands. Templates could be activated leading to synthesized events such as memories of past experiences with one's pet, or with other cats, and a program for petting the cat may be activated. Feelings of affection from the emotional source, memories of past experiences with one's pet or other cats, and a program for petting the cat maybe activated. There are numerous physiological and psychological studies of the interaction between modalities and other qualities, especially for inputs from the sensory source. For instance, a cat observing a mouse may not respond to the same click that normally draws a response. Hernandez-Feen (1955) showed that in this situation the cat's EEG showed no response to the click, even in a nucleus as close to the ear in the auditory pathway as the cochlear nucleus.

Some of the rules for selecting quality bands that can be fitted into the sensory attention channel, i.e. for which programs exist, are illustrated by experiments in which two signals are presented simultaneously. Suppose that the digits 3, 5, 7 are presented to the right ear and the digits 2, 4, 6 are presented to the left ear so that the pairs 2-3, 4-5, 6-7 arrive together. Then the reported groups would be 3, 5, 7 and 2, 4, 6 rather than the simultaneous pairs or a random mix.
However, lists of two very different types of items randomly mixed tend to be grouped by types, even though it means mixing the inputs from different ears. So, the normally dominant quality can be overcome by a channel based on semantic content. When an entire synthesized event must be constructed to process an input, a signal not fitting the channel is strongly suppressed. For example, different stories presented to each ear cannot both be followed.

In searching for targets within messages to both ears, one, but not both are generally reported, when two targets appear at once (Kahneman 1973). A program is needed to sensitize the quality band for a particular ear in the sensory attention channel. A program that requires checking both ears must sequentially check them one at a time. If a template is activated by the input from the first ear, the iconic memory for the second target may be faded before the first is processed. The interference between quality bands in the sensory attention channel seems to be less severe when the units are in different modalities. Perhaps the strongest type of interference is between different parts of the same quality band. For example, it is almost impossible to focus attention on both the upper left and lower right hand corners of the visual field at the same time.

Another way that quality bands interact is through the length of time that signals coming through them are held. It appears that the unattended band decays much more rapidly in storage than the attended band (Krech, et al 1974). This agrees with data showing that the size
D) The program templates seems to be in the lateral frontal cortex. Lesions of the lateral frontal cortex impair the ability to guide behavior with verbal instructions or use programs in general (Luria 1973).

The evidence indicates that processing of each type of template receiving attention can be impaired without greatly affecting processing of the others.

5.4.1- The frontal effects on the dimensions of attention

Within a single attentional shift it was noted above that intensity, and dwell time of attention vary. There seem to be two main regions of the cortex involved in controlling these dimensions of attention, the frontal and posterior association cortex. A number of experiments have shown that a consistent set of effects result from frontal cortex interference. The effects of lateral frontal stimulation on visual attention are to extend the area responding to a particular visual input in the LGN (Pribram 1977). This could probably lead to stronger inhibition between the response to objects in the scene resulting in a tighter focus of attention on the foveal region of the position quality band. In a second experiment, Pribram showed that the reaction time to the second of a pair of stimuli was speeded during frontal stimulation. These effects suggest that frontal stimulation can modify attention by sensitizing a narrow portion of at least one quality band in the attention channel and by decreasing the dwell time on the sensory category. They also seem similar to those found during search as discussed in chapter 5.3. One possible function of the frontal cortex
may be to activate object or quality templates, so that the visual attention channel is sensitized for particular types of inputs. These inputs become figures of high intensity when they are located. In general, high arousal during search produces faster reactions, but less accuracy or a poorer than normal ability to extract details. Noise has been found to lead to physiological signs of arousal (Kahneman 1973), which in turn has been correlated with high frontal activity (Luria 1973). Noise has been found to decrease the threshold of a stimulus while, increasing the threshold of a dim stimulus shown simultaneously (Kahneman 1973). Similarly, the arousal caused by a first stimulus acting as a cue for a second stimulus, speeds the reaction to the second stimulus (Kahneman 1973). These data could result from visual attention channel narrowing in one band, leading to greater interference and greater intensity of attention on the target.

Frontal activity may also indicate another type of attention allocation. EEG data shows that a natural way of getting heightened frontal activity is to perform mental arithmetic (Livanov 1977). During this operation it can easily be seen that the person is less distractible to novel stimuli, i.e. there is lowered visual attention intensity, and that the intersaccadic interval is short (Kahneman 1973). Mental arithmetic and general problem solving is accompanied by a high saccadic rate even in the blind (Amadeo & Gomez 1966). This indicates a narrow width of attention and a short dwell time as in search, but a low intensity of visual attention instead of high. The effect of mental arithmetic on attention to the program source is to
increase the intensity of the program, its specificity and scope, as well as to increase the dwell time on the program. The sense of the present may extend until the problem is completed.

Data from human patients with lesions in the frontal cortex showed fundamentally the opposite effects of frontal stimulation. Such patients are easily distracted by novel stimuli, but stare into space without stimuli (Luria 1973). They respond to questions directed at others, but not at themselves. They seem to have unfocused sensory attention and a long dwell time. The worst behavioral defect these patients show is the inability to use a verbal program to sensitize appropriate attention channels. Their behavior can no longer be guided with verbal instructions, especially in patients with left side lesions. Patients manifest such lack of control when viewing thematic scenes by building a story around the first object in the scene that catches their attention (Luria 1973). Patients will say that the theme of the picture is related, in a nonsensical way, to something about that first object that caught their attention (Luria 1973).

The mistakes these people make are not noticed possibly because of an inability to bring information from all the sources into a synthesized event or scene. The problems that frontal lesion patients have with thematic pictures shows that their programs are not only ineffective in guiding behavior, but that their ability to switch between programs, based on environmental feedback, is also defective. This behavioral problem is exemplified by the persistence of behaviors, once initiated,
seen in frontal lesion patients. For example, a patient told to draw a circle will continue to draw many, overlapping circles (Luria 1973). The author proposes that the two main problems that these patients have is the inability to incorporate templates into synthesized events, and the inability to elevate arousal to meet the demands of the current task. The proposed model requires an increase in arousal before a synthesized event can be formed, and without synthesized events extended programs of action cannot be carried out.

5.4.2 Posterior effects on the dimensions of attention

Posterior association cortex stimulation seems to lead to lower attention intensities, to an expansion of the width of the sensory attention beam, and to longer dwell times.

In the same experiment in which he stimulated the frontal cortex Pribram also stimulated the posterior association cortex. Longer reaction times to a second stimulus and a contraction of the reacting portion of the LGN to a stimulus were observed. During attentive listening or free examination of a picture the posterior association cortex seems very active, the intersaccade times are long, and events in the periphery are easily noted. During these activities there is a comparatively low level of arousal as indicated by pupil dilation (Kahnemen 1973). Most of the attentive capacity that is available seems centered on the sensory source with a little on some of the triggered templates in other sources. During free examination of a scene, the present seems so short that it shifts with each eye movement.
Little sensitization of the quality bands in the sensory attention channel seems to take place under conditions of posterior association cortex stimulation. Thus, posterior stimulation would correspond to periods of information gathering in which connections between templates could be strengthened or new templates formed. The weak attention to the program source may be a prerequisite to learning.

Observations on patients with posterior association cortex lesions also support the inferences made on the basis of the stimulation data. Some of Luria's patients had a condition known as simultaneous agnosia, which means that they could only see one object at a time, no matter what the size. In addition, they seemed to pick out fewer of the qualities of the objects they did see than normals found. Posterior lesion patients are also indistractible by novel stimuli during behavioral sequences (Pribram & McGuiness 1976). In patients with parietal lobe lesions, objects can be picked out, but the relations between them do not reach awareness. These patients are unable to follow directions or to read maps. This problem with relations leads to breakdowns in the recognition of objects. For example, letters cannot be recognized in left side lesions, while faces are often not recognized with right side lesions. This problem seems connected to the inability to build sensory input into a unitary structure with the underlying problem deficit being in the ability to construct composite sensory channels.
The author believes that the posterior lesion data can be related to an inability to construct sensory driven synthesized scenes. The reason for this seems to be related to the behavioral problems of too much sensitization of the visual attention channel. This could prevent the activation of a number of templates which would have to be organized into a compatible synthesized scene. In chapter 4.2, some data was presented supporting the idea that the IT plays a major role in the reassembling of visual input into a non-egocentric scene. This role for the IT could be achieved if it acts as a station for combining visually activated templates. The role of the hippocampus in learning in these terms could be to generate the right behaviors and arousal conditions for synthesized scenes to form.

5.5- A NEURAL MODEL OF ATTENTION

In this chapter, it is shown that there is good agreement between the neural model of the relay system and the model of attention developed above. In chapter 3.5 it was argued that object and route templates could be associated with active circuit cores, while synthesized events and scenes could be associated with turned-on active circuits. In chapter 4.2, a method of incorporating a turned on active circuit core into a turned on active circuit was proposed. These concepts will now be related to attention channels and their control by a synthesized event for a program.
The idea that there are weakly connected regions of the CNS that can act as sources of input for different types of information can be applied directly to the relay system model. This implies is that active circuit cores would rarely bridge these regions of the brain, but turned on active circuits could spread from one region to another when sensitization of relay cells is high. The importance of the regions that are incorporated in the active circuit, in terms of the recognition and the behavior that result from turn on, has already been discussed in chapter 4.2.

The relation between the relay model and attention channels can be illustrated by reexamining the case of a person searching for a red word on a page. A verbal instruction turns on a set of active circuit cores. Arousal is raised by the sound and the active circuit cores are compatible so they are combined into a turned on active circuit centered in the temporal cortex. This turned on active circuit then expands to the frontal area where it turns on an extension of that active circuit that can be associated with the synthesized event for a program. Part of this active circuit will be on visual cortex neurons which respond to red inputs. In addition, it will be sent to motor areas that will center the eyes on the middle of the page. When the visual signal from the eyes arrives and activates some active circuit cores, the driving region of the active circuit will move to the visual cortex. If there is some red on the page, it will tend to activate a strong active circuit core. As this core is incorporated into the active circuit, it will trigger a link to the oculomotor nuclei that can center the fovea on the red region. In the visual cortex, the active circuit will now be
centered on columns in the foveal region of visual cortex, and on columns associated with words in the secondary visual areas. Thus, the foveal part of the visual field will turn on active circuit cores which must be integrated in the posterior association cortex. The strong activity centered in the columns associated with the word will lead to the behavior of pronouncing it.

The limit on the capacity of attention in this model would then be set by the combinations of columns in the visual cortex which could be incorporated in the same active circuit. A set of columns in this area that was not innately compatible, might be made so by the strengthening of any the paths between them (possibly through distant nuclei) and weakening any inhibition. This might be accomplished in the manner described in chapter 4.2. The interference between quality bands is not hard to explain in terms of the neural model. The strong innate inhibition between nearby columns in the visual system would tend to prevent different parts of the same quality band from belonging to the same active circuit. In addition, it would prevent more than a few active circuit cores from sharing the same nuclei. Arguments for how the dwell time and intensity of attention could be controlled by turned on active circuits with columns in the midbrain could also be given. They would, however, be speculative, due to the paucity of information currently available.
6- DISCUSSION AND CONCLUSIONS

Modeling the nervous system is an awesome task. Experimentalists are still in the preliminary stages of collecting relevant data about mechanisms of communication and structural change. Researchers do not agree on which nuclei are connected to which, let alone on which cell types actually receive the connections. There are still disputes on the source of EEG, one of the main measures of physiological activity in the nervous system. Nonetheless, a great deal is known about the nervous system. By posing specific questions, and selecting a reasonable set of assumptions based on empirical observations, models of particular aspects of neural functioning have been constructed. Most of the neural models found in the literature do not attempt to account for actual neural arrangements. Rather, they tend to show that a completely general neural circuit has a processing property reminiscent of a psychological process. The problem with this approach is that many systems have properties in common, and show the same type of activity; for example, both the interneuron network and the relay system show hysteresis effects. Therefore, it is important to determine which of the existing models really are applicable to modeling particular neural functions.

6.1- ASSUMPTIONS IN THE FOUR CELL TYPE MODEL

In this paper, an attempt was made to model a specific, and very complicated function of the human nervous system—visual perception.
The bulk of the work involved selecting those neural structures important enough to be included in the model. Five postulates were proposed as guides in the selection process, and a four cell-type model with variable parameters resulted. Possible objections to the physiological assumptions used in constructing the four cell type model need to be critically reviewed. Three points particularly worth questioning are: the number of cell types, the connections among the cell types, and the consequences of variation within the chosen cell type populations.

Each of the four cell types used in the model was included for a specific reason. Relay cells provide the communications between nuclei, small excitatory interneurons receive the primary visual input in area 17 and have very different axon field shapes than relay cells, small inhibitory interneurons prevent runaway positive feedback, and large inhibitory cells, unlike other s.i.i., have axons which extend beyond the columns defined by the axons of cortical pyramidal cells. Further modifications, such as subdividing the cell types based on anatomical differences, were not added because the performance of a model tends to deteriorate with excessive complexity. It was subsequently found, however, that modeling the extraction of visual qualities did necessitate a further subdivision of the interneurons.

Next, the connections assumed in this model are considered. The assumptions used to build the model were all based, in least in part, on physiological evidence, but some of the assumptions are still
speculative. Certainly, the assumption that each interneuron is in contact with all contiguous interneurons is too simplistic. Yet as a first approximation, this assumption was useful in permitting a functional model of the nervous system to be constructed and examined.

The one type of intercellular connection conspicuously absent in the model, that of relay cells and large inhibitory cells to other interneurons in the same column, have not conclusively been demonstrated to exist. But, the lack of intracolumnar connections between relay cells is not strictly true, since recent findings by Szentagothai (1978) show that axon collaterals contact the basal dendrites of nearby pyramidal cells. If, in fact, there are strong connections from the pyramidal cells back onto interneurons within the same column, then the decomposition of equation 2.1 into the interneuron network and the relay system would no longer be valid. But if only a few, relatively weak intracolumnar connections exist, then the model would not be invalidated.

Another assumption that may not be strictly true concerns the in-register termination of the cortical afferents. The effect of showing this assumption to be false would again depend on the extent to which the assumption was wrong. Due to the 800-fold overlap of columns, there may be no major difference between the system's performance with in-register columns and without them. This question could be addressed with a specific quantitative model of the relay system. A final controversial assumption about connections made in the model is that the lateral spread of dendrites is localized. The assumption was made to
emphasize the between-nuclei connections, and to explain the fine-grained differences in neuron response. Although it is in agreement with some physiological data (see Figures 2.4 and 2.5), it is apparently contradicted by others. Schiebel (1979) found that for the giant Betz cells of the motor cortex, and several other types of cells, a dense plexus of basal dendrites extended a millimeter or more from the cell body. Such an arrangement might allow neurons to monitor the general level of activity around them, but there is no clear role for such an arrangement in the present model.

In addition, it is known that there is a great variation in the relative proportion of cell types in one nucleus versus another, especially in subcortical areas; this variation was not incorporated into the present model. Finally, the shape of the axon field of each type of cell varies. These differences may lead to different basic processing, especially in nuclei as different from the cerebral cortex as the reticular formation. Future models, focusing on particular processes within the nervous system, will need to take this variation into account.

6.2- MODELS OF RECOGNITION, ATTENTION, AND MEMORY BASED ON THE ACTIVE CIRCUIT CONCEPT

Recognition was broken down into component processes; these processes were then associated with particular aspects of the general neural model characterized by equation 2.1. In the discussion of attention (section
5.1) it was pointed out that the attended experience leading to recognition is different during exploration than during searching. In both cases, recognition involves the initial breakdown of the scene and its storage in iconic memory. This leads to a separation of figure and ground plus the activation of associated templates, followed by the synthesis of many of the activated templates into a whole concept. The interneuron network in area 17 was proposed to be primarily responsible for the extraction of visual qualities, due to the different asymmetries within the s.i.i. population. However, qualities such as brightness, movement, and texture might also be extracted in the SC or pretectum. The extracted information could be briefly stored in the form of self-sustaining interneuron activity, called a localized steady state (subsection 2.2.3). It was proposed that the interneuron network requires a prior sensitization associated with mental effort in order to enter this activity mode.

It was postulated that the interneuron network performs only local processing, and that the synthesis and associative memory used in recognition arise from the between-nuclei interactions in the relay system. The main mode of activity in the relay system was proposed to be the active circuit (section 2.3), a feedback loop among columns with multiple convergence, that can maintain self-sustaining activity among nuclei. In chapter 3, a relationship was suggested between the two types of objects in long term memory, templates and synthesized scenes, and the two types of active circuits. The two types of active circuits postulated are: the complete active circuit, requiring only relay cell
sensitization in order to be turned-on, and the active circuit core, which can be turned-on by sensory input without prior sensitization. (The active circuit core is a small portion of a complete active circuit whose participating columns have greatly increased connection strengths.) It was suggested that the elementary units of recognition, templates, might be stored in active circuit cores, while the composite templates, called synthesized scenes, or events, could be stored in complete active circuits. The extraction of figures from a scene may also occur through the activation of active circuit cores, which would lead to the preliminary recognition of individual objects in the scene. Note that lateral inhibition would cause active circuit cores to interfere; thus, the model predicts that only a limited number of figures could be discerned at a time. This would require the sensitization of the relay cells, and would be represented in the model by a change in parameters. It was suggested that the IT is the key region in organizing the integration of active circuits. The consequences of turning-on an active circuit go beyond the simple storage of a pattern of firing in a set of columns. Each column in the active circuit plays two distinct roles, maintaining the turned-on active circuit, and contributing information to overt behavior. Each active circuit contains columns in many nuclei; it is the activity in specific nuclei that leads to the feeling of recognition, and initiates overt behavior.

Active circuits were also credited with a major role in the control of attention. An active circuit can raise the activity in selected
columns in a sensory region, and thus increase the probability that a particular set of active circuit cores will be turned-on by sensory input. In addition, midbrain columns participating in the active circuit may sensitize a subset of the cortical s.i.i. population related to the extraction of a particular quality, thus increasing the response to that quality. The active circuit concept results in a plausible explanation of how incoming visual information is brought to awareness. If one active circuit core were to build up more quickly than the others activated by a scene being analyzed, it would tend to dominate that set of columns turned-on in the IT, and trigger the turn-on of associated active circuit cores, or memory templates. The initial recognition of one figure would trigger other memories, resulting in a synthesized scene that allowed it to be interpreted in this context.

In addition to the similarities they share with long term memory, active circuits have the property of synchronization, which allows them to be used to define neural waves. Establishment of a neural wave equation allowed a strict comparison of the brain's theorized information storage and retrieval system with holography. An I/O function observation morphism was demonstrated between the holographic system and the neural scene analysis system. However, the real interest in the holographic analogy lies in comparing the transfer functions taking the spatial distributions associated with neural and optical input signals into the respective spatial distributions of the recordings in which they are stored. It was found that the transfer function for recording in each system could be divided into two
processes, a transformation going between locations, and the interaction with a recording medium. The transformation process for neural waves traveling between nuclei was compared with the well-known transformation of optical waves. However, the available data were insufficient to decide whether or not consistent transformations are applicable to pairs of nuclei. Certainly, different transforms would be needed for different pairs of nuclei. On the other hand, the interactions with the recording media were shown to be clearly different in the two systems, although, both recording processes involved making the greatest changes at the points of highest activity. In the neural case, there seems to be multiple steps of progressive change, rather than a single one. The change is based on subtle attention drawing characteristics in the scene, rather than simply the total energy coming from each point. The final neural recording also seems to distributed based on spatial separation of many qualities, rather than on just a spatial frequency transformation. The problem of expansion through multiple steps also prevents a mathematical comparison of the two reconstruction subsystems.

6.3- EXPERIMENTAL TESTING

There are a number of experiments that could be used to test the credibility of the models presented. The most basic types are experiments to find the connections among the cell types which are still in question. To test for self-sustaining activity in the interneuron network, one may look for hysteresis and cycling effects in response to a constant LGN response. In fact, Singer (1979) showed that this does
occur. The other aspect of the interneuron network model that should be tested is the extraction of visual qualities. Note that the focusing of attention on a particular quality may sensitize a particular subset of the s.i.i. population. Therefore, some set of neurons may change their sensitivity when red is expected, compared to when it is not, while the sensitivity of other neurons would remain unchanged.

More importantly, the existence of active circuits should be tested. The existence of active circuits would be supported by demonstrating hysteresis in relay cells, without a similar effect in the interneuron network. If a number of area 17 neurons could be recorded and stimulated at the same time, it might be possible to see hysteresis in neurons responding to the same quality.

To test the idea that an active circuit is a memory store, one would wish to show that inhibiting an active circuit could inhibit a specific memory. One way to do this would be to give a radioactive metabolic inhibitor, perhaps glucose with a heavy radioactive metal that is absorbed only by very active cells, while a single concept is kept under attention. This should inhibit the learning process, and disrupt the previously formed memory templates. The uptake of the metabolic inhibitor should be observed in a widely dispersed set of columns, and thus may pin-point the columns of the active circuit.

In conclusion, this paper suggests a number of lines for future modeling, as well as for experimentation. Several modeling questions about the interneuron system have been raised. The interneuron network
should be expanded to include s.i.i. asymmetries of various kinds, and tested for the ability to separate visual qualities. The time delay term in equation 2.1, which was dropped from the interneuron network model, could be reincorporated, to decide whether it really has an effect on the processing modes.

The vector model of the relay system should be tested to see if, as predicted, it could truly support turned-on active circuits with the same waveform throughout. A model of how the IT integrates active circuit cores into a consistent active circuit might also be attempted. There is a seeming contradiction between the ability to integrate different active circuit cores and the idea that synchronization would lead to the same waveform whether it was turned-on visually or internally. However, in both of these cases it was presumed that the active circuit core is turned-on by thalamic input. In the integration performed by the IT, the same columns might be be turned-on in a different way. It is hypothesized that the waveform in an active circuit contains information about the columns in the active circuit, not about the information which the columns contribute to behavior. Finally, models of the transformation of the distribution of active circuit columns going between two nuclei can be constructed. These models would have to be tested experimentally to see if the transform depended on a particular active circuit.

The four cell type model developed in this paper opens new avenues of study in both modeling and research. It forms a general framework for understanding cognitive processes based on single neuron properties. In
addition, the model is a significant step towards using physiological data to iteratively design future research on the various levels of the neural hierarchy.
BIBLIOGRAPHY


2 Aidley, D.J.: The physiology of excitable cells. New York: Cambridge Univ. Press 1971


11 Bartlett, F.C.: Remembering. Cambridge: Cambridge Univ. Press 1932


25 Cowan, J.D.: The possible role of development in the specification of the affluent geometry of the primate visual cortex (To be published)


78 Lawrence, D.H.: Two studies of visual search for word targets with controlled rates of perception. Perception and Psychophysics


Psychol. 8:81-136 (1979)


96 Norman, D.A.: Memory and attention. New York: John Wiley and Sons 1969


123 Schmidt, R.F.: Fundamentals of neurophysiology. New York:


153 Wilson, H.: The significance of frequency gradients in binocular


