Quantum Neural Computing

Subhash C. Kak
Department of Electrical & Computer Engineering
Louisiana State University
Baton Rouge, LA 70803-5901
Email: kak@max.ee.lsu.edu

December 1, 1995

Table of Contents

Chapter 1

Introduction

This article reviews the limitations of the standard computing paradigm and sketches the concept of quantum neural computing. Implications of this idea for the understanding of biological information processing and design of new kinds of computing machines are described. Arguments are presented in support of the thesis that brains are to be viewed as quantum systems with their neural structures representing the classical measurement hardware. From a performance point of view, a quantum neural computer may be viewed as a collection of many conventional computers that are designed to solve different problems. A quantum neural computer is a single machine that reorganizes itself, in response to a stimulus, to perform a useful computation. Selectivity offered by such a reorganization appears to be at the basis of the gestalt style of biological information processing. Clearly, a quantum neural computer is more versatile than the conventional computing machine.

Paradigms of science and technology draw on each other. Thus Newton's conception of the universe was based on the clockworks of the day; thermodynamics followed the heat engines of the 19th century; and computers followed the development of telegraph and telephone. From another point of view, modern computers are based on classical physics. Since classical physics has been superseded by quantum mechanics in the microworld and animal behavior is being seen in terms of information processing by neural networks, one might ask the question if a new paradigm of computing based on quantum mechanics and neural networks can be constructed.

In recent years proposals have been made by Beniof (1982), Deutsch (1985, 1989), Feynman (1986) and others for the development of computers based on quantum mechanics. In these schemes the quantum mechanical basis states are the logic states of the computer and the computation is a unitary mapping of these states into each other. Hermitian Hamiltonians are specified that define the interactions. From another perspective, these computers let the computation of several problems proceed simultaneously as in the evolution of a superimposition of states. By itself that is no better than several computers running in parallel. However, if one were to imagine a problem where the partially evolved computations of some of these superimposed states are used to find the solution then it might be that such a machine offers improved speed in the sense of complexity theory. The idea of a quantum computer has not yet been shown to be practical (Landauer 1991, Gramss 1994). Furthermore, these proposals only deal with the question of the physics underlying basic computation; they do not consider the question of how a computation process leads to intelligence.

The proposal for a quantum neural computer is based not only on an explicit representation of the unity of the computation process, but also an attempt at defining computation as a behavior in relation to other systems. In other words, we wish to introduce computation in relative terms, a perspective that appears to be appropriate for computation related to Artificial Intelligence (AI) problems. This shift in perspective is like the shift urged by Mach in his criticism of the notion of absolute space that had been assumed by Newton. According to Mach, space should be eliminated as an active cause in a system of mechanics; a particle moves not relative to space but rather to the center of all the other masses in the universe. Likewise, one says that intelligent computation can only be defined in terms of the computational behavior related to other such systems.

1.1 The Mind-Body Problem

In order to be able to design machines that are equal to the capabilities of brains, it is essential to understand the nature of brain behavior. Our first hurdle is the mind-brain problem. The brain is defined by its physical and chemical properties, and it is assumed to function in ways that can be predicted by physics and chemistry. In contrast, the mind is an abstract entity that consists of sensations, beliefs, desires, and intentions. How is the physiochemical body related to a nonphysical mind?

In identity theories, the mind is a part of the physical properties of the brain, as is true of the electrical recordings from the brain. Mental states then should be seen as brain states and statements about cognition can be reduced into statements about behavior or a disposition to behave. But the identity theory does not explain how the activity in the brain assumes a unity which constitutes an awareness of the self. If stable brain states are identified as memories, as is done in many current models, can the problem of self be reduced to the problem of a bundling of these memories? Experiments with split-brain patients have shown that memory does not require consciousness. The converse, that consciousness does not require memory, has been known for a longer time. The split-brain subject appears to experience consciousness via the dominant cerebral hemisphere. By contrast, the minor cerebral hemisphere does not enable the subject to have conscious experience, although it subserves memory, nonverbal intelligence, and concepts of spatial relations. We know ourselves as unique indivisible minds, and not as a unique brain.

In dualist theory, the mind has a reality independent of the body. Dualists cite introspection by which knowledge of things such as pleasure, pain, happiness, emotions and so on is gained as proof of the existence of mind. The argument against dualism is that the recognition of the color green can be reduced to photons of a certain frequency striking the retina. Another objection is that a nonphysical entity cannot influence a material object without violating conservation laws.

1.2 Emergent Behavior

Mind, with its concomitant intelligence, is often taken to be an emergent property of the complexity and the organization of the interconnections between the neurons. The concept of emergent property comes from chemistry, where the properties of water are taken to be "emerging" from the properties of hydrogen and oxygen. Although the properties of water could not originally be predicted from the properties of hydrogen and oxygen, it is assumed that in future, given the properties of atoms and the rules of their combinations, the properties of water would be completely explainable from the properties of its constituents. The notion of emergent behavior is a consequence of a reductionist approach to phenomena. Such an approach is a reasonable foundation for scientific theories. But the workings of the mind appear to be so much different from the substratum of theories about the brain, that doubts have been expressed about it being an emergent phenomenon.

Neuroscience is also based on a reductionist agenda. It is believed that once the physiochemical basis of the activity and the interactions of the nerve cells are understood, mental events will be explained in terms of physiochemical events taking place in the nerve cells. But how awareness can emerge from principles of physics or chemistry has not been understood. Neither does current theory explain how awareness could arise from complexity alone.

The idea of emergent behavior does not rule out the need to add new laws at the higher level. Does the mind have unique properties that are not reducible to physical laws? In going to higher levels related to brain behavior, we confront the tyranny of complexity: the explosion of combinatorial possibilities may make it impossible to answer whether the higher level behavior is reducible to that of lower levels.

1.3 Indivisibility and Quantum Models

Reductionist approaches to brain function do not capture the richness of biological information processing. According to the complementarity interpretation of quantum theory any indivisible phenomenon must be described by a wavefunction. This is why "elementary" particles, which in turn have "sub-particles" as constituents, are described by wavefunctions. One can even speak of a wavefunction for a macro-object and indeed for the entire universe. If consciousness is taken to be indivisible, one has no choice but to model it in a quantum mechanical fashion.

Evidence for the unity of self-awareness or consciousness is provided by many neuropsychological experiments. These include split-brain research as well experiments on dissociation of behavior from its awareness as in prosopagnosia, amnesia, and blindsight (Weiskrantz 1986). Schrödinger (1965), Penrose (1989) and other scientists have argued that, as a unity, consciousness should be described by a quantum mechanical type wavefunction. No representation in terms of networking of classical objects, such as threshold neurons, can model a wavefunction. Therefore, current computing machines, which are based on mechanistic laws of information processing, are unlikely to lead to machines that would match human intelligence.

It is sometimes argued that since quantum mechanics is not needed to describe neural processes, therefore it should not enter into any higher level descriptions of cognitive processes or of consciousness. The noisy environment characterizing nerve impulse flow should drown out any quantum mechanical behavior. On the other hand, it has been experimentally determined (See Schnapf and Baylor (1987)) that the retina does respond to single photons although the process that leads to such response is not understood. This was determined by directing dim flashes of light into one eye of a subject sitting in total darkness. The subject perceived a flash when only seven photons were absorbed. Since a population of about 500 rods in the eye absorbed the photons in a random spatial pattern, there was no likelihood that any single rod had absorbed more than one photon. The response itself is a macroscopic nerve signal.

Now consider social computing (Kak 1988), that is processing that takes place in a society of individuals without a conscious notion of a collective self. Current machines are not capable of matching such performance. Social computing appears to be less powerful than computing that is accompanied with awareness, therefore current computing models seem to fall considerably short of the capabilities of biological systems.

There also exist speculations regarding consciousness in quantum physics. It has been argued that the collapse of the wavefunction occurs owing to its interaction with consciousness. On the other hand, physicists like Wigner have argued that science, as it stands now, is incomplete since it does not include consciousness. In a variant of the Schrödinger cat experiment, Wigner (1967) visualized two conscious agents, one inside the box and another outside. If the inside agent makes an observation that leads to the collapse of the wavefunction, then how is the linear superposition of the states for the outside observer to be viewed? Wigner argued that in such a situation, with a conscious observer as a part of the system, linear superposition must not apply. Clearly, these questions are a restatement of the ancient debate regarding the notion of free will in a deterministic or causal universe. Regardless of its origin, consciousness appears to bring entirely new, and paradoxical, aspects into the nature of physical reality. Thus in Chapter 5.3.1 we describe a scenario dealing with quantum mechanics, due to John Archibald Wheeler, that admits the possibility of current actions influencing the past if common-sensical, mechanistic interpretations are sought. Wheeler has likened the physical universe to a gigantic information system whose unfolding depends on how we question it.

1.4 A Historical Note

The ancient Vedic science of cognition, based primarily on consciousness examining itself, that has been traced in India to at least 2000 B.C. (Kak 1994c) led to a rich conceptual structure for the mind. The individual was seen as a system with the body, the energy field, mind, intellect, and emotion as different hierarchical levels. The mind itself was seen as a system consisting of the sense organs, an emergent characteristic of I-ness, associative memory, logic, and an underlying universal principle of consciousness. The whole conception was sometimes viewed as a dualistic system of matter and consciousness, but more often as a unitary (monistic) system where a universal consciousness is the ground-stuff of reality, but individual awareness is an emergent characteristic related to brain organization and behavioral associations.

There were other more advanced structural models that spoke of consciousness centers and agents. Animal and human minds were taken to be different in the sense that humans possess a language richer than that of animals; animals were also taken to be sentient, conscious beings. Details of this fascinating tradition, which is not widely known in the West, may be found in Sinha (1958), Aurobindo (1970), Dyczkowski (1987), and Kak (1993c, 1994b).

Can the categories of the Vedic cognitive science be related to current neurophysiological discoveries, we do not know. It is significant that the tradition claims to define a science of consciousness and the description is in terms of a holistic framework that is reminiscent of that of quantum mechanics. It is not surprising, therefore, that at least for Schrödinger, one of the creators of quantum mechanics, Vedic ideas were a direct source of inspiration (Schrödinger 1961; Moore 1989, pages 170-3). Schrödinger (1965) also believed that the Vedic conception provided the resolution to the paradox of consciousness. Walter Moore, Schrödinger's biographer, summarizes:

> The unity and continuity of Vedanta are reflected in the unity and continuity of wave mechanics. In 1925, the world view of physics was a model of a great machine composed of separable interacting material particles. During the next few years, Schrödinger and Heisenberg and their followers created a universe based on superimposed inseparable waves of probability amplitudes. This new view would be entirely consistent with the Vedantic concept of All in One. (Moore 1989, page 173)

1.5 Overview of the Article

This article presents several perspectives on the problem of machine and animal intelligence so as to set the framework for introducing the notion of a quantum neural computer. The emphasis of this article is on concepts and not on practical considerations. We do not speak about a design for a quantum neural computer, but the examination of this concept is very useful in understanding the limitations of current techniques of machine intelligence.

We begin by revisiting the Turing test and suggest that a new test that measures gradations of intelligence should be devised. Recent research on the cognitive capabilities of some animals, that validates the notion of levels of intelligence, is reviewed. This research establishes that certain cognitive abilities of animals, such as abstract generalization, are beyond the capabilities of conventional computers.

This is followed by a survey of the state of current neural network research. We consider new neuron and network models including one where, in analogy with quantum mechanics, the neuron outputs are taken to be complex. We also review rapid training of feedforward networks using prescriptive learning, chaotic dynamics in neuron assemblies, models of attention and awareness, and cytoskeletal microtubule information processing. Recent discoveries in neuroscience that cannot be placed in the reductionist models of biological information processing are examined. We learn from these studies that all biological systems cannot be viewed as connectionist circuits of components or systems; there exist dynamic structures whose definition is, in part, related to the environment and interaction with other similar systems. Although there are biological structures that are well modelled by artificial neural networks, in general biological systems define a concept of interdependence that is much stronger than the notion of connectionism that has been used in artificial neural networks. This is based on a recursive relationship between the organism and the environment. We call such interdependence as connectionism in its strong sense. One may postulate systems that are equivalent in their connectionist complexity to biological systems.

We define a quantum neural computer as a strongly connectionist system that is nevertheless characterized by a wavefunction. In contrast to a quantum computer, which consists of quantum gates as components, a quantum neural computer consists of a neural network where quantum processes are supported. The neural network is a self-organizing type that becomes a different measuring system based on associations triggered by an external or an internally generated stimulus. We consider some characteristics of a quantum neural computer and show that information is not a locally additive variable in such a computer. Models for a subneuron field that could explain the unity of awareness are reviewed. Uncertainty relations connecting the programs for structure and environment are described. Issues related to the design of intelligent systems, different in their programmatic structure from those found in nature, are discussed.

The article argues that quantum neural computing has parallels with brain behavior. Owing to this perspective, a considerable part of this article is devoted to pointing out the limitations of the conventional computing paradigm and the characteristics of quantum systems. We have adopted the philosophical point of view that quantum reality reveals itself through questions, as embodied by different experimental arrangements. It is natural then to first explore the information theoretic issues related to the concept of a quantum neural computer.

Chapter 2

Turing Test Revisited

If current computers are so fundamentally deficient in comparison with biological organisms at certain kinds of tasks, one might ask why has the discipline of computer science not generally recognized it. It appears that a pre-occupation with comparisons with human thought distracted researchers from considering a rich set of possibilities regarding gradations of intelligence. The prestige of the famed test for machine intelligence proposed by Alan Turing (1950), that now bears his name, was partly responsible for such a focus.

2.1 The Test

The Turing test proposes the following protocol to check if a computer can think: (1) The computer together with a human subject are to communicate, in an impersonal fashion, from a remote location with an interrogator; (2) The human subject answers truthfully while the computer is allowed to lie to try to convince the interrogator that it is the human subject. If in the course of a series of such tests the interrogator is unable to identify the real human subject in any consistent fashion then the computer is deemed to have passed the test of being able to think.

It is assumed that the computer is so programmed that it is mimicking the abilities of humans. In other words, it is responding in a manner that does not give away the computer's superior performance at repetitive tasks and numerical calculations.

The trouble with the popular interpretation of the Turing test is that it focused attention exclusively on the cognitive abilities of humans. So researchers could always claim to be making progress with respect to the ultimate goal of the program, but there was no means to check if the research was on the right track. In other words, the absence of intermediate signposts made it impossible to determine whether the techniques and philosophy used would eventually allow the Turing test to be passed.

In 1950, when Turing's essay appeared in print, his test embodied the goal for machine intelligence. But a statement of a goal, without a definition of the path that might be taken to reach it, detracts from its usefulness. Had specific tasks, that would have constituted levels of intelligence or thinking below that of a human, been defined then one would have had a more realistic approach to assessing the progress of AI.

Perhaps this happened because the dominant scientific paradigm in 1950, following old

Cartesian ideas, took only humans to be capable of thought. The rich tradition of cognitive philosophy that emerged in India about four thousand years ago (Kak 1993b) where finer distinctions in the capacity to think were argued was at that time generally unknown to AI researchers. The dominant intellectual ideas ran counter to the folk wisdom of all traditions regarding animal intelligence. That Cartesian ideas on thinking and intelligence were wrong has been amply established by the research on sub-human intelligence of the past few decades.

Another reason why Turing's test found a resonance in the intellectual climate of his times was the then prestige of operationalism as a scientific philosophy. According to the operationalist view, a machine is to be seen as having thoughts if its behavior is indistinguishable from that of a human. The ascendancy of operationalism came about from the standard interpretation of quantum mechanics, according to which one could only speak in terms of the observations by different experimental arrangements, and not ask questions about what the underlying reality was.

2.2 On Animal Intelligence

According to one view, generally attributed to Descartes, animal behavior is a series of unthinking mechanical responses. Such behavior is an automatic response to stimuli that originate in the animal's internal or external environments. In this view, complex behavior can always be reduced to a configuration of reflexes where thought plays no role. According to Descartes only humans are capable of thought since only they have the capacity to learn language.

Recent investigations of sub-human animal intelligence not only contradict Cartesian ideas, but also present fascinating riddles. It had long been thought that the cognitive capacities of the humans were to be credited in part to the mediating role of the inner linguistic discourse. Terrace (1985) claims that animals do think but cannot master language, so the question arises as to how thinking can be done without language:

> Recent attempts to teach apes rudimentary grammatical skills have produced negative results. The basic obstacle appears to be at the level of the individual symbol which, for apes, functions only as a demand. Evidence is lacking that apes can use symbols as names, that is, as a means of simply transmitting information. Even though non-human animals lack linguistic competence, much evidence has recently accumulated that a variety of animals can represent particular features of their environment. What then is the non-verbal nature of animal representations?...[For example] learning to produce a particular sequence of four elements (colours), pigeons also acquire knowledge about a relation between non-adjacent elements and about the ordinal position of a particular element. (Terrace, 1985, page 113)

Clearly the performance of animals points to representation of whole patterns that involves discrimination at a variety of levels. In an ingenious series of experiments, Herrnstein and Loveland (1964) were able to elicit responses about concept learning from pigeons. In another experiment, Herrnstein (1985) presented 80 photographic slides of natural scenes to pigeons who were accustomed to pecking at a switch for brief access to feed. The scenes were comparable but half contained trees and the rest did not. The tree photographs had full views of single and multiple trees as well as obscure and distant views of a variety of types. The slides were shown in no particular order and the pigeons were rewarded with food if they pecked at the switch in response to a tree slide; otherwise nothing was done. Even before all the slides had been shown the pigeons were able to discriminate between the tree and the non-tree slides. That this ability, impossible for any machine to match, was not somehow learnt through the long process of evolution and hardwired into the brain of the pigeons, another experiment was designed to check the discriminating ability of pigeons with respect to fish and non-fish scenes and once again the birds had no problem doing so. Over the years it has been shown that pigeons can also distinguish: (1) oak leaves from leaves of other trees, (ii) scenes with or without bodies of water, (iii) pictures showing a particular person from others with no people or different individuals.

Herrnstein (1985) summarizes the evidence thus:

> Pigeons and other animals can categorize photographs or drawings as complex as those encountered in ordinary human experience. The fundamental riddle posed by natural categorization is how organisms devoid of language, and presumably also of the associated higher cognitive capacities, can rapidly extract abstract invariances for some (but not all) stimulus classes containing instances so variable that we cannot physically describe either the class rule or the instances, let alone account for the underlying capacity.

Amongst other examples of animal intelligence are mynah birds who can recognize trees or people in pictures, and signal their identification by vocal utterances—words—instead of pecking at buttons (Turney 1982), and a parrot who can answer, vocally, questions about shapes and colors of objects, even those not seen before (Pepperberg 1983).

The question of the relationship between intelligence and consciousness may be asked. Griffin infers animal consciousness from a variety of evidence:

> I. Versatile adaptability of behavior to novel challenges;
> II. Physiological signals from the brain that may be correlated with conscious thinking;
> III. Most promising of all, data concerning communicative behavior by which animals sometimes appear to convey to others at least some of their thoughts. (Page 27, Griffin 1992)

Consciousness implies using internal images and reconstructions of the world. Purposive behavior is contingent on these internal representations. These representations may be based on the stimulus from the environment, memories, or anticipation of future events.

2.3 Gradation of Intelligence

The insight from experiments of animal intelligence, that one can attempt to define different gradations of cognitive function, is a useful one. The theory of evolution not only posits the evolution of human brains but also that of behavior. In this theory one would expect to see different levels of intelligence.

It is obvious that animals are not as intelligent as humans; likewise, certain animals appear to be more intelligent than others. For example, pigeons did poorly at picking a pattern against two other identical ones, as in picking an A against two B's. This is a very simple task for humans. Herrnstein (1985) describes how they seemed to do badly at certain tasks:

1. Pigeons did not do well at the categorization of certain man-made and three-dimensional objects.

2. Pigeons seem to require more information than humans for constructing a three-dimensional image from a plane representation.

3. Pigeons seem to have difficulty with dealing with problems involving classes of classes. Thus they do not do very well with the isolation of a relationship among variables, as againt a representation of a set of exemplars.

In a later experiment Herrnstein et al (1989) trained pigeons to follow an abstract relational rule by pecking at patterns in which one object was inside, rather than outside of a closed linear figure. It is to be noted that pigeons and other animals are made to respond in extremely unnatural conditions in Skinner boxes of various kinds. The abilities elicited in research must be taken to be merely suggestive of the intelligence of the animal, and not the limit of it.

Animal intelligence experiments suggest that one can speak of different styles of solving AI problems. Are the cognitive capabilities of pigeons limited because their style has fundamental limitations? Or is the cognitive style of all animals similar and the differences in their cognitive capabilities arise from the differences in the sizes of their mental hardware? And since current machines do not, and cannot, use inner representations, is it right to conclude that their performance can never match that of animals?

This raises the question whether one can define a hierarchy of computational tasks that could be quantified as varying levels of intelligence. These tasks could be the goals defined in a sequence that could be set for AI research. If the simplest of these tasks proved intractable for the most powerful of computers then the verdict would be clear that computers are designed based on principles that are deficient compared to the style at the basis of animal intelligence.

2.4 Animal and Machine Behavior

2.4.1 Linguistic behavior

A classification of computer languages in the 1950s provided an impetus to examine the nature of animal communication. The focus of these studies on primates was not to decode this communication in their natural setting, but rather to see if animals could be taught to communicate with humans. Since the development of language, as part of behavior, is critically a component of social interaction, studies in unnatural laboratory settings are inherently limited.

Many projects have examined the grammatical competence in apes (e.g. Savage-Rumbaugh et al 1985, Griffin 1992). This research has clarified questions related to what we understand by language. Responding to signals in a consistent way by the subject is not sufficient to demonstrate linguisitic competence. It is essential for us to know that the subject uses different symbols for naming.

Savage-Rumbaugh (1986) argues that to qualify as a word, a signal must satisfy the following attributes: (1) it must be an arbitrary symbol that stands for some object, activity, or relationship; (2) it must be used intentionally to convey knowledge; and (3) the recipient must be able to decode and respond appropriately to the symbols. Many early ape language projects did not satisfy all these attributes. Savage-Rumbaugh et al (1985) report striking success in the acquisition of symbolic skills on a lexigram keyboard by an ape named Kanzi.

How to grade competence in symbolic manipulation remains an interesting problem. If this were possible, one would have the picture of different levels of symbol manipulation leading finally to the development of language. Grammatical competence could likewise be classified in different categories. The relationship of these levels to the ability to solve AI problems could be explored. A hierarchy of intelligence levels will be useful in the classification of animal behavior.

2.4.2 Recursive behavior

Another useful perspective on animal behavior is its recursive nature. Life can be seen at various levels, but consciousness is a characteristic of the individual alone.

Considering this from the bottom up, animal societies have been viewed as "superorganisms." For example, the ants in an ant colony may be compared to cells, their castes to tissues and organs, the queen and her drones to the generative system, and the exchange of liquid food amongst the colony members to the circulation of blood and lymph. Furthermore, corresponding to morphogenesis in organisms the ant colony has sociogenesis, which consists of the processes by which the individuals undergo changes in caste and behavior. Such recursion has been viewed all the way up to the earth itself seen as a living entity. Parenthetically, it may be asked whether the earth itself, as a living but unconscious organism, may not be viewed like the unconscious brain. Paralleling this recursion is the individual who can be viewed as a collection of several "agents" where these agents have sub-agents which are the sensory mechanisms and so on. But these agents are bound together and this binding defines consciousness.

A question that arises out of this reasoning is whether an organism as well as a superorganism could be simultaneously conscious. It appears that such nested consciousness should be impossible. In multiple personality disorder, the person's consciousness remains a unity. Such a person might claim that "yesterday I was Alice and today I am Mary," but each time the assertion would be in terms of an indivisible consciousness. In split-brain patients, the minor hemisphere does not have a conscious awareness. In other words, there is no nested consciousness.

2.4.3 Logical tasks

Logical tasks are easy for machines whereas AI tasks are hard. We have two perspectives. The first is: Why build machines in a hierarchy that mimics nature? After all, wheeled machines do better than animals at locomotion. The second is that there is something fundamental to be gained in building machines that have recursively defined behavior in the manner of life, and the same facility for solving AI problems that subhumans possess. If the basis for animal competence at AI tasks were understood then new kinds of AI machines could be designed.

Logical tasks are difficult for animals to complete. Human language is more than reflexive behavior, and it has a logical component. A logical structure underpins human language. On the other hand, when apes have been taught associations of symbols with objects, they have found it hard to string these together into rule-based combinations that could be termed language. Although communication amongst animals seems to be very rich, it is clear that animal communication lacks many features that are present in human language. To see these differences in terms of less developed neural structures would connect neural hardware with specific function.

Chapter 3

Neural Network Models

3.1 The Brain

The central nervous system of vertebrates is organized in an ascending hierarchy of structures—spinal cord, brain stem, and brain. A human brain (Figure 1) contains about 10^{11} neurons. The structure, composition, and functioning of central nervous systems in all vertebrates have general similarity. The peripheral nervous system consists of peripheral nerves and the ganglia of the autonomic nervous system.

At the gross level, the brain is bilaterally symmetric and the two hemispheres are connected together by a large tract of about half a billion axons called the corpus callosum. At the base are structures such as the medulla, which regulates autonomic functions, and the cerebellum, which coordinates movement. Within lies the limbic system, a collection of structures involved in emotional behavior, long-term memory and other functions.

The convoluted surface of the cerebral hemispheres is called the cerebral cortex. The cerebral cortex is a part of the brain more developed in humans than in other species. Approximately seventy percent of all the neurons in the central nervous systems of primates are found in the cortex. This is where the most complex processing and coding of sensory information occurs. The most evolutionary ancient part of the cortex is part of the limbic system. The larger, younger neocortex is divided into frontal, temporal, parietal, and occipital lobes that are separated by deep folds. Each of these lobes contains yet further subdivisions. The parietal, temporal, and occipital lobes receive sensory information such as hearing and vision. These sensory-receiving areas occupy a small portion of each lobe, the remainder being termed the association cortex. Each sensory area of the neocortex receives projections primarily from a single sensory organ. Thus, the visual-receiving area in the occipital lobe processes sensations received by the retina; the auditory-receiving area in the temporal lobe processes sensations received by the cochlea; and the body-sense-receiving area in the parietal lobe processes sensations received by the body surface. Within each sensory-receiving area, the sense-organ projections form a map of the sensory organ on the neocortical surface.

3.2 Neurons and Synapses

The nervous tissue is made up of two types of cells: neurons and satellite cells. In the central nervous system the satellite cells are called neuroglia and in the periphery, Schwann cells. The satellite cells provide insulating myelin sheathing around the neurons. This insulation increases the propagation velocity of the electrical signals.

A neuron consists of tree-like networks of nerve fiber called dendrites, the cell body (or soma), and a single long fiber called the axon, which branches into strands and substrands (Figure 2). At the end of these substrands are the axon terminals, the transmitting ends of the synaptic junctions, or synapses (Figure 3). The receiving ends of these junctions can be found both on the dendrites and on the cell bodies themselves. The axon of a typical neuron makes about a thousand synapses with other neurons.

It may be argued that brain computations must really be seen to take place in the synapses, rather than neurons. This change in point of view has important philosophical implications. Not only are there more synapses than neurons, say about 10^{15}, but this changed perspective can be taken to mean a view of the brain as a chemical computer rather than an electrical digital system. This is an issue that will not be explored in this article.

Two major types of synapse are electrical and chemical. In electrical synapses the channels formed by proteins in the presynaptic membrane are physically continuous with the similar channels in the postsynaptic membrane, or the synaptic gap is only about two nanometers, allowing the two cells to share some cytoplasm constituents and electrical currents.

At chemical synapses, the gap of about 20 to 30 nanometers between the presynaptic and postsynaptic membranes prevents a direct flow of current. The signal is now transmitted to the next neuron through an intermediary chemical process in which neurotransmitter substances are released from spherical vesicles in the synaptic bouton. This raises or lowers the electrical potential inside the body of the receiving cell. When this potential reaches a threshold, a pulse or action potential of fixed strength and duration is sent down the axon. This constitutes the firing of the cell. After firing, the cell cannot fire again until the end of the refractory period.

Ionic currents flowing across surface membranes of nerve cells lead to electrical potential changes. Ions such as sodium, potassium, calcium, and chloride participate in this process. The nerve maintains an electrical polarization across its membrane by actively pumping sodium ions out and potassium ions in. When ion channels are opened by voltage changes, neurotransmitters, drugs and so on, sodium and potassium rapidly flow through the channel creating a depolarization. Action potentials occur on an all or none basis from integration of dendiritic input at the cell body region of the neuron. The frequency of firing is related to the intensity of the stimulus. Action potential velocity is fixed for given axons depending on axon diameter, degree of myelinization, and distribution of ion channels. Typical values of action potential velocity is 100 meters per second.

The neurons come in a great variety of structures. The diversity is even greater if molecular differences are considered. Although all cells contain the same set of genes, individual cells exhibit, or activate, only a small subset of these genes. Selective gene expression has been found even in seemingly homogeneous populations of neurons. Evidence does not support the thesis that each neuron is unique; but certainly the brain cannot be viewed as a collection of identical elements. An order is imposed on the enormous diversity of the neurons and their connection patterns by the nested arrangement of certain circuits and subcircuits (Mountcastle 1978, Sutton et al 1988). For example, neurons of similar functions are grouped together in columns that extend through the thickness of the cortex. A typical module is the visual cortex, which could contain more than 100,000 cells, the great majority of which form local circuits devoted to a particular function.

3.2.1 The McCulloch-Pitts model

McCulloch and Pitts (1943) proposed a binary threshold model for the neuron. The model neuron computes a weighted sum of its inputs from other inputs and outputs a one or a zero according to whether this sum is above or below a certain threshold. The updating of the states of a network of such neurons was done according to a clock. Although the McCulloch-Pitts (MP) neurons do not model the biological reality well, networks built up of such neurons can perform useful computations. Variants of the MP neurons have graded response; this is achieved if the neuron performs a sigmoidal transformation on the sum of the input data. But networks of such graded neurons do not provide any fundamental computing advantage over binary neurons.

Networks of MP-like neurons have been extensively used to model brain circuits in the hope that these networks will somehow capture the essentials of the processing done in the brain. Various kinds of networks have been used in these models.

3.3 Feedback and Feedforward Networks

A broad distinction may be made between feedback and feedforward neural networks. In feedback networks the computation may be taken to be over when the networks has reached a stable state. The updating for the network may be represented by the rule

$$X_{new} = f(X_{old})$$

where f represents the concatenation of the transformation by the synaptic interconnection weight matrix and the sign or the sigmoidal function. This updating may be viewed to occur either synchronously or asynchronously. The operation of such a network in the asynchronous mode may be viewed as a descent to minimum energy value in the state space. Such minimas can thus represent stored memories in a feedback model. Learning shapes attraction areas around exemplars. The notion of the attraction basis associated with each energy minimum provides the model for generalization. The major limitation of the feedback model is that along with the useful memory, a very large number of spurious memories are automatically generated.

Feedforward networks perform mapping from an input to an output. Once the training input-output pairs have been learned to define certain classes, this network also has the capacity to generalize. This means that neighbors of the training inputs are automatically classified.

Although feedforward networks are quite popular as far as signal processing applications are concerned, they suffer from inherent limitations. This is owing to the fact that such networks perform no more than mapping transformations or signal classification. Feedback networks, on the other hand, offer greater parallels with biological memory, and biological structures, such as the vision system, have aspects that include feedback. It is due to this reason that recurrent networks, which are feedforward networks with global feedback, are being increasingly studied. If one accepts the proposition that artificial neural architectures would gain from insights obtained in the study of biological systems, then a further study of feedback networks is called for.

The neuron model that we examine in this chapter is the on-off or the bipolar MP model. This model and its variant, where the neuron output can be any analog value over a range, have generally been used for signal processing artificial neural networks. But biological neurons exhibit extremely complex behavior and on-off type neurons are a gross simplification of the biological reality (Libet 1986). For example, neuron output is in terms of spikes, this output cannot be taken to be held to any fixed value. There is, furthermore, the question of the refractory phase of a neuron's response. If one considers waves of neural activity, then a continuum model may be called for (Milton et al 1994). Neuron outputs may also be considered to arise from the partial response of other neurons. In addition, there exist neurons that remain quiescent in a computation. We look at some of these issues and investigate how the bipolar model can be made more realistic.

In many artificial intelligence problems, such as those of speech or image understanding, local operators can be properly designed only if some understanding of the global information is available. Networks of MP neurons do not provide any choice as far as resolution of the data is concerned. A generalization of the MP model of neurons appears essential in order to account for certain aspects of distributed information processing in biological systems. One particular generalization allows one to deal with some recent findings of Optican and Richmond (1987) that indicate that in neuron activity the spike rate as well as spike distribution carry information. This supposes complex valued neurons with relationship between the real and imaginary activations; this can form the first step in the storage of patterns in such a fashion so that together with these corresponding context clues are represented in the complex domain.

3.4 Iterative Transformations

Consider a fully connected neural network, composed of n neurons, with a symmetric synaptic interconnection matrix T_{ij} where $T_{ii} = 0$. The matrix T is obtained by Hebbian learning, which is an outer product learning rule, or its generalization called the delta rule. In Hebbian learning

$$T_{ij} = \sum_{k=1}^{l} X_i^k X_j^k$$

where l is the number of memories, and X^k represents the kth memory. Let x_i be the output of neuron i. The updated value of this neuron is

$$x_i^{'} = f(\sum_{j=1}^{n} T_{ij} x_j)$$

where f is a sigmoid or a step function. Without loss of generality we will now confine ourselves to the sign function (sgn) for f and we will use the convention that $sgn(0) = 1$. Also note that the updating of the neurons may be done in any random order, in other words asynchronously.

The characteristics of such a feedback network have parallels with that of one-variable iterative systems. Consider the equation $g(x) = 0$, the solutions for x defining the roots of the equation. Let x_0 be a point near a root x. Then one may expand $g(x)$ in a Taylor series so that

$$g(x) = g(x_0) + (x - x_0)g^{'}(x_0)/1! \ + (x - x_0)^2 \ g^{''}(x_0)/2! \ + ...$$

For values of x near x_0, $(x - x_0)$ will be small. Assuming that $g^{'}(x_0)$ is large compared to $(x - x_0)$, and that $g^{''}(x_0)$ and the higher derivatives are not unduly large, the above equation may be rewritten as a first approximation:

$$g(x) = g(x_0) + (x - x_0)g^{'}(x_0)$$

Since we seek x such that $g(x) = 0$, we can simplify this equation as:

$$x = x_0 - g(x_0)/g^{'}(x_0)$$

The new x will be closer to the root than x_0 and this forms the basis of the Newton-Raphson iteration formula:

$$x_{n+1} \ = \ x_n \ - g(x_n)/g^{'}(x_n)$$

If the equation $g(x) = 0$ has several solutions then the variable x will get divided into the same number of ranges. Starting in a specific range would take one to the corresponding solution. Each range can therefore be seen as an attraction basin.

This may be generalized by the consideration of the transformation on the complex plane. For an analytic function one can use the Taylor series to give:

$$z_{n+1} \ = \ z_n \ - g(z_n)/g^{'}(z_n)$$

For $g(z)$, a nth degree polynomial, one would see n attraction basins. A good example is $g(z) = z^3 - 1 = 0$ which leads to the iterative mapping:

$$z_{n+1} \ = \ (2z^3 \ + 1)/3z^2$$

This has three attraction basins. The boundaries of the basins are fractals. Das (1994) has examined how neural networks that are a generalization of the above idea can be designed.

We see that for an iterative transformation the concepts of attractors and attraction basins are defined similar to those for the feedback neural network. However, we do not get a regular structure of connected points defining an attraction basin. Moreover, the nature of the attraction basins is dependent on the iterative transformation although the underlying polynomial may be the same.

3.4.1 Neuron models and complex neurons

Biological neurons are characterized by extremely complex behavior expressed in a distribution of voltage and time dependence of current. Different types of membrane currents may be involved. Perhaps the simplest model is the classic Hodgkin-Huxley model of the squid giant axon. It has three main characteristics: 1) an action potential, 2) a refractory period after an action potential during which the slow recovery of the potassium and sodium conductance has a great impact on electrical properties, and 3) the ability to capacitively integrate incoming current pulses. This model is approximated by two first order differential equations in terms of four time-dependent dynamical variables. It has been suggested that a further approximation in terms of two dimensions is reasonably good. This description shows that the MP binary or bipolar neurons do not capture the intricacies of even the simple Hodgkin-Huxley model.

Considering the problem at a gross level, it has been suggested that neurons carry information in the distribution as well as the number of spikes in a response (Optican and Richmond 1987). The nature of the relationship between these two types of information is not clear. In the most general case one may take these two channels of information to provide us with independent information, although they may actually be connected through some transformation.

On the other hand, from an analytical perspective related to information processing, there exists the problem of local/global information linkage whereby local information features are linked into global concepts. This may be seen to occur through stages of transformation so that information is delivered to neurons that deal directly with global concepts.

One may use a straightforward generalization of the standard neuron model to include complex sequences where the real and the imaginary activations are related to each other. It may be assumed that usual observations, being based on time-averaged behaviour, ignore the information in the distribution of the spikes. One might further assume that the real activation is related to the spike rate and the imaginary activation expresses the information in the spike distribution. In the general case the real and the imaginary quantities are connected to each other through a transformation. Note that each neuron in the standard network model does receive information from all other neurons and, therefore, there exists the potential of the delivery of global information at each neuron. However, this information is presented in a mixed fashion from where it may not be retrievable. It is for this reason it is useful to postulate an explicit transformation.

Let the stored memories be represented by z^j where

$$z^j = x^j + iy^j$$

and

$$y^j = g(x^j)$$

The transformation g would in general be nonlinear. As a starting point this transformation may be taken to be Fourier followed by a saturation function. Transformations such as Fourier decorrelate redundant patterns. This means that far fewer components of the pattern in the transformed domain are necessary to represent most of the pattern information.

For ease of illustration one may consider bipolar neurons. The Hebbian learning rule for the interconnection weight matrix T can be stated to have the form:

$$T = \sum_{j=1}^{m} z^j z^{j*'}$$

The neural network operates in the usual mode, $Z_{new} = sgm\{TZ_{old}\}$.

The only difference with the standard Hebbian rule is that the column memory vector is multiplied with its complex conjugate transposed form. Also, if the staggering is done correctly for the patterns in a set, Hebbian learning will automatically associate the imaginary components of these patterns in a concatenated form.

3.4.2 Chaotic maps on the complex plane

The dynamics of a synchronous feedback network is represented by $X_{new} = f(X_{old})$, where f is a linear map followed by a sigmoidal function and X is a n-component vector. In a synchronous update one can encounter oscillatory behavior. When the components of the X vector are continuous then the behavior can be chaotic. This will be seen in the context of a specific non-linear transformation below.

Consider $z_{n+1} = f(z_n)$, an iterative transformation on the complex plane. Such an iterative transformation must be considered for points on the unit circle because for $|z| < 1$ the asymptotic value would end up at 0 whereas for $|z| > 1$ it would end up at ∞. The set of points on the unit circle is the Julia set.

Consider now

$$z_{n+1} = z_n^2$$

for the unit circle. This corresponds to:

$$z_{n+1} = x_{n+1} + iy_{n+1} = x_n^2 - y_n^2 + i2x_n y_n$$

Since $x_n^2 + y_n^2 = 1$, we obtain a pair of transformations for the x and y variables:

$$x_{n+1} = 2x_n^2 - 1, \quad -1 \leq x_n \leq 1$$

and

$$y_{n+1}^2 = 4y_n^2\ (1 - y_n^2)$$

Let us define another variable $r_n = y_n^2$, then the above equation can be rewritten as:

$$r_{n+1} = 4r_n(1 - r_n), \quad 0 \leq r_n \leq 1$$

This represents the well known logistic map. If r_o is represented by an irrational point then the iterative transformation will never return to its starting point. In other words, the logistic map will generate an infinite period sequence, or a chaotic sequence. The same will also be true of the other maps listed above.

To see this consider $z_n = |z_n| \; exp(i\theta)$. It is clear that this map implies multiplying the angle θ by 2 at each iteration. Now, if the starting angle θ_o is represented in a binary form, then each step implies shifting the expansion one bit to the left. Therefore, if the original angle is represented by an infinite expansion (as for an irrational number) the iterative transformation will lead to chaotic behavior.

It is also clear from the above that the two-dimensional structure of the Julia sets merely expresses the randomness properties in different scales of the underlying irrational number.

One may consider different generalizations of the map on the unit circle. Thus one may consider

$$x_{n+1}^2 = \lambda x_n^2 \left(1 - \alpha + \frac{\alpha}{\beta} y_n^2\right)$$

$$y_{n+1}^2 = \beta - \lambda x_n^2 y_n^2$$

When $\alpha = \beta = 1$ and $\lambda = 4$, we obtain the logistic map. When only $\alpha = \beta = 1$, we have:

$$r_{n+1} = \lambda r_n (1 - r_n)$$

To determine the stability of a fixed point one looks at the slope $|f'(r)|$ evaluated at the fixed point. The fixed point is unstable if $|f'| > 1$. For this logistic function when $1 < \lambda < 3$, there are two fixed points $r = 0, (\lambda - 1)/\lambda$ of which $r = 0$ is unstable and the other point is stable.

For $\lambda > 3$ the slope at $r = (\lambda - 1)/\lambda$, which is $f' = 2 - \lambda$ becomes greater than 1 and both equilibrium points become unstable. At $\lambda = 3$ the steady solution becomes unstable and a two-cycle orbit becomes stable. For $3 < \lambda < 4$, the map shows many multiple period and chaotic motions. The first chaotic motion is seen for $\lambda = 3.56994....$ Before the onset of chaos at this value if the nth period doubling is obtained for λ_n, we get the well-known result that in the limit

$$\frac{\lambda_{n+1} - \lambda_n}{\lambda_n - \lambda_{n-1}} \rightarrow 4.6692016...$$

which is the Feigenbaum number (Feigenbaum 1978). This represents an example of the period-doubling route to chaos. There are other routes to chaos, such as quasi-periodicity and intermittency, that do not concern us in the present discussion.

Considering that chaos in the brain arises from a similar mechanism, one might ask whether the universality of the route to chaos provides a certain normative basis.

3.4.3 Implication

The study of complex valued neurons and iterative transformations shows that these alone are not capable of defining a rich enough framework in which cognitive functions can be understood. Thus a complex network may be replaced by four real ones and so nothing fundamental is gained by such a generalization. That should not be surprising since the transition from classical to quantum mechanics is much more than generalization from scalar to complex representations.

3.4.4 Performance of interacting units

We now consider the question of a computation being carried out by a neural system of which a feedback network is a part. A network that deals with non-trivial computing tasks will be composed of several decision structures. If the neural network computation can be expressed in terms of a straightforward sequential algorithm then it is inappropriate to call the computation style neural. We may call neurons that perform decision computations to be cognitive neurons. These cognitive neurons may take advantage of information stored in specialized networks some of which perform logical processing. There would be other neurons that simply compute without cognitive function. Such cognitive neurons could lie at the end of the brain's many convergence regions. It is essential to assume that many cognitive neurons compute in parallel for a computation to be neural. It was shown in Kak (1992, 1993b) that independent cognitive centers will lead to competing behaviour.

In the cortex certain cells seem to be responsive to specific inputs. In other words, certain neurons process global information. Thus Hubel and Wiesel (1968) have shown that certain cells in the visual cortex of a cat are responsive to lines in the visual field which have a particular angle of slope. Other cells respond to specific colors, and others respond to the differences between what each eye has received, leading to depth perception. One might assume that neurons are organized hierarchically, with the higher level neurons dealing with more abstract information. It may be assumed that the two different kinds of neurons do interact with each other? Does the information processing of such a network have special characteristics or limitations?

Another interesting phenomenon is where damage to parts of the visual cortex makes a person blind in the corresponding part of the visual field. However, experiments have shown that often such subjects are able to "guess" objects placed in this region with accuracy even though they are blind in that region. The information from the retina is also processed by obscure regions lying in the lower temporal lobe. Clearly, these independent cognitive centres supply information to the subject's consciousness. But this is done without direct awareness.

Consider two non-linearly interacting cognitive centers A and B with three courses of action labeled 1, 2, and 3 with energies $(-1, 2, 1)$ and $(3, 5, -1)$ respectively. If the non-linear interaction is represented by multiplication at the systems level then the system states will be shown as in the table below.

Energy	A_1	A_2	A_3
B_1	-3	6	3
B_2	-5	10	5
B_3	1	-2	-1

The optimal system performance is represented by A_1 B_2 which corresponds to the minimum energy of -5. However, the two centers would choose A_1 and B_3 considering individual energy distributions.

This is indication that an energy function cannot be associated with a neural network that performs logical processing. But if there is no energy function, then how do the neurons operate cooperatively?

Neural network models do not deal with or describe inner representations, so from our experience with animal intelligence (Chapter 2.2) it is clear their behavior cannot be purposive. Current theory does not show how such inner representations could be developed.

Chapter 4

The Binding Problem and Learning

4.1 The Binding Problem

How is a stimulus recognized if it belongs to a class that has been learnt before? If the stimulus triggers the firing of certain neurons, how is the firing information "gathered" together to lead to the particular perception? Considering visual perception, there exist an unlimited number of objects that one is capable of seeing. Early theories postulated "grandmother neurons," one for each image; this concept leads to a homunculus—the person's representation inside the brain which observes and controls, and also represents the self. The postulation of grandmother neurons leads to logical problems: How can we have a number that would exhaust all possible objects, and how does the homunculus organize all the sensory input that is received? Thus grandmother neurons cannot exist, although there might exist some specialized structures for certain features. If a large number of neurons fire in response to an image, the problem of how this set fires in unison, defining the perception of the image, is called the "binding problem." No satisfactory solution to the binding problem is known at this time.

In reality the problem is much worse than the binding problem as stated above. The bound sets are in turn bound at a higher level of abstraction to define deeper relationships.

The reductionist approach seeks explanations of brain behavior in terms of a sum of the behaviors of lower level elements. Brain behavior has traditionally been examined at three different levels of hierarchy: (i) synaptic level or in terms of biochemical changes; (ii) neuron level as in maps of the retina in the visual cortex; (iii) that of neural networks where the information is distributed over entire neuron cell assemblies. But the explanation of brain function in terms of behavior at these levels has proved to be inadequate.

Skarda and Freeman (1990) have argued that while reductionist models have an important function, they suffer from severe limitations. They claim that "perceptual processing is not a passive process of reaction, like a reflex, in which whatever hits the receptors is registered inside the brain. Perception does not begin with causal impact on receptors; it begins within the organism with internally generated (self-organized) neural activity that, by re-afference, lays the ground for processing of future receptor input... Perception is a self-organized dynamic process of interchange inaugurated by the brain in which the brain fails to respond to irrelevant input, opens itself to the input it accepts, reorganizes itself, and then reaches out to change its input (page 279)." In other words, the neural networks that perform the perceptual processing function together within a holistic framework that defines the self-organizing principle.

4.2 Learning

4.2.1 The biological basis

Two types of learning or memories have been described by researchers. Memories that require a conscious record have been termed declarative or explicit. Memories where conscious participation is not needed are called nondeclarative or implicit. It is not known whether other types of memories exist.

Explicit learning is fast and may take place after only one training trial. It involves associations of simultaneous stimuli and defines memory of single events. In contrast, implicit learning is slow and requires repetition over many trials. It often connects sequential stimuli in a temporal sequence. Implicit learning is expressed by improved performance on certain tasks without the subject being able to describe what he has learnt. Experiments indicate that the storage of both these memory types proceeds in stages. "Storage of the initial information, a type of short-term memory, lasts minutes to hours and involves changes in the strength of existing synaptic connections. The long-term changes (those that persist for weeks and months) are stored at the same site, but they require something entirely new: the activation of genes, the expression of new proteins and the growth of new connections." (Kandel and Hawkins 1994)

Since gene activation at neuron sites, and the reorganization of the brain, are two of the consequences of learning, the question arises: Do we perceive as we do as a consequence of the genetic potential within us via its embodiment in the organization of the brain? In other words, is learning a part of the response to the universe? If the ape, or the pigeon, does not have the appropriate genes to activate specific development and reorganization of the brain it will not respond to certain patterns in its environment.

Hebb's rule that coincident activity in the presynaptic and postsynaptic neurons strengthens the connection between them, is one learning mechanism. According to another rule, called the pre-modulatory associative mechanism, the synaptic connection can be strengthened without activity of the postsynaptic cell when a third neuron, called a modulatory neuron, acts on the presynaptic neuron. The associations are formed when the action potentials on the presynaptic cell are coincident with the action potentials in the modulatory neuron.

4.2.2 Prescriptive learning

Although neural networks and other AI techniques provide good generalization in many signal and image recognition applications, these methods are nowhere close to matching the performance of biological neural systems. As explained in Chapter 2, pigeons, parrots and other animals can quite easily perform abstract generalization, recognizing abstract classes such as those of a tree, a human, and so on which lies completely beyond the capability of any AI machine. This raises the question if current learning strategies being employed are fundamentally deficient. One would expect that biological systems in their evolution through millions of years would have evolved learning mechanisms that are optimal in a certain sense. This section examines biological learning to conclude that artificial neural networks do not learn in the same manner as animals. We claim that a two-step learning scheme, where the first step is a prescriptive design, offers a more versatile model for machine learning.

Our starting point is developmental neurobiology. During embryogenesis several kinds of cell death occur. One is histogenetic cell death, or the death of neurons that have ceased proliferation and have begun to differentiate; this is what is of relevance for our discussion. One may also speak of phylogenetic cell death, responsible for the regression of vestigal organs or of larval organs during metamorphosis (such as the tadpole tail), and morphogenetic cell death, which involves degeneration of cells during massive changes in shape as in bending and folding of tissues and organs during early embryogenesis.

According to Oppenheim (1981): "The neurons that exhibit cell death include such a wide variety of cells ... that it is impossible at this time to argue that any particular feature is characteristic of all neurons exhibiting cell death. Although it has been suggested that cell death may only occur among neurons that have greatly restricted or limited numbers of synaptic targets, the variety of cell types [that exhibit cell death] would appear to contradict such a suggestion."

Soon after the issuance of Darwin's *The Origin of Species* it was suggested by T.H. Huxley, G.S. Lewes, and Wilhelm Roux that Darwinian ideas of natural selection might apply to the cells and tissues of the developing organism. The Spanish neuroanatomist Ramòn y Cajal also suggested in 1919 that "during neurogenesis there is a kind of competitive struggle among the outgrowth (and perhaps even among the nerve cells) for space and nutrition. The victorious neurons ... would have the privilege of attaining the adult phase and of developing stable dynamic connections." But none of these early pioneers anticipated the possiblity of massive neuronal death during normal development which was reported first by Hamburger and Levi-Montalcini (1949).

We suggest that cell death is a result of many neurons being rendered redundant during a stage of reorganization that follows the first phase of development of the system. During the first phase all the neurons fulfil a specific function; this is followed by a structural reorganization. This two-step strategy can be used to devise a new approach to machine learning.

4.2.3 Reorganization

The two-step learning scheme allows us to view the developmental process in a different light. The first step consists of prescriptive learning which defines the layout and the synaptic weights of the interconnections. This learning may be called the biologically programmed component of the learning.

The nature of the prescriptive learning requires many more neurons than are necessary in a more efficiently organized network. During the second stage of learning a reorganization of the network takes place that renders a large number of neurons redundant. These are the neurons that suffer death as they do not receive appropriate signals from target neurons

(Purves 1988).

In a more general setting this new learning strategy may be termed a structured approach. We begin with a general kind of architecture for the intelligent machine which is refined during a subsequent phase of training.

4.2.4 The example of feedforward neural networks

Backpropagation is the commonly used method for training feedforward networks. But this does not explicitly use any specific architecture related to the nature of the mapping.

It has recently been shown that a prescriptive learning scheme can be devised (Kak 1993a, 1994a). In its basic form this does not require any training whatsoever. But this form does not generalize. The weights are now changed slightly by adding random noise which allows the network to generalize. Characteristics of this model are summarized in the work of Madineni (1994) and Raina (1994). This variant scheme provides significant learning generalization and a speed-up advantage of 100 over fastest versions of backpropagation. The penalty to be paid for this speed-up is the much larger number of hidden neurons that are required.

The second step in this training would be to reorganize from the architecture obtained using prescriptive learning and reduce the number of hidden neurons.

The speed of learning, and the simplicity of this procedure make the new model a plausible mechanism for biological learning in certain structures. Interesting issues that remain to be investigated include further development of this approach so that one can obtain continuous valued outputs and the development of competitive learning. But we cannot expect such models to provide any insight regarding holistic behavior.

But to explain the nature of memory at the level of the neural network model, one needs to understand learning at the more fundamental level of organization of the brain.

4.3 Chaotic Dynamics

Freeman and his associates (e.g. Freeman 1992, 1993) have argued that neuron populations are predisposed to instability and bifurcations that depend on external input and internal parameters. Freeman (1993) claims that chaotic dynamics makes it "possible for microscopic sensory input that is received by the cortex to control the macroscopic activity that constitutes cortical output, owing to the the selective sensitivity of chaotic systems to small fluctuations and their capacity for rapid state transitions."

One significance of this work is to point out the gulf that separates simple input-output mapping networks used in engineering research from the complexity of biological reality. But the claim that chaotic dynamics in themselves somehow carry the potential to explain macroscopic cortical activity relating to the binding problem seems to be without foundation.

The nonlinearities at the basis of chaos are seen as the basis of a self-organizing principle. Considering these ideas for the olfaction in a rabbit, Freeman (1993) says: "The olfactory system maintains a global chaotic attractor with multiple wings or side lobes, one for each odor that a subject has learned to discriminate. Each lobe is formed by a bifurcation during learning, which changes the entire structure of the attractor, including the pre-existing lobes and their modes of access through basins of attraction. During an act of perception the act of sampling a stimulus destabilizes the olfactory bulbar mechanism, drives it from the core of its basal chaotic attractor by a state transition, and constrains it into a lobe that is selected by the stimulus for as long as the stimulus lasts, on the order of a tenth of a second... In this way the cortical response to a stimulus is "newly constructed"... rather than retrieved from a dead store."

The above theory is very attractive for the olfactory system but it remains very limited in its scope. It is hard to see how it could be generalized for more complex information processing.

4.4 Attention, Awareness, Consciousness

Milner (1974) speculated that neurons responding to a "figure" fire synchronously in time, whereas neurons responding to the background fire randomly. More recently von der Malsburg and Schneider (1986) have proposed a correlation theory to explain a temporal segregation of patterns.

Crick and Koch (1990) have considered the problem of visual awareness. They distinguish between two kinds of memory: "very short term" or "iconic," and a slower "short term" or working memory. Iconic memory appears to involve visual primitives, such as orientation and movement, and it appears to last half a second or less. On the other hand, short term or working memory lasts for a few seconds, and it deals with more abstract representations. This memory also has a limited capacity and it has been claimed that this capacity is about seven items.

For the "short term" memory Crick and Koch postulate an attentional mechanism that transiently binds together all those neurons whose activity relates to the different features of the visual object. They argue that semi-synchronous coherent oscillations in the 40–70 Hz range are an expression of this mechanism. These oscillations are sensory-evoked in response to auditory or visual stimuli. But postulating such a mechanism merely trade one problem for another: how the neurons that participate in these oscillations get bound is still not explained. This theory is an extension of the earlier "searchlight" hypothesis relating to awareness. It shifts the basis of awareness to a more abstract mechanism of attention.

This work has led to a re-examination of the question of what may be taken to define the loss of consciousness. In the standard view, anesthesia leads to four distinct effects: motionlessness in the face of surgery, attenuation or abolition of the autonomic responses like tachycardia, hypertension, and so on that would normally accompany surgery, lack of pain, and lack of recall. Use of EEG's shows little difference between natural sleep and the anesthetized state. On the other hand, Kulli and Koch (1991) argue that the loss of consciousness, as defined by these four effects, is best represented by the loss of the 40 Hz sensory-evoked oscillations. Deepening anesthesia has also been seen as progressive loss of complexity of the EEG phase plot (Hameroff 1987). Hameroff claims that anesthetics retard mobility of electrons, and doing so may disrupt hydrophobic links among proteins which interconnect microtubules which are filamentous substructures of the neuron (See next section). From a functional point of view one may see a prevention of memory consolidation from input to long-term memory storage as one of the consequences of anesthesia; other cognitive and motor functions may be similarly inhibited.

4.4.1 Scripts

Yarbus (1967) recorded the eye movements when a picture is viewed. He found that there is a continual scanning of the scene with a predominance of the fixation points on parts which carry important and complex features. From this one may infer that the brain constructs the regular image from fragments, obtained through the scanning process, that may be termed scripts. Such scripts may be defined both with respect to space as well as time.

The structuring of events in dreams gives us important clues regarding time scripts. My father explained to me in 1955 that certain dreams run as scripts. Thus a foot slipping off another in sleep may be accompanied by a dream about falling off a precipice. That such a script includes an appropriate sequence of events preceding the climax indicates that the mind rearranges the events so that the dream appears to have started before the slipping of the foot. Similar scripts should be a part of normal awake cognition. For example, in delirious speech ideas and words are strung together in an illogical manner. It is as if different brief scripts have been jumbled together without comparing the stream to the inner model of reality.

Scripts might be seen as chunks of visualization or linguistic behavior, just as neural subsystems might be viewed as chunks of structure.

4.4.2 Time effects

Experiments focusing on time delays and rearrangement of events by consciousness have been performed by H.H. Kornhuber and his associates and by Benjamin Libet. Kornhuber and his associates (Deeke et al 1976) found that the readiness potential, the averaged EEG trace from the precentral and parietal cortex, of a subject who was asked to flex his finger built up gradually for a second to a second and a half before the finger was actually flexed. This slow build-up of the readiness potential may be viewed as a response of the unconscious mind before it it passes into the conscious mind and is expressed as a voluntary movement. In Libet's work (1976) the subjects were undergoing brain surgery for some reason unconnected to the experiment and they agreed to electrodes being placed at points in the brain, in the somato-sensory cortex. It was found that when a stimulus was applied to the skin, the subject became aware of this half a second later although the brain would have received the signal of the stimulus in barely a hundredth of a second. Furthermore, the subjects themselves believed that no delay had taken place in their becoming aware of the stimulus!

In further experimentation the somatosensory cortex was electrically stimulated within half a second of skin stimulus. A backward masking phenomenon occurred and the subject did not become aware of the earlier skin sensation. Now Libet initiated a persistent cortical stimulation first and then within half a second he also touched the skin. The subject now was aware of both the stimulations but he believed that the skin stimulation preceded that of the cortex. In other words, this established that the subject did extrapolate the skin-touching sensation backwards in time by about half a second. These experiments also demonstrate how brain works as an active agent, reorganizing itself as well as the information.

4.5 Cytoskeletal Networks

Recursive definition is one of the fascinating characteristics of life. For example, natural selection does not work only at the level of species but also at the level of the individual and that of the nerve cells of the developing organism (Cowan 1981). Purposive behavior characterizes human societies as also the societies of other animals such as ants. Recursion may be seen regarding information processing as well down from animal societies to the neural structures of the individual or perhaps further down to the cytoskeleton of the cell. C.S. Sherrington (1951) argued that even single cells possess what might be called minds: "Many forms of motile single cells lead their own independent lives. They swim and crawl, they secure food, they conjugate, they multiply. The observer at once says 'they are alive'; the amoeba, paramaecium, vorticella, and so on. They have specialized parts for movement, hair-like, whip-like, spiral, and spring-like... [Of] sense organs ... and nerves there is no trace. But the cell framework, the cyto-skeleton, might serve."

Hameroff and his associates have argued that microtubules, hollow cylinders 25 nm across, that are the cytoskeletal filamentous polymers to be found in most cells, perform information processing. Hameroff (1987) and Rasmussen et al (1990) propose that the cytoskeleton may be viewed as the cell's nervous system and that it may be involved in molecular level information processing that subserves higher, collective neuronal functions ultimately relating to cognition. They further propose that microtubule automata may have a function in the guidance and movement of cells and cell processes during the morphogenesis of the brain's network architecture, and that in learning they could regulate synaptic plasticity.

In Hameroff (1994) interactions between the electric dipole field of water molecules confined within the hollow core of microtubules and the quantized electromagnetic radiation field were considered. These and Bose-Einstein condensates in hydrophobic pockets of microtubule subunits were taken to be responsible for microtubule quantum coherence. It was suggested that optical signalling in microtubules would be free from both thermal noise and loss, and that this may provide a basis for biomolecular cognition and a substrate for consciousness.

Irrespective of the precise function of the microtubule information transmission, it could carry additional features that may be useful in biological information processing. The notion of a connectionist network within the neuron, which in turn is an element of the connectionist neural network, defines a recursive relationship that has many desirable features from the point of view of ability to model biological behavior.

4.6 Invariants and Wholeness

The brain's task can also be viewed as one that involves the extraction of invariants. For example, an image cannot be recognized based on the visual stimulus that is compared to a stored exemplar. The reason is that the details of the visual stimulus depend on illumination, the distance and orientation of the object, motion and many other factors. In order to extract the invariant features of the image, the brain must, from the constantly changing cascade of photonic data, construct an internal visual world. Zeki (1992) has shown that color, form, motion and possibly other attributes are processed separately in specialized parts of the visual cortex. These parts function as parallel processing modules, although there exist considerable connections amongst them. How the processing of these parallel modules is put together remains a puzzle. Zeki summarizes: "The entire network of connections within the visual cortex must function healthily for the brain to gain complete knowledge of the external world. Yet as patients with blindsight have shown, knowledge cannot be acquired without consciousness, which seems to be a crucial feature of a properly functioning visual apparatus. Consequently, no one will be able to understand the visual brain in any profound sense without tackling the problem of consciousness as well."

It is the notion that awareness possesses a unity and that many aspects of consciousness are distributed over wide areas of the brain, that has driven the search for quantum neural models. These issues are summarized in Kak (1993c) and Penrose (1989). But the presence of noise and dissipation inside the brain makes the development of such models a daunting task.

Artificial neural network research has stressed pattern recognition and input-output maps. The development of machines that can match biological information processing requires much more than just pattern recognition. Corresponding to an input not only are many sub-patterns bound together, but there is also generated other relevant information defining the background and the context. This is the analog of the binding problem for biological neurons and therefore further progress in the design of artificial neural networks appears to depend on the advances in understanding brain behavior.

From the perspective of wholeness it appears that a drastic change of perspective may be necessary to solve the current problems. Consciousness is a recursive phenomenon: not only is the subject aware but he is also aware of this awareness. If one were to postulate a certain region inside the brain from where the searchlight is shown on the rest of the brain and which provides the unity and wholeness to the human experience, the question of what would happen if this searchlight were to be turned on itself arises.

Chapter 5

On Indivisible Phenomena

To achieve the marvellous information processing ability of animals it is natural to investigate the neural structure of the brain and relate it to a hypothesized nature of the mind. If brain structure is neuronal then cognitive capabilities should be found for networks of neurons. But note the argument (Sacks 1990) that neuropsychology itself is flawed since it does not take into account the notion of self, which is why it is hard put to explain phenomena such as that of phantom limbs (Melzack 1989).

The study of neural computers was inspired by possible parallels with the information processing of the brain. It was proposed that artificial neural networks constituted a new computing paradigm. But our experience with these networks has shown that such a characterization is incorrect. When simulated they represent sequential computing. In hardware, they may be viewed as a particular style of parallel processing but as we know parallel computing is not a departure from the basic Turing model. Neither does the use of continuous values provide us any real advantage because such continuous values can always be quantized and processed on a digital machine. Artificial neural networks were assumed to offer a real advantage in the solution of optimization problems. However, the energy minimization technique, while sound in theory, fails in practice since the network gets stuck in local minima.

Cognitive abilities may be seen to arise from a continuing reflection on the perceived world. This question of reflection is central to the brain-mind problem and the problem of determinism and free-will (see for example Kak 1986, Penrose 1989). A dualist hypothesis (for example Eccles 1986) to explain brain-mind interaction or the process of reflection meets with the criticism that this violates the conservation laws of physics. On the other hand a brain-mind identity hypothesis, with a mechanistic or electronic representation of the brain processes, does not explain how self-awareness could arise. At the level of ordinary perception there exists a duality and complementarity between an autonomous (and reflexive) brain and a mind with intentionality. The notion of self seems to hinge on an indivisibility akin to that found in quantum mechanics. This was argued most forcefully by Bohr, Heisenberg, and Schrödinger (e.g. Moore 1989).

5.1 Uncertainty

Uncertainty is a fundamental limitation in a quantum description arising out of indivisibility. Since a quantum system is characterized by the wavefunction ψ, which is a superposition of states, and the measurement leads to the collapse of the wavefunction, one can never completely infer the state before the measurement.

The wavefunction evolves according to the Schrödinger equation:

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi$$

where $\hat{H}$ is the Hamiltonian. According to the probability interpretation of the wavefunction, the probability of finding a quantum object in the volume element $d\tau$ is given by

$$\psi\psi^* d\tau$$

Thus $\psi\psi^*$ is a probability density.

For a particle the Heisenberg's uncertainty relation

$$\delta x \delta p \geq \frac{\hbar}{2}$$

limits the simultaneous measurements of the coordinate x and momentum p, where δx and δp are the uncertainties in the values of the coordinate and momentum. This relation also implies that one cannot speak of a trajectory of a particle. Since the initial values of x and p are inherently uncertain, so also is the future trajectory of the particle undetermined. One consequence of this picture is that we cannot have the concept of the velocity of a particle in the classical sense of the word.

There are many scholars (see Jammer 1974 for a bibliography) who champion a statistical interpretation of the uncertainty relations, according to which the product of the standard deviations of two canonically conjugate variables has a lower bound given by $\hbar/2$. Such an interpretation is based on the premise that quantum mechanics is a theory of ensembles rather than individual particles. But there is considerable evidence, including that related to the EPR experiment, described in the next chapter, that goes against the ensemble view. It appears, therefore, that the correct interpretation is the non-statistical, according to which one cannot simultaneously specify the precise values of conjugate variables that describe the behavior of a single object or system.

5.2 Complementarity

The principle of complementarity, as a commonly used approach to the study of the individuality of quantum phenomena, goes beyond wave-particle duality. In the words of Bohr:

> The crucial point [is] the *impossibility of any sharp separation between the behavior of atomic objects and the interaction with the measuring instruments which serve to define the conditions under which the phenomena appear.* In fact,

> the individuality of the typical quantum effects finds its proper expression in the circumstance that any attempt of subdividing the phenomena will demand a change in the experimental arrangement introducing new possibilities of interaction between objects and measuring instruments which in principle cannot be controlled. Consequently, evidence obtained under different experimental conditions cannot be comprehended within a single picture, but must be regarded as complementary in the sense that only the totality of the phenomena exhausts the possible information about the objects (Bohr 1951, page 39).

Observe that complementarity is required at different levels of description. But just as one might use a probabilistic interpretation instead of complementarity for atomic descriptions, a probabilistic description may also be used for cognitive behavior. However, such a probabilistic behavior is inadequate to describe the behavior of individual agents, just as notions of probability break down for individual objects.

According to complementarity, one can only speak of observations in relation to different experimental arrangements, and not an underlying reality. If such an underlying reality is sought then it is seen that the framework of quantum mechanics suffers from paradoxical characteristics. One of these is non-local correlations that appear in the manner of action at a distance (Bell 1987). But quantum mechanics remains a very successful theory in its predictive power.

Consider again the similarity between the thought process and the classical limit of the quantum theory. The logical process corresponds to the most general type of thought process as the classical limit corresponds to the most general quantum process. In the logical process, we deal with classifications. These classifications are conceived as being completely separate but related by the rules of logic, which may be regarded as the analogue of the causal laws of classical physics. In any thought process, the component ideas are not separate but flow steadily and indivisibly. An attempt to analyze them into separate parts destroys or changes their meanings. Yet there are certain types of concepts, among which are those involving the classification of objects, in which we can, without producing any essential changes, neglect the indivisible and incompletely controllable connection with other ideas.

5.3 What Is A Quantum Model?

Wave-particle duality is not a characteristic only of light. Interference experiments have been conducted with neutrons. In one experiment, the interference pattern formed by neutrons, diffracted along two paths by silicon crystal, could be altered by changing the orientation of the interferometer with respect to the earth's gravitational field. This demonstrated that the Schrödinger equation holds true under gravity.

Interference experiments are also being designed for whole atoms. If we use complementarity in its widest sense as applying to all "indivisible" phenomena, reality is seen to be such that it cannot be captured by a single view. This indicates that one should be able to define complementary variables in terms of higher level attributes for "complex", indivisible objects.

From another perspective, a quantum system is characterized by probabilistic behavior. In Born's probability interpretation, the wavefunction is not a material wave but rather a probability wave. Nevertheless, in relativistic quantum theory, the wavefunction cannot be used for the definition of a probability of a single particle.

The orthodox Copenhagen interpretation of quantum mechanics is based on the fundamental notion of uncertainty and that of complementarity. According to this interpretation we cannot speak of a reality without relating it to the nature of measurement.

If one seeks a unified picture one is compelled to accept the existence of effects propagating instantaneously. One example of paradoxical results arising from an insistence on a description independent of the nature of observation is the delayed choice variant of the double-slit experiment of the well-known physicist John Archibald Wheeler.

5.3.1 Delayed choice

In the delayed choice experiment considering a specific reality (picture) before observations leads to the inference that we can influence the past. In the words of Wheeler (1980):

> A choice made in the here and now has irretrievable consequences for what one has the right to say about what has *already* happened in the very earliest days of the universe, long before there was any life on earth.

Wheeler considers a source that emits light of extremely low intensity, one photon at a time, with a long time interval between the photons. The light is incident on a semi-transparent mirror M_1 and divides into two parts r and t and after reflections by the totally reflecting mirrors A and B it reaches the photomultipliers P_1 and P_2 (Figure 4). Since a photon will travel one of the two trajectories rAr or tBt, it will be detected either by P_1 or P_2. This experiment shows the corpuscular nature of photons.

Now a semi-transparent mirror M_2 is inserted at DCR (the delayed-choice region). The thickness of the mirror is so chosen that if light is considered wave-like then the superposition of the waves toward P_2 combines in a destructive manner and the waves toward P_1 interfere in a constructive manner. With the insertion of M_2 only the detector P_1 will register any reading and it should be assumed that each photon followed both trajectories.

Wheeler now wishes for us to imagine a situation where the mirror M_2 has been inserted at the last moment, when the photon has already passed through M_1 and is along the way either down rAr or tBt. If M_2 is inserted the photon will behave as a wave, as if it had followed both the paths. On the other hand, if M_2 is not inserted then the photon chooses only one of the two paths. Since, the insertion of M_2 is done at the last moment, it means that this choice modifies the past with regard to the behavior of that photon.

Wheeler points out that astronomers could perform such a delayed choice experiment on light from quasars that has passed through a galaxy or other massive object that acts as a gravitational lens. Such a lens can split the light from the quasar and refocus it in the direction of the distant observer.

The astronomer's choice of how to observe photons from the quasar appears to determine whether each photon took both paths or just one path around the gravitational lens in its journey commenced billions of years ago. When they approached the galactic beam-splitter, did the photons make a choice that would satisfy the conditions of an experiment to be performed by unborn beings on a still nonexistent planet? Clearly, this fallacy arises out of the view that a photon has a physical form before the astronomer's observation.

Chapter 6

Of Quantum And Neural Computation

6.1 Parallels With A Neural System

By a quantum neural computer we mean a (strongly) connectionist network which is nevertheless characterized by a wavefunction. Let us assume that we are only interested in considering the computational potential of such a supposition. In parallel to the operational workings of a quantum model we can sketch the basic elements of the working of such a computer. Such a computer will start out with a wavefunction, reflecting the state of the self-organization of the connectionist structure, that is a sum of several different problem functions. After the evolution of the wavefunction the measurement operator will force the wavefunction to reduce to the correct eigenfunction with the corresponding measurement that represents the computation.

States of quantum systems are associated with unit vectors in an abstract vector space V, and observables are associated with self-adjoint linear operators on V. Consider the self-adjoint operator $\hat{f}$. We can write

$$\hat{f}\psi_i \;=\; f_i\psi_i$$

where ψ_i are the eigenvectors and f_i are the eigenvalues. The ψ_i are the wavefunctions and f_i represent the corresponding measurements. When ψ_i are a complete orthonormal set of vectors in V, we have $\psi_i\psi_j^* = \delta_{ij}$. Also any wavefunction ψ in V can be written as a linear combination of the ψ_i. Thus

$$\psi \;=\; \sum_i c_i\psi_i$$

where the complex coefficients are given by

$$c_i \;=\; \int \psi\psi_i^* dq$$

and q represents the configuration space. The sum of the probabilities of all possible values f_i equals unity:

$$\sum |c_i|^2 = 1$$

When an observation is made on an object characterized by a wavefunction ψ, the measurement process causes the wavefunction to collapse to the eigenfunction ψ_i that corresponds to the measurement f_i. This measurement itself is obtained with the probability $|c_i|^2$.

Since the wavefunction evolves irrespective of whether any observations are being made, this will endow the system to compute several problems simultaneously. Analogously, animals can simultaneously perform several tasks, although such performance has traditionally been explained as being reflexive.

From a functional point of view this has parallels with the workings of a feedback neural computer. In such a computer the final measurement is one of the stored "eigenvectors" X_i where the neural computer is itself characterized by the synaptic interconnection weight matrix T, so that

$$sgm\{TX_i\} = X_i$$

where sgm is a nonlinear sigmoid function, so as to define a X with discrete component values.

One difference between the quantum mechanical framework and the above equation is the non-linearity introduced by the sigmoidal function. But this non-linearity may be seen to be at the end of a chain of transformations, where the first step is a linear transformation as in quantum mechanics. On the other hand, the matrix T is like one of the many measurement operators of a quantum system. In other words, a neural network is associated with a single measurement operator, whereas a quantum computer is associated with a variety of measurement operators.

Equivalently, we may view a quantum computer to be a collection of many neural computers that are "bound" together into a unity. The view of the cognitive process representing self-organized neural activity implies that each such process does set up what may be considered a different neural computer. The set of the self-organized states may then be viewed as representing a collection of neural structures that are bound together. The search for a local mechanism for this binding is as difficult as for the binding problem mentioned in Chapter 4. If on the other hand, we postulate a global function defining such a binding, then we speak of a quantum neural computer.

If the wavefunction was associated with several operators then the neural hardware for a specific problem will secure its solution. The measurement process will appear instantaneous after the decision to choose a specific measurement has been made. But how this choice might be made constitutes another problem.

Some of the characteristics of such a model are:

1. It explains intuition, or the spontaneous computation of the kind performed in a creative moment, as has been reported by Poincaré, Hadamard, and Penrose (1989).

2. A wavefunction that is a sum of several component functions explains why the free-running mind is a succession of unconnected images or episodes. The classical neural model does not explain this behavior.

3. One can admit the possibility of tunnelling through potential barriers. Such a computer can then compute global minima, which cannot be done by classical neural computer, unless by the artifice of simulated annealing. (The simulated annealing algorithms may not converge.)

4. Being a linear sum of a large (or infinite) number of terms, the individual can shift the focus to any desired context by the application of an appropriate measurement hardware that has been designed through previous exposure (reinforcement) or through inheritance. Such a shifting focus is necessary in speech or image understanding.

6.2 Measurements

Let us consider a pure state ψ to be a superposition of the eigenfuctions. Or

$$\psi = \sum_i c_i \psi_i$$

The overlapping realities collapse into one of the alternative worlds of a specific eigenfunction when a measurement by a non-quantum device is made. However, not all measurement devices are completely non-quantum; thus superconductivity represents a quantum phenomenon at the macroscopic level. If the measuring system is also a quantum system, it too should be described by a wavefunction. The measurement state would then arise out of the interference of the two wavefunctions and, in itself, it would represent a superposition of several states. As shown by von Neumann this requires that further measurements be made on the measuring apparatus to resolve the wavefunction and so on.

6.2.1 The example of vision

It has been argued by Gibson (1979) and Schumacher (1986) that it is not necessary to view vision as resulting from the passage of signals through the optic nerve, but rather as a reorganization of the brain as a response to an environment. According to Gibson:

> It is not necessary to assume that *anything whatever* is transmitted along the optic nerve in the activity of perception. We can think of vision as a perceptual system, the brain being simply part of the system. The eye is also part of the system, since retinal inputs lead to ocular adjustments and then to altered retinal inputs, and so on. The process is circular, not a one-way transmission. The eye-head-brain-body system registers the invariants in the structure of ambient light. The eye is not a camera that forms and delivers an image, nor is the retina simply a keyboard that can be struck by fingers of light.

Consider the parallel between the tracks in the photosensitive emulsion used to detect deflected particles and the eye. Bohm (1980) has argued that the track in the emulsion is to be viewed as resolving the relevant wavefunction; likewise, the track in the eye may be viewed as resolving the wavefunction associated with the incident beam. But we have already seen in Chapter 1.1 that the eye cannot be seen to be entirely classical. It appears then that the vision system itself should be decomposed into a concatenation of subsystems where this resolution proceeds.

From this perspective the vision system may be viewed as quantum and classical measuring instruments associated with a quantum process.

6.3 More On Information

That a quantum neural computer will have characteristics different from that of a collection of classical computers is seen from an examination of information. In contrast to classical systems, information in a quantum mechanical situation is affected by the process of measurement. If a linearly polarized photon strikes a polarizer oriented at 90° from its plane of polarization, the probability is zero that the photon will pass to the other side. If another polarizer tilted at 45° is placed before the first one, there is a 25 percent chance that a photon will pass both polarizers. By the process of measurement the 45° polarizer transforms the photons so that half of the initial photons can pass through it. The collapse of the wavefunction also causes non-local effects. These characteristics can be looked for in determining whether biological information systems should be considered quantum.

6.3.1 The EPR experiment

We consider the thought experiment described by Einstein, Podolsky, and Rosen (1935) (Bell 1987) now known as the EPR experiment. EPR assume the following condition for an element of a mathematical model to represent an element of reality:

> If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical reality. (page 777)

They also assert that "every element of the physical reality must have a counterpart in the physical theory." They argue that if a pair of particles has strongly interacted in the past then, after they have separated, a measurement on one will yield the corresponding value for the other. Since the position and momentum values are supposed to be undefined before the measurement is made and, nevertheless, the value is revealed after the measurement on the remote particle is made, EPR argue that the measurements should correspond to aspects of reality. They conclude that "the quantum-mechanical description of physical reality given by wave functions is not complete."

Since after separation each particle is to be considered as physically independent of the other (locality condition), this can also be taken to imply that effects propagate instantaneously. Bell (1987) has shown that the EPR reality criterion is incompatible with the predictions of quantum mechanics. Experiments have confirmed the predictions of the latter.

We now consider the EPR experiment from an information theoretic viewpoint. In its Bohm variant, a pair of spin one-half particles, A and B, have formed somehow in the singlet spin state and they are moving freely in opposite directions. The wavefunction of this pair may be represented by

$$\frac{1}{\sqrt{2}}\,(V^+W^- \;-\; V^-W^+)$$

where V^+ and V^- represent the measurements of spin $+1/2$ and $-1/2$ for particle A and W^+ and W^- represent the measurement of $+1/2$ and $-1/2$ for particle B.

The EPR argument considers the particles A and B to have separated. Now if a measurement is made on A along a certain direction, it guarantees that a measurement made on B along the same direction will give the opposite spin. The important point here is that the spin is determined as soon as, but not before, one of the particles is measured. This has been interpreted to mean that knowledge about A somehow reduces the wavefunction for B in a specific sense. In other words, the EPR correlation has been taken to imply a non-local character for quantum mechanics, or instantaneous action at a distance. To reexamine these questions in an information-theoretic perspective, it is essential to determine the extent of information obtained in each measurement.

Given a spin one-half particle, an observation on it produces $\log_2 2 \;=\; 1$ bit of information. A further measurement made along a direction at an angle of θ to the previous measurement leads to a probability $cos^2\theta/2$ that the measurement would give the same sign of spin, and a probability $\sin^2\theta/2$ that it will give spin of the opposite sign. The information associated with the second measurement is

$$H(\theta) \;=\; -\cos^2\frac{\theta}{2}\;\log_2\cos^2\frac{\theta}{2} \;-\; \sin^2\frac{\theta}{2}\;\log_2\sin^2\frac{\theta}{2}$$

The average information considering all angles is:

$$H_{av} \;=\; \frac{1}{\pi}\int_0^{\pi} H(\theta)d\theta = \;1 \;-\; \log_2\sqrt{e}\ \text{bits} \;=\; 0.27865\ \text{bits}.$$

In other words, the information obtained from the second measurement is somewhat less than a quarter of a bit. These sequential experiments are correlated. It is also important to consider that all further measurements provide exactly 0.27865 bits of information. This indicates that information in a quantum description is not a locally additive variable.

6.3.2 Information in the experimental setup

Now consider the information to be obtained from the measurement of spin of two half-spin particles that are correlated in the EPR sense. If information were additive then the first measurement provides one bit of information and after the end of the second measurement we have a total of 1.27865 *bits*.

Since the EPR correlation reveals the spin of the particle B, as soon as the measurement of A has been made, one might infer that the information that the two arms of the experimental setup provide equals 0.27865 bits. But in reality we cannot do so as the calculus for information, in a quantum description, is unknown. Not being locally additive, the information in the experimental setup cannot be used in subsequent repetitions of the experiment.

6.4 A Generalized Correspondence Principle

In a study several years ago (Kak 1976, 1984) it was argued that the fundamental Heisenberg uncertainty was compensated by information in terms of new symmetries of the quantum description. In other words one can generalize the correspondence principle to include uncertainty:

$$I_{classical} = I_{qm} + uncertainty$$

A calculation of this information yields plausible results regarding the number of such symmetries. The analysis given here provides further elaboration of that idea. Although we appear to be unable to use the information in the various measurements, it is clear that the experimental setup itself plays a fundamental role in our knowledge. From another perspective, information can be associated with the angular (as well as spatial) position of an object. Nevertheless, this information cannot be considered in a local fashion. Or in other words, this information cannot be considered separately for the components of the system.

In the context of a quantum neural computer it means that behavioral characteristics of such a machine can only be defined in terms of manner in which measurements on it are made. If such measurements are stored in the machine, then this experience can only be regarded as an *interaction* between the machine and the environment. Berger (1977, p. xiv) argues that one needs to use a similar relativism in the study of biological systems. "A necessary condition of experience is interaction between the organism and the environment, so that the structure of experience reflects the structure of both the organism and the environment with which it interacts. It follows that an organism cannot experience its own structure or that of the environment independently of the other."

Assume that an emergent phenomenon arises from interactions that are noncomputable in terms of the descriptions at the lower level. Assume further that the higher level description is associated with a fundamental uncertainty. Will the generalized correspondence principle apply to this situation? And if it does, will the emergent phenomenon be associated with new attributes that provide information equal in magnitude to this uncertainty?

If information in biological information processing exhibits other non-local characteristics as well, such behavior may represent further parallels between the quantum mechanical paradigm and biological reality.

Chapter 7

Structure And Information

7.1 A Subneuron Field

In our earlier chapters we have critiqued many materialist models, where an identity is assumed between mental events and neural events in the higher centers of the brain. Our main criticism was that these models did not deal with the binding problem. We presented evidence in support of the view that biological systems perform quantum computing.

We now raise the question of the location of the quantum field that associates experience with its unity. Two proposals are reviewed in this section.

7.1.1 A dualist hypothesis

Popper and Eccles (1977) presented the view in which the world of mental events (they call it World 2) is sharply separated from the world of brain processes (World 1). Both these worlds are supposed to have autonomous existence. These ideas were further developed in Eccles (1986, 1990).

In his theory, Eccles (1990) assumes that the entire set of the inner and outer senses is composed of elementary mental events, that he calls psychons. He further proposes that each psychon is linked to a corresponding functional aggregate of dendrites, that he names a dendron. A dendron is the basic receptive unit of the cerebral cortex, and it is a bundle or cluster of apical dendrites of certain pyramidal cells. Each dendron could involve about 200 neurons and as many as 100,000 synapses. There are about 40 million dendrites in the human cerebral cortex. He assumes that the mind exerts a minimal additional excitation on the dendronic vesicles, which are at a state just below the threshold for excitation.

In order to meet the objection that immaterial mental events such as thoughts cannot act on material structures, such as the neurons in the cerebral cortex, Eccles has recourse to quantum mechanics. He says that "a calculation on the basis of the Heisenberg uncertainty principle shows that a vesicle of the presynaptic vesicular grid could conceivably be selected for exocytosis by a psychon acting analogously to a quantal probability field. The energy required to initiate the exocytosis by a particle displacement could be paid back at the same time and place by the escaping transmitter molecules from a high to a low concentration. In quantum physics at microsites energy can be borrowed provided it is paid back at once. So the transaction of exocytosis need involve no violation of the conservation laws of physics." (Eccles 1990, page 447)

Although the central idea of this theory is the linkage between the psychons and the microsite dendrons, Eccles' theory is a field theory of psychons. A dendron represents the measurement hardware that each unitary or elementary mental event is associated with. Psychons are autonomous entities that may exist apart from dendrons and be related amongst themselves. In the brain to mind transaction, Eccles proposes that each time a psychon selects a vesicle for exocytosis, this "micro-success" is registered in the psychon for transmission through the mental world. This signal carries into the mental world the special experiential character of that psychon.

Eccles has only sketched the outline of his theory. He sees the psychons and the brain processes to be interlinked entities, which is why he labels his theory to be dualistic. But what is the physical support of psychons, he does not explain. Seen from this perspective, psychons are nothing but an artifice to circumvent the brain-mind problem. If they need the dendrons to emerge as mental events, then we are again confronted with the problem of the processes that lead to such an emergent phenomenon.

Eccles does see the need for a quantum theoretic basis to explain psychon-dendron interaction of mental events. It appears, therefore, that the psychon field ought to be a quantum field. He does not explain how the world of mental events, as a separate, autonomous reality, functions?

7.1.2 Microtubule field

As a field one cannot take awareness to be localized. Considering the parallel of the wavefunction, it is a unity and it cannot be taken to be identical to the physical body although it is associated with it. Hameroff and his collaborators (Hameroff 1994, Jibu et al 1994) have suggested that the sub-neural structures called microtubules provide coherent information transmission that leads to the development of a quantized field that solves the "binding problem" and explains the unitary sense of self. These cytoskeletal microtubules, which provide structural support to cells, are hollow cylinders 25 nanometers in diameter whose walls are made of subunit proteins known as tubulin. Each tubulin subunit is a 8-nm dimer that is made up of two slightly different classes of 4-nm dimers known as α and β tubulin. Hameroff has argued that the quantum dynamical system of water molecules and the quantized electromagnetic field confined inside the hollow microtubule core manifests a collective dynamics by which coherent photons are created inside the microtubule. But the question of why consciousness characterizes only certain cells, although microtubules are to be found in all cells, is not clearly answered. Hameroff (1994) suggests that consciousness might be an attribute of all quantum phenomena. Commenting on the results of the EPR experiment, he says that "quantum entities are 'aware' of the states of their spatially separated relatives!"

It is possible also to see the microtubule information as providing the "imaginary" component in the complex neural information that was described in an earlier chapter. If this is the mechanism that provides the simultaneous global and local information which is essential for artificial intelligence tasks then it is clearly not feasible to build such quantum neural computers at this time because microtubule information processing is still imperfectly understood.

7.1.3 A universal field

If one did not wish for a reductionist explanation as in inherent in the cytoskeletal model, one might postulate a different origin for the quantum field. Just as the unified theories explain the emergence of electromagnetic and weak forces from a mechanism of symmetry breaking, one might postulate a unified field of consciousness-unified_force-gravity from where the individual fields emerge.

The notion of a universal field still requires for one to admit the emergence of the individual's I-ness at specialized areas of the brain. This I-ness is intimately related to memories, both short-term and long-term. The recall of these memories may be seen to result from operations by neural networks. Lesions to different brain centers effect the ability to recall or store memories. For example, lesions to the area V1 of the primary visual cortex lead to blindsight (Weiskrantz 1986). These people can "see" but they are unaware that they have seen. Although such visual information is processed, and it can be recalled through a guessing game protocol, it is not passed to the conscious self.

7.2 Uncertainty Relations

By structure we mean a stable organization of the neural system. The notion of the stability may be understood from the perspective of energy of the neural system. Each stable state is an energy minimum. For example, the following energy expression (Hopfield 1982) is suitable for a feedback neural network:

$$E = -\frac{1}{2}\sum_{ij} T_{ij} X_i X_j$$

Any structure may be represented by a number, or a binary sequence. Thus in a one dimension, the sequences

$$abc, ba, cab$$

represent three structures that can be coded into numbers by a binary code.

Assume that a neural structure has been represented by a sequence. Since this representation can be done in a variety of ways, the question of a unique representation becomes relevant.

Definition 1 *Let the shortest binary program that generates the sequence representing the structure be called p.*

The idea of the shortest program gives us a measure for the structure that is independent of the coding scheme used for the representation. The length of this program may be taken to be a measure of the information to be associated with the organization of the system. This length will depend on the class of sequences that are being generated by the program. Or in other words, this reflects the properties of the class of structures being considered. Evidence from biology requires that the brain be viewed as an active system which reorganizes itself in response to external stimuli. This means that the structure p is a variable with respect to time.

Assuming, by generalized complementarity, that the structure itself is not defined prior to measurement, then for each state of an energy value E, we may, in analogy with the Heisenberg's uncertainty principle, say that

$$\delta E \delta t \geq k_1$$

where k_1 is a constant based on the nature of the organizational principle of the neural system.

The external environment changes when the neural system is observed, due to the interference of the observer. This means that as the measurement is made, the structure of the system changes.

This also means that at such a fundamental level, a system cannot be associated with a single structure, but rather with a superposition of several structures. Might this be a reason behind pleomorphism, the multiplicity of forms of microbes?

The representation described above may also be employed for the external environment.

Definition 2 *Let the shortest binary program that generates the external environment be called x.*

If the external environment is a eigenstate of the system, then the system organization will not change; otherwise, it will.

We may now propose an uncertainty principle for neural system structure:

$$\delta x \delta p \geq k_2$$

This relation says that the environment and the structure cannot be simultaneously fixed. If one of the variables is precisely defined the other becomes uncontrollably large. Either of these two conditions implies the death of the system. In other words, such a system will operate only within a narrow range of values of the environment and structure.

We conjecture that $k_1 = k_2 = k$.

One may pose the following questions:

- Are all living systems characterized by the same value of k?
- Can one devise stable self-organizing systems that are characterized by a different value of k? Would artificial life have a value of k different from that of natural life?
- What is the minimum energy required to change the value of p by one unit?
- Does a Schrödinger type equation define the evolution of structure?

It is also clear that before a measurement is made, one cannot speak of a definite state of the machine, nor of a definite state of the environment.

Chapter 8

Concluding Remarks

In the previous sections we have argued that humans and other animals perform what might be called quantum neural computing. If we accept the reverse of this claim then a quantum neural computer would be characterized with life and consequently consciousness. On the other hand, ordinary machines cannot be conscious since they do not come with a set of potentialities that consciousness provides. Machines are therefore like the neural hardware that provides extension. A machine that is so designed so that it has infinite set of potentialities would be alive. But such an alive machine need not be based on the organic molecules of normal life.

Our review has highlighted the following points:

- Conventional computing has been unable to devise schemes for holistic processing. This is why computers cannot recognize faces or understand text. Conventional computing is based on a reductionist algorithmic approach. But such an approach presupposes that the particles, or objects of the computation have been *selected*. The selection of these objects is normally left to a human in the solution of any real world problem. Conventional computing has failed at deriving methods for such a selection.

- Biological computing may be viewed as a potentially infinite collection of conventional computing devices. This versatility arises from the brain reorganizing its structure in response to an association, presented by the environment or by a self-generated state, to become a special-purpose computing machine suited to the solution of that problem. We postulate a wavefunction associated with brain behavior that allows the biological system to get into a definite structural state corresponding to the stimulus.

- We can simulate a quantum neural computer by a collection of a large number of conventional computers. To build a true quantum neural computer one will have to develop a framework for organization. In other words, one would then know how the different conventional computers that together constitute an approximation to the quantum neural computer could have a *plasticity* built into them so that the same machine would be able to reorganize itself.

- We speculate that self-awareness is associated with the wavefunction that allows for the selectivity in the biological organism. If this is true then a quantum neural computer will be self-aware.

This article presents arguments for a holistic computing paradigm that has parallels with biological information processing and with quantum computing. In this paradigm inputs trigger an internal reorganization of the connectionist computer that makes it selective to the input. But such a holistic paradigm suffers from several paradoxical aspects. No wonder, the determination of the phenomenological correlates of the holistic function, and therefore the design of a computer that operates in this paradigm, remain perplexing problems.

A quantum neural computer represents an underlying quantum system that interacts with classical measurement structures composed of neural networks. This parallels the basic quantum theory framework where measurement is to be performed using macroscopic apparatus. Learning by a biological system then represents the development of the measuring instruments, which are the neural structures in the brain. This picture still needs to address other questions, such as how in response to a stimulus does the brain reorganize itseld so as to pick the appropriate neural network for measurement? And how does the unity of the wavefunction lead to self-awareness? Unless we consider all matter to be conscious, we must take consciousness to be an emergent property of a quantum system that requires a certain structural basis supporting specific neuronal activity.

References

Aurobindo, S. (1970). "The Synthesis of Yoga." Aurobindo Ashram, Pondicherry.

Baylor, D.A., Lamb, T.D., and Yau, K.-W. (1979). *J. Physiol.* 288, 613-634.

Bell, J.S. (1987). "Speakable and Unspeakable in Quantum Mechanics". Cambridge University Press, Cambridge.

Beniof, P. (1982). *J. Stat. Phys.* 29, 515-546.

Berger, R. (1977). "Psyclosis: The Circularity of Experience". W.H. Freeman and Company, San Francisco.

Bohm, D. (1951). "Quantum Theory." Prentice-Hall, New York.

Bohm, D. (1980). "Wholeness and the Implicate Order." Routledge and Kegan Paul, London.

Bohr, N. (1951). "Atomic Physics and Human Knowledge." Science Editions, New York.

Cowan, W.M. (ed.), (1981). "Studies In Developmental Neurobiology." Oxford University Press, New York.

Crick, F. and Koch, C., (1990). *Seminars in the Neurosciences*, 2, 263-275.

Das, S. (1994). "Second Order Neural Networks." Ph.D. dissertation, LSU.

Deeke, L., Grötzinger, B., and Kornhuber, H.H., (1976). *Biological Cybernetics* 23, 99-119.

Deutsch, D. (1985). *Proc. Royal Society London*, A 400, 97-117.

Deutsch, D. (1989). *Proc. Royal Society London*, A 425, 73-90 (1989).

Dyczkowski, M.S.G. (1987). "The Doctrine of Vibration". State University of New York Press, Albany.

Eccles, J.C. (1986). *Proc. Royal Society London,* B 227, 411-428.

Eccles, J.C. (1990). *Proc. Royal Society London,* B 240, 433-451.

Einstein, A., Podolsky, B., Rosen, N. (1935). *Phys. Rev.* 47, 777-780.

Feigenbaum, M.J. (1978). *J. Statistical Physics* 19, 25-52.

Feynman, R.P. (1986). *Found. Phys.* 16, 507-531.

Freeman, W.A. (1992). *International J. of Bifurcation and Chaos*, 2, 451-482.

Freeman, W.A. (1993). *In* "Rethinking Neural Networks: Quantum Fields and Biological Data" (Karl H. Pribram ed.). Lawrence Erlbaum Associates.

Freeman, W.A. and S. Jakubith, S. (1993). *In* "Brain theory" (A. Aertsen, ed.). Elsevier Science Publishers, New York.

Gibson, J.J. (1979). "The Ecological Approach to Visual Perception." Houghton and Mifflin, Boston.

Gramss, T. (1994). On the speed of quantum computation. *Technical Report* 94-04-017, Santa Fe Institute.

Griffin, D.R. (1992). "Animal Minds." The University of Chicago Press, Chicago.

Hamburger, V. and Levi-Montalcini, R. (1949). *J. Exp. Zool.* 111, 457-502.

Hameroff, S.R. (1987). " Ultimate Computing: Biomolecular Consciousness and NanoTechnology". North-Holland, Amsterdam.

Hameroff, S.R. (1994). *J. of Consciousness Studies* 1, 91-118.

Hameroff, S.R., Dayhoff, J.E., Lahoz-Beltra, R., Samsonovich, A.V., Rasmussen, S. (1992). *IEEE Computer* 25(11), 33-45.

Herrnstein, R.J., Loveland, D.H. (1964). *Science* 146, 549-551.

Herrnstein, R.J. (1985). *Phil. Trans. R. Soc. Lond.* B 308, 129-144.

Hopfield, J.J. (1982). *Proc. Natl. Acad. Sc. USA* 81, 2554-2558.

Jammer, M. (1974). "The Philosophy of Quantum Mechanics". John Wiley, New York.

Jibu, M. Hagan, S., Hameroff, S.R., Pribram, K.H., Yasue, K. (1994). *BioSystems*, 30, in press.

Kak, S.C. (1976). *Nuovo Cimento*, 33B, 530-534.

Kak, S.C. (1984). *Proceedings of the Indian National Science Academy*, 50, 386-396.

Kak, S.C. (1986). "The Nature of Physical Reality". Peter Lang, New York.

Kak, S.C. (1988). *Proc. of the 22nd Annual Conference on Information Science and Systems.* Princeton, pp. 131-133.

Kak, S.C. (1993a). *Pramana -J. of Physics*, 40, 35-42.

Kak, S.C. (1993b). *Circuits, Systems, and Signal Processing*, 12, 263-278.

Kak, S.C. (1993c). "Reflections in clouded mirrors: selfhood in animals and machines". Symposium on *Aliens, Apes, and AI: Who is a person in the postmodern world*, Southern Humanities Council Annual Conference, Huntsville, AL, February 1993.

Kak, S.C. (1994a). *Pattern Recognition Letters*, 15, 295-298.

Kak, S.C. (1994b). "From Vedic science to Vedanta." *Fifth International Congress of Vedanta*, Miami University, Oxford, Ohio, August 1994.

Kak, S.C. (1994c). "The Astronomical Code of the Ṛgveda." Aditya, New Delhi.

Kak, S.C. (1995). *Information Sciences*, 83, 143-160.

Kandel, E.R. and Hawkins, R.D. (1992). *Scientific American*, 267(3), 79-86.

Koch, C. (1993). *Current Opinion in Neurobiology* 3, 203-208.

Kulli, J. and Koch, C. (1991). *Trends in Neurosciences* 14, 6-10.

Landauer, R. (1991). *Physics Today* 44, 23-29.

Libet, B. (1986). *Federation Proceedings, American Societies for Experimental Biology*, 45, 2678-2686.

Libet, B., Wright, E.W., Feinstein, B. and Pearl, D.K. (1979). *Brain* 102, 193-224.

Madineni, K.B. (1994). *Information Sciences* 81, 229-234.

von der Malsburg, C. and Schneider, W. (1986). *Biological Cybernetics* 54, 29-40.

Melzack, R. (1989). *Canadian Psychology* 30, 1-16.

Milner, P.M. (1974). *Psychological Review* 81, 521-535.

Milton, J., Mundel, T., an der Heiden, U., Spire, J.-P., and Cowan, J. (1994). Neural activity waves and the relative refractory states of neurons. *Technical Report* 94-09-052, Santa Fe Institute.

Moore, W. (1989). "Schrödinger: Life and Thought". Cambridge University Press, Cambridge.

Mountcastle, V.B. (1989). *In* "The Mindful Brain" (G.M. Edelman and V.B. Mountcastle, eds.), pp. 7-50. MIT Press, Cambridge.

Oppenheim, R.W. (1981). *In* "Studies in Developmental Neurobiology." (W.M. Cowan, ed.) Oxford University Press, New York.

Optican, L.M. and Richmond, B.J. (1987). *J. Neurophys.* 57, 162-178.

Penrose, R. (1989). "The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics." Oxford University Press, Oxford.

Pepperberg, I.M. (1983). *Anim. Learn. Behav.* 11, 179-185.

Popper, K.R. and Eccles, J. C. (1977). "The Self and its Brain." Springer International, Berlin.

Purves, D.. (1988). "Body and Brain: A Trophic Theory of Neural connections." Harvard University Press, Cambridge.

Raina, P. (1994). *Information Sciences* 81, 261-274.

Rasmussen, S., Karampurwala, H., Vaidyanath, R., Jensen, K.S., and Hameroff, S. (1990). *Physica*, D 42, 428-449.

Sacks, O. (1990). "Awakenings, A Leg to Stand On, The Man Who Mistook His Wife for a Hat". Book of the Month Club, New York.

Savage-Rumbaugh, E.S., Sevcik, R.A., Rumbaugh, D.M., and Rubert, E. (1985). *Phil. Trans. R. Soc. Lond.* B 308, 177-185.

Savage-Rumbaugh, E.S. (1986). "Ape Language: From conditioned response to symbol." Columbia University Press, New York.

Schnapf, J.L. and Baylor, D.A. (1987). *Scientific American*, 256(4), 40-47.

Schommers, W. (1989), *Quantum Theory and Pictures of Reality.* Springer-Verlag, Berlin.

Schrödinger, E. (1961). "Meine Weltansicht." Paul Zsolnay, Wien.

Schrödinger, E. (1965). "What is Life?" Macmillan, New York.

Schumacher, J.A. (1986). *In* "Fundamental Questions in Quantum Mechanics." (L.M. Roth and A. Inomata, eds.), pp. 93-109. Gordon and Breach, New York.

Sherrington, C.S. (1951). "Man On His Nature". Cambridge University Press, Cambridge.

Sinha, J. (1958). "Indian Psychology." Calcutta.

Skarda, C.A. and Freeman, W.J. (1990). *Concepts in Neuroscience* 1, 275-285.

Sutton, J.P., Beis, J.S., and Trainor, L.E.H. (1988). *J. Phys. A: Math. Gen.* 21, 4443-4454.

Terrace, H.S. (1985). *Phil. Trans. R. Soc. Lond.* B 308, 113-128.

Turney, T.H. (1982). *Bull. Psychon. Soc.* 19, 56-62.

Turing, A.M. (1950). *Mind.* 59, 433-460.

Weiskrantz, L. (1986). "Blindsight." Oxford University Press, London.

Wheeler, J.A. (1980). "Delayed choice experiments and the Bohr-Einstein Dialog." Paper presented at the Joint Meeting of the American Philosophical Society and the Royal Society, London, June 5.

Wigner, E., (1967). "Symmetries and Reflections." Indiana University Press, Bloomington.

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

Zeki, S. (1992). *Scientific American.* 267(3), 69-76.