Reflections in Silicon: Artificial and Natural Neural Networks

by

Rick Willard Tanney, B.A., B.S., M.Div., MSCIS

Report

Presented to the Faculty of the Graduate School

of The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 1997

Reflections in Silicon: Artificial and Natural Neural Networks

APPROVED BY

SUPERVISING COMMITTEE

Dedicated to the Old English scribe, Gerefa, who recorded this witness speaking in exasperation at a trial:

Ic gecende be ðam ðe ic cuðe;
se ðe bette cunne gecyðe his mare .[1]

1 Old English : "I have set forth [the subject] as best I could; let him who knows [it] better make more of it known." *Gerefa*. Baugh, Albert C. *A Literary History of England*. New York: Appleton-Century-Crofts, 1967, 36.

Preface: A Look Before the Leap

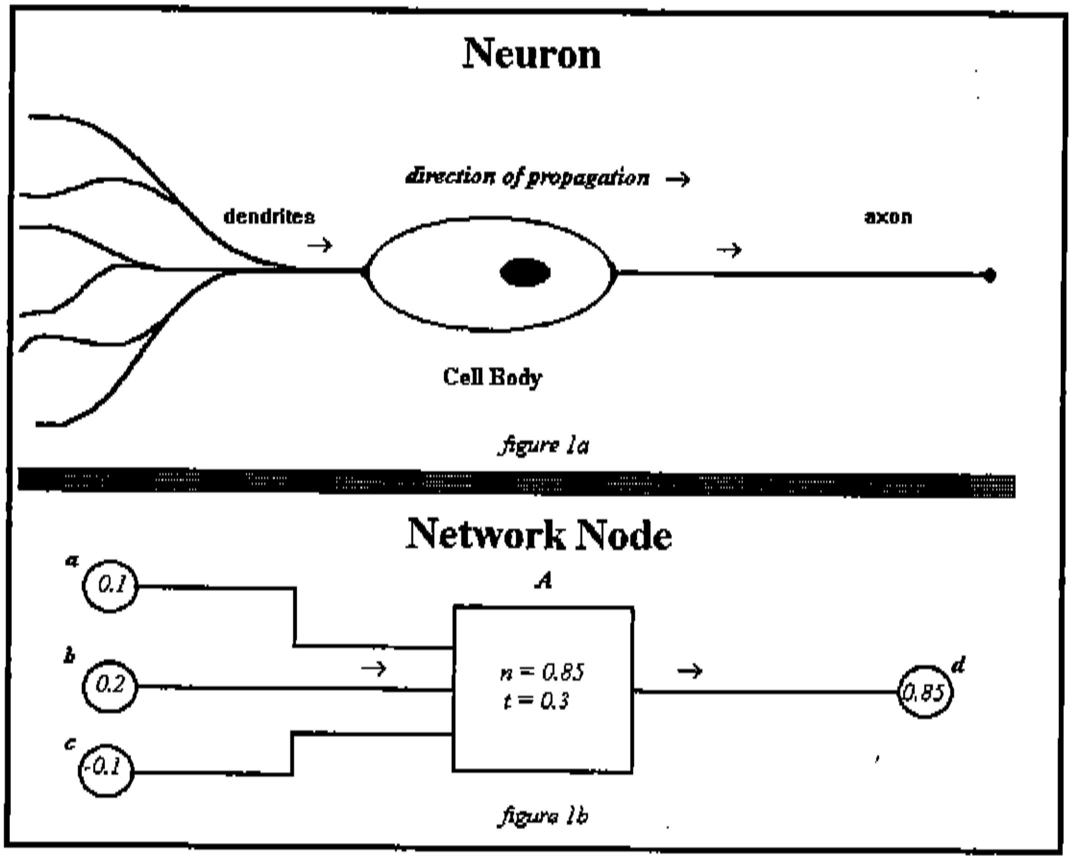


Figure 1. Comparing Neurons and Network Nodes

The basic function of the neuron is this (figure 1a. Neuron)[2]:

1. Input signals arrive from the axons of other neurons through the dendrites
    a. excitatory signals increase the likelihood of activation
    b. inhibitory signals decrease the likelihood of activation

---

2. Rahmann, H., and M. Rahmann. *The Neurobiological Basis of Memory and Behavior*. New York: Springer-Verlag. 1992, 3.

2. Signals are summed and compared to the activation threshold

3. If the activation threshold is reached

    then an output signal is sent through the axon to the dendrites of other neurons

    else the activation level of the neuron decreases

The basic computer neural network neuron simulation is this (figure 1b Network Node):

1. Input signals arrive from other network nodes through the program algorithm

    a. excitatory signals increase the likelihood of activation

    b. inhibitory signals decrease the likelihood of activation

2. Signals are summed and compared to the activation threshold

3. If the activation threshold is reached

    then an output signal is sent through the program algorithm

    else the activation level of the node decreases

Many factors affect the flow of activation through the nervous system:

1. Placement in the structure being activated

2. Chemical environment

Factors affecting the flow of activation through a computer neural network are less complex, and include only the algorithm used for propagation, and the data set fed to the system. The basic operation of the individual nodes or neurons, however, seems to be very simple: sum inputs, if the sum is greater than some threshold value, propagate a signal.

This basic operation goes on continuously in the brain. Significant variations occur in the structures active at any given moment. Much text has been dedicated to "brain states," but this term is applicable in a very limited sense. Physically, the only time a brain can be said to be in an actual state is when it is dead. Otherwise, there is a continuous flow. The brain does not maintain a static state while it waits for more input. Nothing is done in the brain by achieving some static arrangement, but only by moving through structures which are themselves constantly subject to modification. Heraclitus' dictum *vis à vis* the river of time applies equally to the brain: you can't progress through the same brain twice. Many general structures remain relatively the same over time, but all are constantly subject to minor modifications. The rate of attrition from disuse of language demonstrates that, without rehearsal, knowledge fades.[3] [4] [5] The brain does not permanently encode its contents, but relies on occasional refreshment or reinforcement. Long-term memory implies long-term reinforcement. The reinforcement need not be from external sources. The unconscious seems to keep busy reinforcing structures by training the brain's neural network through dreams.[6]

3. Brown, Douglas H. *Principles of Language Learning and Teaching, 2d ed.* Englewood Cliffs, NJ: Prentice Hall Regents, 1987, 68-70.
4. Rahmann. *Ibid.* 220.
5. Klimesch, Wolfgang. *The Structure of Long-Term Memory.* Hillsdale: Lawrence Erlbaum Associates, 1994, 12.
6. Fuster, Joaquin M. *Memory in the Cerebral Cortex*: An Empirical Approach to Neural Networks in the

In a computer neural network, the nodes are trained on a given set of data, then frozen while the system is in use doing what it is intended to do. The dynamism which is constant in the brain is only present in the initial training phase of a neural network. Further training may be done, but for the most part, the net is frozen for testing. This paper will focus on the philosophical questions involved in comparing brain function with neural networks, but first I must introduce the basic structures involved in both brain activity and the related neural network paradigms. Both areas are currently in a dynamic state of flux (to use "state" loosely). New brain studies are revealing faults in recent theories of brain function as well as more detailed data regarding how the structures in the brain interact. Computer neural networks are constantly undergoing refinement and facing new challenges. Does knowledge in one of these systems imply similar knowledge in the other? We will first look at the current view of brain function, then the current state of neural networks, before plunging into the details of philosophical arguments about connectionist versus symbolic paradigms, dynamism versus states, single unit versus functional structure, imposed structure versus self-organization, local versus global representation, etc.

Cognitive science is truly a multi-disciplinary field: the philosophical problems are not resolvable by reason alone: reference must be made to the physical truths made manifest by on-going research. The computerized approach will be somewhat isolated

Human and Nonhuman Primate. Cambridge: The MIT Press, 1995, 284--288.

in its uses and implications for humanity without some correlation to the philosophical questions relating to *What is man? What is thought? What is consciousness?*. Psychology and linguistics are also deeply engaged in this exploration of *what is?*, *what happens?*, and *how do we value it?* Especially psychology has a great stake in the outcomes of research into the workings of the brain. Who is closer to the truth, Pavlov or Freud, Freud or Jung, Jung or Piaget? The investigation of neural networks is relevant to all these areas. If computer neural networks can be shown to be in a meaningful way equivalent to the neural network of the brain, these investigations could lead to dramatic progress in many significant areas of human daily life, especially in how we think about *how we think*.

The word "model" has a great many uses. In logic it is: "An interpretation I is called a *model* of Γ iff it satisfies Γ. Γ is *satisfiable* iff it has a model."[7] We use it in everyday language to mean a paradigm, as in "a model citizen." It refers to small planning versions of a building. "Model" has many common uses. Consider:

> *I met a girl at a party. She told me she was a model.*
> *I asked, "What scale?"*
> *"Union," she sighed.*[8]

There are scientific and literary models, economic and artist's models, models of trains, planes and automobiles. I will use the word "model" many times in this paper. I mean a particular sense when I use it. This definition is no substitute for the

7. Causey, Robert L. *Logic, Sets, and Recursion*. Boston. Jones and Bartlett Publishers, 1994, 298. "Γ" is a set of Predicate Calculus *(PC)* sentences, and I is a *(PC)* interpretation.
8. Tanney, Rick W. "Toward a Connectionist Lexicon and the Possibility of Disambiguation." Unpublished.

function of either an artificial neural network or of a natural neural network (the nervous system), but I think it may help to link those concepts in both types of network that we seek to elucidate and understand. When I use the word "model", I mean a structure which includes three associated parts: 1) *what is*, 2) *how it works*, and 3) *what value it has*. When I speak of *what is*, I include what physically or psychically exists, to include both existential properties and actions associated with the object or concept. This means our evaluation (conscious or not) of whether the thing in question has physical existence and what its physical properties are, or whether it has a conceptual existence. "Cat" has a physical existence; "addition" exists only as a concept. *What is*, for "cat", will include "a physical presence which is a furry, quick, short mammal with (usually) a long tail and whiskers"; for "addition", it will include "a process by which things are serially joined together." For "cat," this includes such things as jumping, purring, and eating birds; for "addition," it includes combining two elements to form a sum. When I speak of *how it works*, I include what causal relations apply to *what is*. When the cat is stroked, it purrs, that is, *stroking* causes *purring*; when it wants to catch a bird to eat, it jumps -- this exemplifies two causal relations -- *wanting* causes *jumping* which results in *catching*. *What value it has* includes what our interest in it is, its survival value, pleasure value, its relative desirability or acceptability. For "cat", this includes the pleasure I feel when I stroke it and it purrs. For "addition", it includes the practical value in being able to join many things together to be treated as a group rather than going through the tedium of treating each element individually. These three elements combined comprise all relevant factors I can think of that are needed to form a working black box which can process incoming data relating to the thing we have in mind and output a meaningful result. I call this the "Model model."

Science is developing a model of brain function which has implications for our philosophical models of thought, perception and knowledge. Neural network models can be tools for manipulating and evaluating these models. If we can develop valid working models which are essentially the same in both areas, we may be able to use neural networks for many practical as well as philosophical investigations. Natural neural networks are the basis not only of thought, but of memory -- short, medium, and long-term -- which is necessary for our sense of identity. Through artificial neural network simulation, we may discover more about who we are, and how we are aware of who we are.

Reflections in Silicon: Artificial and Natural Neural Networks

by

Rick Willard Tanney, M.A.
The University of Texas at Austin, 1996
SUPERVISOR: Robert L. Causey

Abstract

The concept of neural networks as serious analogs for human brain function is very appealing, but controversy continues over how significant they are, and in what contexts they may be applicable. They are of serious interest for philosophers for their implications for theories of mind, for issues involving thought, perception and knowledge. There is serious doubt about whether an electronic mechanism, no matter how elaborate, can meaningfully mimic the functions of a complex electro-biochemically active organ. Proponents believe that neural networks can perform cognitive tasks at a functional level, and that we may be able to learn much about our own cerebral activities from their study. I conclude that neural networks are an important and essential model for mental function, but that the neuronal circuitry alone is not sufficient to explain motivation or overall system guidance.

Table of Contents

List of Tables Page

List of Figures Page

Introduction

The brain is amazingly complex, yet scientists have been making great strides in discovering its workings.[1] To a great extent, brain functions seem to be based on a few simple principles and procedures involving neurons which are made powerful through the structure of their synaptic connections. New techniques are constantly being developed to allow an ever finer-grained examination of the brain at work. The functions of smaller and smaller portions of the brain's rich structure are being defined with increasing clarity. The black boxes are getting ever smaller. Effective models are being developed at many functional levels, and the contribution of each level to over-all function is being integrated into an overall picture. The picture is far from complete, and we will have scant space in this paper for a detailed view of even what is currently known. I will try to paint a broad enough picture not to get bogged down in unnecessary detail, but clear enough to indicate the significance of each level and how it applies to questions of neural network simulation and philosophical implications for theories of mind.

Neural networks as brain simulators are of relatively recent origin, having begun with the pioneering work of McCulloch and Pitts in 1943.[2] There are many forms of neural network, the primary ones being back-propagating networks and self-organizing maps. I will try to paint a broad picture, and focus on the pertinent

1. The 1990's have been designated the "Decade of the Brain."
2. Haykin, Simon. Neural Networks. New York. Macmillan College Publishing Company, 1994, 36.

implications. As computers become faster and increase in capacity, simulations will become more and more effective, and come closer to modeling the real neural networks of the brain. Their range is currently very limited compared to the vast capacity of the brain, megabytes compared with multi-gigabytes. But it is the emulation of structure which is most important, and which is most controversial. Can a mere electronic device act like a very complex electro-biochemical organ? I hope to lay before you a lucid discussion of the issues as they currently stand.

What can we hope to explain using neural networks? Learning and memory -- how do we learn? How do we forget what we've learned? Knowledge representation -- knowledge is not stored within the neurons, but in the structure which connects and activates them. Can we use this structural approach to good advantage elsewhere? What is the relation of the real world to our perception of it? Will our concept of the real world change once we discover *how* we understand it internally? Humor. Puns. Irony and ambiguity. I believe that humor is the result of the overlap of models stored in the neural network. Volition, motivation. What are the mental mechanisms which drive us to do what we do? Spoonerisms. Language -- it's not straightforward symbol processing, so how do we coordinate printed words, their sounds, and our internal representation of what they mean in the real world? *Déjà vu, Presque vu, Jamais vu* -- I've done this before; I almost saw something; I feel like I've never seen this before (although

I see it every day). Intuition -- I know I'm right, but I can't express why. Symbol manipulation -- if the brain does not actually directly manipulate symbols, how does it learn to identify and usefully manipulate what we describe as a symbolic representation? Consciousness -- is it an emergent property at a high level, or is it like the concerted effort of a group of muscle fibers, each synapse or neuron having a little bit of consciousness? The difference between "my arm goes up" and "I raise my arm" is consciousness. Psychological development -- how does our brain and its functions develop from the overabundance of neurons that we start with to the set we grow up using? What are the incremental changes involved? Can we effectively apply this information to treatment of mental problems or to educational programs? Intelligence -- is it a measure of how many models of a certain sophistication we can usefully recognize and manipulate? Is there one basic type of intelligence or seven or more? Phantom limbs -- they are no longer there -- how does an amputee continue to have the sensations of the missing limb? Dreams -- are dreams the brain's way of retraining the nervous system while we sleep?[3] [4] Systems effects, emergent properties. How is it that unexpected abilities or functions become apparent when elements of the substrate act in concert? What structural dynamics account for these properties? The combined efforts of neural scientists and computer neural network experimenters can help eliminate inappropriate theories of mental function using the Sherlock Holmes Rule of

---

3. Wolf, Fred Alan. *The Dreaming Universe*. New York: Touchstone/Simon & Schuster. 1994.
4. Rahmann, *Ibid.*

Evidence: when all the impossibilities have been eliminated, whatever remains, however improbable, must be the correct answer. Neural networks can be trained to translate neural signals to movement in a prosthesis. There is a possibility of mapping visual activity in the brain to a neural network. We may be able to directly read the visual content of the brain, perhaps even recollections of things seen or merely imagined. To view the imagination at work is becoming at least a theoretical possibility. We have reached the outer limits of speculation.

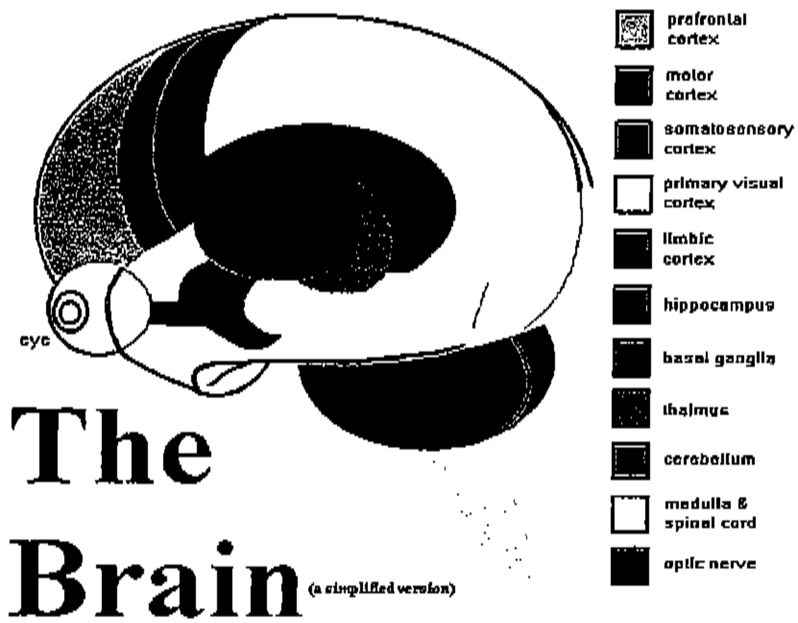


Figure 02: "...the most complex object in the universe. It comprises a trillion cells, 100 billion of them neurons linked in networks that give rise to intelligence, creativity, emotion, consciousness and memory." [5]

---

[5]. Fischbach, Gerald D. "Mind and Brain", Scientific American, September 1992, 51. This article provides the bulk of the discussion and description of brain structure in this exposition.

I. A Natural Neural Network: The Brain

Composition

The brain is composed largely of neurons which are connected to each other in multiple ways, neurons which operate as tiny individual computers, summing inputs and sending out signals to other neurons, both locally and at a distance, neurons which are integrated into highly complex functional systems in a hierarchy which begins with the lowly synapse and culminates in human behavior. Analysis of these structures at a fine level is made difficult by several obstacles: (1) the system is immensely complex, composed of several hundred billion neurons; (2) the number of connections is extremely large -- as many as 10,000 terminals per neuron; (3) there are diverse modes of synaptic association between two neurons -- excitatory-inhibitory, electrical-chemical; and (4) there is a great diversity of transmitter substances, probably around 50.[6] [7] This lack of more detailed information about brain structures and their functions will make the comparison to artificial neural networks even more difficult.

Structure

The number of identifiable levels of brain structure and their significance for cognition are issues for debate. The human brain weighs three to four pounds and contains about 100 billion neurons. This immense number alone does not account for the brain's complexity:

6. Rahmann, *Ibid.* 99.
7. Fischbach. *Ibid.* 48-57. As of this writing, there are (at least) between 100 and 200.

> Although that extraordinary number is of the same order of magnitude as the number of starts in the Milky Way, it cannot account for the complexity of the brain. The liver probably contains 100 million cells, but 1,000 livers do not add up to a rich inner life.[8]

There is a great diversity of cells -- Purkinje, basket, motor neurons, etc. A Purkinje cell alone is amazingly complex. The challenge to define the function of one such cell -- or even a small group taken together -- is daunting. The brain is not one smooth net of similar cells performing identical functions.

There are some simplifications in brain structure which do facilitate analysis of structure and function: 1) Groups with similar functions are grouped together in columns (or "slabs") that extend through the thickness of the cortex. For example, a module in the visual cortex which responds to a line of a particular orientation could contain up to 100,000 cells, most of which participate in local circuits devoted to a particular function; 2) All neurons conduct information in much the same way: information travels along axons in the form of brief electrical impulses called action potentials, which measure about 100 millivolts in amplitude and one millisecond in duration. These result mainly from the exchange of potassium and sodium ions across the surface membrane of the neuron from the extracellular fluid into the cell interior or cytoplasm. At a critical potential called the threshold, an electrical charge is generated and positive feedback produces a regenerative event that forces the membrane potential to reverse sign. The sodium permeability mechanism remains refractory for a

8. Fischbach. *Ibid.* 49.

few milliseconds after each pulse which limits the rate at which action potentials can be generated to 200 per second or less . The need for signal boosting limits the maximum speed at which an impulse travels to about 100 meters per second, less than one millionth the speed at which an electrical signal moves in a copper wire. Thoughts must depend on the relative timing of impulses conducted over many axons in parallel and on the thousands of connections made by each one.

Transmitters

Communications between neurons is mediated by chemical transmitters that are released at specialized contacts called synapses, which, we will see later, are the best candidates to be the basic computational unit of the network.[9] Each neuron must continually integrate up to 10,000 synaptic inputs, which do not add up in a simple linear manner, which means each neuron is a sophisticated computer in itself.

> Transmitter receptors can be grouped into two large (and growing) superfamilies based on their amino acid sequence and on presumptions about the shape that the molecules assume as part of the cell membrane in which they are embedded . . . one receptor superfamily consists of ion channels, [which] underlie . . . changes in permeability . . . The other superfamily . . . does not form channels. Instead its members interact with a neighboring membrane protein . . . This process initiates a cascade of biochemical reactions.[10]

This chemical interaction must take place at an incredible pace in order to have a timely effect on the electrical activity in the system. The pace of chemical change must be at least that of the synaptic interactions with the neurons, faster than 100-200

9. See Appendix 4. Defining the Basic Computational Unit.
10. Fischbach. *Ibid.* 54.

transformations per second.

Plasticity

The account of structural, functional and molecular variety given so far would seem to be sufficiently complete as a basis for mental function. We have yet to consider plasticity, the tendency of synapses and neuronal circuits to change as a result of activity. "Plasticity weaves the tapestry on which the community of mental life depends."[11] The metabolic after-effects of action potentials not only encode information, they alter the circuits over which they are transmitted.

Synapse plasticity is the basis for connectionist neural models. It "multiplies the complexity provided by any fixed cast of molecular characters or cellular functions," providing an even richer substrate for mental phenomena. These changes may alter function of synapses as well as the number or location of synapses themselves. Axons generate new endings when nearby neurons cease transmission, and the terminal branches of dendritic arbors are constantly remodeled.

> Although the forces leading to plastic changes in the mature brain are frequent and ineluctable, it is important to emphasize the precision and overall stability of the wiring diagram. We could not sense the environment or move in a coordinated manner, let alone think, if it were otherwise. All studies of higher brain function must take into account the precise way in which neurons are connected to one another.[12]

11. Fischbach. *Ibid.* 54.
12. Fischbach. *Ibid.* 55.

Topographic Mapping

> Pathways in the brain have been traced by means of a variety of molecules that are transported along axons. Such reporter molecules can be visualized once the tissue is properly prepared. Connections have also been traced by fine-tipped microelectrodes positioned close enough to a nerve cell body or an axon to detect the small currents generated as an action potential passes by. Each technique has revealed ordered, topographic maps in the cerebral cortex. The body surface is represented in the postcentral gyrus of the cerebral cortex even though the cortical neurons are three synapses away from sensory receptors in the skin. Likewise, a point-to-point map of the visual world is evident in the primary visual cortex at the occipital pole at the back of the brain. Order is evident at each of the early relays on route to the cortex, and topographic order has also been found in projections from the primary cortices to higher centers.[13]

Hubel and Wiesel discovered 30 years ago that neurons in the primary visual cortex (V1) respond to line segments or edges of a particular orientation rather than to the small spots of light that activate the input neurons in the retina and lateral geniculate nucleus of the thalamus, implying that neurons in V1 are connected, via the lateral geniculate nucleus, to retinal ganglion cells that lie along a line of the preferred orientation. "There is a general topographical correspondence between a region of simulation of the sensory organ and its representation in the central nervous system . . . While maintaining topographical coherency, the nature of the information stored in the brain can change as it progresses from one way-station to the next."[14]

13. Fischbach. *Ibid.* 55.
14. Cook, Norman D. *The Brain Code*. New York: Methuen, 1986, 33.

Brain Code

We know the anatomy of the major sensory and motor systems in some detail. However, the pattern of connections within the intervening association cortices and the large subcortical nuclei of the cerebral hemispheres is not clearly defined. Visual information is mapped so that the retinal image in the eye maintains its basic configuration as the pattern of stimulation is transferred from the retina to the lateral geniculate body in the thalamus to striate cortex. Somatosensory information is also topologically intact, with the entire surface of the skin being represented contiguously and specially detailed representation for the hand. Cook makes a distinction here between study of the neuron and its synaptic activity, which he terms *neurophysiology*, and the study of larger aggregates of neural activity which account for specific cognitive functions or for functions which taken together account for mental and physical behavior, which he terms the *Brain Code* or *neuropsychology*.[15]

The pattern of information flow in the brain during the performance of mental tasks cannot easily be determined by anatomic studies of the circuit diagram or by studies of plasticity. Neural correlates of higher mental functions are being sought directly in awake primates trained to perform tasks that require judgment, planning or memory, or all three capacities.

15. Cook. *Ibid.* xiii.

Hierarchies

> One of the most important principles is that sensory systems are arranged in a hierarchical manner. That is, neurons respond to increasingly abstract aspects of complex stimuli as the distance -- measured in numbers of synapses from the source -- grows. The fact that neurons in V1 respond to lines rather than spots makes the case. Another important principle . . . is that information does not travel along a single pathway. Rather, different features of a single percept are processed in parallel pathways . . . the movement, color and shape of a tennis ball are processed in different cortical visual centers.[16]

The auditory system has a similar topographical organization, with nearby structures responding to similar frequencies. But, in the barn owl, phase and amplitude signals which account respectively for location along the azimuth and elevation are processed in different pathways through three synaptic relays in the brain. It seems likely that this type of parallel processing characterizes other sensory systems, association cortices and motor pathways as well. For years, psychologists have seen the normal brain as a black box whose inner workings were the speculations of many different schools of thought. Occasionally some area of a brain would be damaged, and the corresponding loss of function would give a clue as to the function performed by that part of the brain. In this way Broca's Area and Wernicke's Area were respectively found to process speech production and language comprehension.[17] We are now reaching the point at which it is becoming possible to reach deeper into the structure and to more clearly identify the functions of smaller and smaller local

---

16. Fischbach. *Ibid.* 56.
17. Cook. *Ibid.* 19.

structures in the brain.

Cerebral Hemispheres

> Where is the information reassembled? When does the subject become aware of the approaching ball? The receptive fields of neurons in higher centers are larger than those found in earlier relay stations, so they monitor a larger fraction of the external world. Zeki describes a model that depends on feedback connections from cells with large receptive fields to the cells in the primary visual cortex that have high spatial resolution. Such feedback circuits might coordinate the activity of cells in the primary cortex that have high spatial resolution and cells that respond to more abstract features of the stimulus no matter where it is located. Francis Crick and Christof Koch address the role in visual awareness of a 40-cycle-per-second oscillation in firing rate that is observed throughout the cortex . . . [which] may synchronize the firing of neurons that respond to different components of a perceptual scene and hence may be a direct neural correlate of awareness.[18]

How do we close the big loop?[19] I mean the feedback loop that accounts for hand-eye coordination, that doesn't just see the ball coming, but takes steps to intercept it. Fischbach asks where it all comes together. Cook has a suggestion for where to look: the corpus callosum. The brain is composed of a number of identifiable parts, the most identifiable being the two lateral hemispheres. The cerebral cortex is outer 2-3 mm of the entire surface of the cerebral hemispheres. Its surface forms a series of elevated ridges, or gyri, separated by shallow depression, called sulci, or deeper grooves, called fissures.[20] The total surface area is approximately 25 square

18. Fischbach. *Ibid.* 56.
19. Dreyfus. Hubert L. "Phenomenology of Embodiment." *The Electronic Journal of Analytic Philosophy*, 4 (Spring 1996). Bloomington: Indiana University. 1996. ¶ 32.
20. Martini, Frederick H. *Fundamentals of Anatomy and Physiology. 3d ed.* Englewood Cliffs: Prentice Hall. 1995, 468.

feet of flat surface -- it is significant that the topology of the cortex is actually two- rather than three-dimensional.

The prefrontal cortex is involved with conscious intellectual functions. The frontal lobe includes the primary motor cortex, and processes voluntary control of skeletal muscles. The parietal lobe is associated with conscious perception of touch, pressure, vibration, pain, temperature, and taste. The occipital lobe contains the primary visual cortex, and the temporal lobe accounts for auditory and olfactory cortices. The sensory and motor regions of the cortex are connected to nearby association areas that interpret incoming data or coordinate motor response. Integrative centers include the prefrontal cortex, which integrates information from sensory association areas and performs abstract intellectual functions, such as predicting the consequences of possible responses, the general interpretive area (for language and mathematical calculation), which is usually confined to one (the left) hemisphere, and the speech center.[21]

The corpus callosum contains over 200 million axons, carrying an estimated 4 billion impulses per second.[22] Its central location and function as the carrier of signals between hemispheres are what made it the focus of Cook's *Brain Code* book. He lists the accompanying table (Table 1.) of *dichotomies attributed to the cerebral hemispheres*:[23] Martini calls the left hemisphere the *categorical hemisphere*, the right

21. Martini. *Ibid.* 470-2.
22. Martini. *Ibid.* 469.
23. Cook. *Ibid.* 18.

hemisphere the *representational hemisphere*.[24]

Cook argues the necessity of one hemisphere to dominate in certain activities -- for example, the coordination of movement and expression. You can't clap your hands very well if each hand is controlled independently, so one side must coordinate the effort. The communication for this coordination goes through the corpus callosum.

There is still much to be found out about brain function in the various areas. A recent study challenged the traditional theory of cerebellar function, showing that the cerebellum is not directly responsible for fine motor control, but is involved "in sensory discrimination rather than in movement per se."[25] And we still wonder where it all comes together. Our most persistent interest in this pursuit of knowledge is the source of consciousness, which we frequently associate with cognition. But the brain performs many unconscious functions at many levels. One wonders just how many of those levels we are privileged to experience as consciousness. Are we limited to one? Do we occasionally get a glimpse into others, as when we are between dream and wakefulness?[26]

Firing Rate and Vectors

How many neurons must change their firing rate to signal a

---

24. Martini. *Ibid.* 471-2.
25. Gao, Jia-Hong, Lawrence M. Parsons, James M. Bower, Jinhu Xiong, Jinqi Li, peter T. Fox. "Cerebellum Implicated in Sensory Acquisition and Discrimination Rather than Motor Control." Science VOL N.d. 1996, 1-3.
26. Wolf. *Ibid.*

Table 01. Dichotomies attributed to the cerebral hemispheres

Left Hemisphere	*Right Hemisphere*
Experimentally Derived	
sequential	parallel/simultaneous
analytic	synthetic
linguistic	visuospatial
verbal	nonverbal
focal	global
linear	spatial
mathematical	geometric
Philosophically Motivated	
expressive	perceptive
propositional	appositional
symbolic	imaginative
intellection	emotion
reason	affect
discrete	holistic
active	passive

coherent percept or gestalt? The most extreme view holds that one cell may do the job. Is there one face cell per face? Such a supposition seems unlikely on first principles: we lose thousands of neurons every day, so overcommitment to one would be unwise. A more compelling argument comes from recent experiments that have shown face cells to be broadly tuned, responding to faces with similar features rather than to one face alone. The number of neurons that must be activated before recognition emerges is not known, but the data are consistent with a sparse coding rather than global or diffuse activation.

Face cells have their counterparts on the motor side. "Command" neurons have been identified in certain invertebrates that trigger all-or-none, fixed-action patterns, such as stereotypical escape behaviors . . . like face cells in the temporal lobe, individual motor cortex neurons are broadly tuned.

The vector obtained by summing the firing frequencies of many neurons is better correlated with the direction of movement than is the activity of any individual cell. The vector becomes evident several milliseconds *before* the appropriate muscles contract and the arm actually moves. It must be a sign of motor planning. The vector is usually derived from less than 100 neurons, so sparse coding may be the rule in the motor cortex as it is in the temporal sulcus.[27]

Experimental Evidence

Studies are underway to produce mental phenomena by focal electrical stimulation. This harkens back to the memory stimulation exercises by Wilder Penfield, in which he evoked vivid memories from patients by stimulating their brains directly electronically while they were undergoing brain surgery.[28] Strokes and other unfortunate "experiments of nature" have also provided important insights regarding neural correlates of mental phenomena.

The future of cognitive neuroscience depends on our ability to study the living human brain. PET and MRI hold great promise in this regard [29]. The brain is never completely at rest. At present, neither technique provides the spatial resolution to visualize single cortical columns.

Fischbach believes we can expect advances at an increasing rate on all levels of investigation relevant to the mind: We will soon know exactly how many transmitters and transmitter receptors there are in the brain and where each one is concentrated . We will also have a more complete picture of neurotransmitter actions, including multiple interactions of jointly released modulators. And we will learn much more about molecules that affect neuronal differentiation and degeneration. The great

---

27. Fischbach. *Ibid.* 57.
28. Penfield, Wilder. *The Mystery of the Mind.* Princeton: Princeton University Press, 1975.
29. Gao, et al. *Ibid.* 1-3.

challenge is to determine how these molecules modulate the functional wiring diagram of the brain and how this functional nerve network gives rise to mental phenomena.

Emergence

Ultimately, it will be essential to specify what exactly it means to say that mental events are correlated with electrical signals. Is the mind an emergent property of the brain's electrical and metabolic activity? An emergent property is one that cannot be accounted for solely by considering the component parts one at a time . . . biological explanations of mental events may become evident once the component neural functions are more clearly defined. We will then have a more appropriate vocabulary for describing the emergent mind.[30] *Emergence* is frequently discussed when systems produce behavior which can not be expected from a minute analysis of all of the subcomponents, in other words, when a system exhibits behavior which can not be predicted by examining the properties of its constituent parts. Consider *sand*. Consider *dune*. Even if you know all properties of each individual grain, and the dynamic relations among the grains, and properties of the grains when aggregated, you do not have enough information at that level to explain an individual dune and its systemic properties.[31] An appropriate explanation can be made at a higher level of organization, one that recognizes the applicable models for friction and fluid flow of granular aggregations. Once all factors have been considered and weighed,

30. Fischbach. *Ibid.* 57.
31. Nagel, Ernest. *The Structure of Science*. New York: Harcourt, Brace & World, Inc. 1961. 366-80.

both at the substrate level and at the local level, environmental factors may have a role which is yet unexplained. When we understand all the factors affecting a system, we hope we will have a set of laws which completely predict its behavior, but until such a comprehensive system description is confirmed, we may profitably consider unexplained system phenomena as *emergent*. Supposedly, equally mysterious properties *emerge* from other types of aggregates, such as neuron bundles. Consider muscle. Consider muscle fiber. Here, the properties of the muscle inhere to each individual fiber, and the power of the muscle is merely an aggregation of smaller powers.[32] Is the mind merely an aggregate of mindlets? Or is there something that must be explained by appeal to a higher level model? We are ever closer to plausible explanations for the physical basis of mental phenomena, but much more experimental evidence must be processed before we can claim to know how these systems work. In the meantime, the interplay between investigations of artificial and natural neural networks provides insights for all neural networks.

There is a brilliant visualization of what one layer of a neural network might look like and how it might function in "Pinscreen Excerpts," featuring the animation work of Alexander Alexeieff:[33] A pinscreen consists of thousands of dark pins closely clustered and thrust through a white flat background. The height of the pins above the

32. Shepherd, Gordon M. "The Significance of Real Neuron Architecture for Neural Network Simulations." Chapter 8 in Eric Schwartz, *Computational Neuroscience*, (q.v.)
33. Alexander, R. McNeill. *Exploring Biomechanics: Animals in Motion.* New York: Scientific American Library. 1992.

surface and the angle of lighting determines how much white shows through, creating a very finely-grained black-and-white picture. The pins are adjusted frame after frame to create a moving picture in the same way that other animation techniques create an impression of movement through gradual change. The difference is that the subtleties of the representation are made using small changes in individual elements which are not simply on or off or in a certain color or contrast, but which are arbitrarily finely adjustable, and which yet form a readily identifiable whole. To an outside observer, of course, this representation is no more enlightening than a table of node weights. But some mechanism monitors these weights, compares the results of applying the current network, and retrains the network to respond with a different weighting on the next input. A picture emerges from a pinscreen, but this is due to conscious human activity. We are not conscious of the activity which governs our own natural neural network. We trust that, at some level, there is an explanation for everything. As advances are made, not just in neurological science, but in other sciences, both hard and soft, for clues which will lead us to a clear understanding of what is really meant when we say we think, or that we have thoughts, or that we perceive, or understand. New ground is being broken in neural studies which show us that the brain, even now, is not necessarily doing exactly what we think it is.[34] Studies in physics show us that there is a natural tendency towards self-organization in dynamic systems.[35] Sand.

---

34. Gao, et al. *Ibid.*
35. Umbanhowar, Paul B., Francisco Melo & Harry L. Swinney. "Localized excitations in a vertically vibrated granular layer." *Nature* 382, 793-796. 1996. I suspect there may be similar activities in natural neural

Dune. Two different models constituted of the same substance, the one the sole substrate of the other. Pile up the sand -- which particle made the dune? How we see it and understand it as a model differs as our model of sand and our model of dune differ. Our model of dune has nothing directly to do with our model of sand particle. A dune has topological implications completely missing in the sand model.

A recent issue of Scientific American[36] showed magnified images of sand particles form many locations around the world. It showed the particles to be extremely varied, and of many beautiful designs. Our normal sight can barely differentiate one particle from another. This beauty of form is missing from our model of sand.

Anything seen macroscopically presents a different model than that seen microscopically, with the exception of fractal patterns (little fleas have lesser fleas . . .) The significance of a model at one level need not have any bearing on the significance of a higher level model, even though it is the sole substrate of that model.[37] On the other hand, some models are so close together that operations on one readily translate into operations on the other. Many such models are mathematically expressible.[38] Tools are available to help us study actual natural neural networks in action: CAT scan, PET, MRI, etc. The use of artificial neural networks as not simply beneficiaries

---

network to those in the granular layers investigated here.

36. Mack, Walter N. and Elizabeth A. Leistikow. "Sands of the World." *Scientific American*. Volume 275, Number 2, August 1996. 62-67.
37. See Appendix 4. Defining the Basic Computational Unit.
38. See Appendix 2. Emergent Properties.

of connectionist theory, but also as tools of experimentation can enhance our knowledge of neural network behavior economically and safely.

II - Artificial Neural Networks

Introduction

The introduction of the idea of neurons as structural constituents of the brain is attributed to Ramón y Cajál (1911).[39] Eighty-five years later, many questions remain concerning just how the system performs the work it performs. Against a background of controversy over just what level of operation is appropriate for discussion of the various levels of function of the natural neural network of the brain, computer scientists are trying to emulate the brain's functions by simulating its operation in artificial neural networks either in a massively parallel computing machine or in a von Neumann architecture which can simulate parallel processing. Neurons are much slower than silicon logic gates, five to six orders of magnitude slower. Events in a silicon chip cycle in the nanosecond (10^{-9} s) range, whereas neural events are in the millisecond (10^{-3} s) range. The brain makes up for the individual neuron's slow rate of operation by having on the order of 10 billion neurons in the human cortex, with 60 trillion synapses or connections among them.[40] The net result is that the brain is an enormously efficient structure, with an *energetic efficiency* of approximately 10^{-16}

39 Haykin. *Ibid.* 1. This source represents the bulk of substantive information on the structure and theory of artificial neural networks in Part II of this report.

40. Churchland, Paul M. "A Deeper Unity: Some Feyerabendian Themes in Neurocomputational Form." Chapter 4 in Steven Davis (q.v.). *Connectionism: Theory and Practice*. 41.

joules (J) per operation per second, whereas the corresponding value for the best computers currently (late 1996) in use is about 10^{-6} joules per operation per second. It is possible on current computers to simulate only a small portion of this complexity in any given instance. It is not yet possible to copy the exact structure of natural networks in artificial networks. The brain is a highly *complex, nonlinear, and parallel computer* (information-processing system) which performs complex perceptual recognition tasks in 100-200 ms, whereas tasks of much lesser complexity will take days on a huge conventional computer. We are currently limited in the extent to which we may attempt to emulate the brain by the availability of, both in terms of computational structure and of raw memory.

At birth the brain has many more neurons, and great flexibility in how they are to be connected. Experience -- perception and interaction with the real world -- leads to restructuring of the brain's circuits, with many neurons disappearing while new synapses continue to be formed. Although development continues well beyond that stage, the most dramatic development (i.e., hard-wiring) of the human brain takes place in the first two years. During this early stage of development, about 1 million synapses are formed per second. The synapse is arguably the *basic computational unit* of the brain. Rahmann states " . . . a neuron, as the elemental unit of the nervous system, constitutes both an ontogenetic (developmental) and a physical (i.e., trophic and functional) unit that is responsible for perception, processing, transmission, and, above all, storage of information in an organism." [41]

Synapses are elementary structural and functional units that mediate the interactions between neurons. The most common kind of synapse is a *chemical synapse*, which operates as follows. A presynaptic process liberates a transmitter substance that diffuses across the synaptic junction between neurons and then acts on a postsynaptic process. Thus a synapse converts a presynaptic electrical signal into a chemical signal and then back into a postsynaptic electrical signal. In electrical terminology, such an element is said to be a *nonreciprocal two-port device*. In traditional descriptions of neural organization, it is assumed that a synapse is a simple connection that can impose *excitation* or *inhibition*, but not both on the receptive neuron.[42]

Haykin identifies a developing neuron with a plastic brain:

> "*Plasticity* permits the developing nervous system to adapt to its surrounding environment. In an adult brain, plasticity may be accounted for by two mechanisms: the creation of new synaptic connections between neurons, and the modification of existing synapses. *Axons*, the transmission lines, and *dendrites*, the receptive zones, constitute two types of cell filaments that are distinguished on morphological grounds; and axon has a smoother surface, fewer branches, and greater length, whereas a dendrite (so called because of its resemblance to a tree) has an irregular surface and more branches. Neurons come in a variety of shapes and sizes in different parts of the brain . . . In its most general form, a *neural network* is a machine that is designed to model the way in which the brain performs a particular task or function of interest; the network is usually implemented using electronic components or simulated in software on a digital

41. Rahmann. *Ibid.* 2.
42. Haykin. *Ibid.* 2.

computer . . . a neural network viewed as an adaptive machine:
A neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:
1. Knowledge is acquired by the network through a learning process.
2. Interneuron connection strengths known as synaptic weights are used to store the knowledge. (adapted from Aleksander and Morton (1990))[43]

The procedure used to perform the learning process is called a *learning algorithm*, the function of which is to modify the synaptic weights of the network in an orderly fashion so as to attain a desired design objective.

Neural networks are also referred to in the literature as neurocomputers, connectionist networks, or parallel distributed processors. A number of benefits are claimed for neural networks. Among them, Haykin cites the following nine:[44]

1. *Nonlinearity* -- In actually using a parallel processing mechanism or simulating one on a sequential computer, it is possible to mimic natural neural networks. Just how valid this emulation may be is a crucial part of our later discussion.

2. *Input-Output Mapping* -- By using *supervised learning*, the network learns from examples, repeating *training examples* until a steady state is reached, and then applying the network to *task examples*, theoretically emulating natural neural network functions.

3. *Adaptivity* -- Networks can be retrained to perform new tasks or modify

43. Haykin. *Ibid.* 2.
44. Haykin. *Ibid.* 4-6.

their performance of old tasks. This flexibility is desirable in many contexts, but adaptivity is not always consistent with robustness. Too great a sensitivity to random noise or other spurious interference can degrade performance unnecessarily. This is the *stability--plasticity dilemma.*

4. *Evidential Response* -- In addition to which selection to make, the system can determine the level of confidence in that particular choice.

5. *Contextual Information* -- is dealt with naturally in neural networks, since knowledge is represented in the very structure and activation state of the network.

6. *Fault Tolerance* -- Neural networks provide a *graceful degradation* in performance when the network is damaged, and may continue to give reliable responses long after a symbolic program will have failed altogether.

7. *VLSI Implementability* -- Implementation using *Very Large Scale Integrated* technology makes possible real-time applications involving pattern recognition, signal processing, and control.

8. *Uniformity of Analysis and design* -- The *universality* of neural networks as information processors facilitates the use of the neuron model as an ingredient common to all neural networks, makes it possible to share theories and learning algorithms, and allows construction of modular networks through a seamless integration of modules.

9. *Neurobiological Analogy* -- Network design is "motivated by analogy with

the brain, which is a living proof that fault-tolerant parallel processing is not only physically possible but also fast and powerful. Neurobiologists look to (artificial) neural networks as a research tool for the interpretation of neurobiological phenomena. For example, neural networks have been used to provide insight on the development of premotor circuits in the oculomotor system (responsible for eye movements) and the manner in which they process signals. On the other hand, engineers look to neurobiology for new ideas to solve problems more complex than those based on conventional hard-wired design techniques . . ."

In Part I, I identified an eight-level structural organization for natural neural networks. Haykin identifies two types of structural model which may be applied, a three-level model, and an eight-level model. The three-level model is comprised of an *input* level, the *stimulus*, a *processing* level which includes a *neural network* sandwiched between *receptors* and *effectors*, and a *response*:

$$\text{Stimulus} \rightarrow [\ \text{Receptors} \leftrightarrow \text{Neural net} \leftrightarrow \text{Effectors}] \rightarrow \text{Response.}$$

This model seems appropriate for psychological applications among others. This model includes feedback between the receptors and the network, and between effectors and the network. The extent and type of feedback that is appropriate in neural networks is a matter for debate. Haykin's eight-level analysis is very similar to that of the natural network scientists:[45]

45. Haykin. *Ibid.* 6-7.

1. Central nervous system

2. Interregional circuits

3. Local circuits

4. Neurons

5. Dendritic trees

6. Neural microcircuits

7. Synapses

8. Molecules

A neural microcircuit refers to an assembly of synapses organized into patterns of connectivity so as to produce a functional operation of interest. These are grouped to form *dendritic subunits* within the *dendritic trees* of individual neurons. The whole *neuron* contains several dendritic subunits. At the next level of complexity are *local circuits* made up of neurons with similar or different properties which perform operations characteristic of a localized region in the brain. Then come *interregional circuits* made up of pathways, columns, and topographic maps, which involve multiple regions located in different parts of the brain. *Topographic maps* are organized to respond to incoming sensory information. Whether this topography is an essential part of cognition is also a matter for contention. These maps are often arranged in sheets, as in the superior colliculus, where the visual, auditory, and somatosensory maps are stacked in adjacent layers in such a way that stimuli from corresponding points in space lie above each other. Finally, the topographic maps, and other interregional circuits mediate specific types of behavior in the *central nervous system (CNS)*.

It is important to recognize that the structural levels of organization which I described are a *unique* characteristic of the brain. They are not found in a digital computer, and we are far from realizing them with artificial neural networks. However, we are gradually making our way toward a hierarchy of computational levels similar to that just described. The artificial neurons we use to build our neural networks are truly primitive in comparison to those found in the human brain. We are not presently able to design a neural network with anywhere near the complexity of the local circuits and the interregional circuits in the brain.

Neuron Models

For artificial neural networks, a *neuron* is a fundamental information-processing unit with three basic elements:

1. A set of *synapses* or *connecting links*, each of which is characterized by a weight or strength of its own. The weight is positive if the associated synapse is excitatory; it is negative if the synapse is inhibitory.

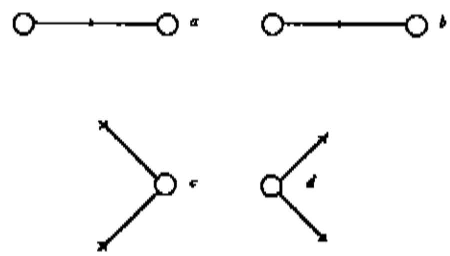


Figure 03. Basic rules for constructing signal-flow graphs.
(Based on Haykin, p.14)

2. An *adder* for summing the input signals, weighted by the respective synapses of the neuron; the operations described here

constitute a *linear combiner.*

3. An *activation function* for limiting the amplitude of the output of a neuron. The activation function is also referred to in the literature as a *squashing function* . Activation functions vary from straightforward threshold functions which have a sharp effective signal cut-off, to sigmoid functions which can have a gradual effect or a sharp effect depending on the slope parameter. The particular mathematical functions used to calculate activation are of little philosophical import except at the sub-synapse level. As we saw in the description of brain function, the synapse is the basic computational unit, and the substrate of the basic computational unit may be safely ignored for the purposes of this paper. For this reason we will eschew a study of activation function mathematics and focus more on structural detail and activation patterns.

Haykin[46] suggests simplifying the appearance of the model of an artificial neuron by using the idea of signal -flow graphs with a well-defined set of rules (Fig. 03). A *signal-flow graph* is a network of directed links (branches) that are interconnected at certain points called nodes. The flow of signals is dictated by three basic rules:

RULE 1. Signal flows along a link only in one direction defined by the arrow on the link. Two different types of links may be distinguished: (a) *Synaptic links,*

46. Haykin. *Ibid.* 13.

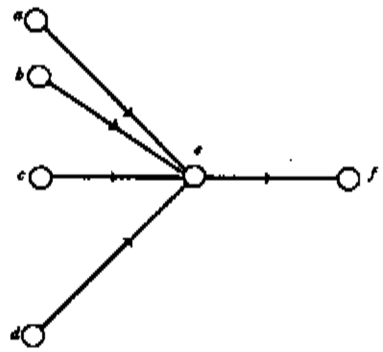


Figure 04: Typical neuron.

governed by a *linear* input-output relation, and (b) *Activation links*, governed in general by a *nonlinear* input-output relation.

RULE 2. A node signal equals the algebraic sum of all signals entering the pertinent node via the incoming links. This rule (c) represents *synaptic convergence* or *fan-in*.

RULE 3. The signal at a node is transmitted to each outgoing link originating from that node, with the transmission being entirely independent of the transfer functions of the outgoing links. This rule (d) represents *synaptic divergence* or *fan-out*.

This signal-flow graph model (Fig. 04) results in the following mathematical definition of a neural network:[47]

> *A neural network is a directed graph consisting of nodes with interconnecting synaptic and activation links, and which is characterized by four properties:*
> 1. *Each neuron is represented by a set of linear synaptic links, an externally applied threshold, and a nonlinear activation link. The threshold is represented by a synaptic link with an input signal fixed at a value of -1.*
> 2. *The synaptic links of a neuron weight their respective input signals.*

47. Haykin. *Ibid.* 14-15.

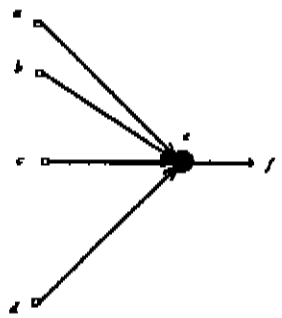


Figure 05. Architectural Graph of a Neuron

3. *The weighted sum of the input signals defines the total internal activity level of the neuron in question.*
4. *The activation link squashes the internal activity level of the neuron to produce an output that represents the state variable of the neuron.*

Such a directed graph is defined as *complete* in that it describes not only the signal flow from neuron to neuron, but also the signal flow inside each neuron. Other representations are also used, such as an *architectural graph* (Fig. 05), which is the result of omitting the details of signal flow inside the individual neurons. Such a directed graph is said to be partially complete, and is characterized as follows:
1. Source nodes supply input signals to the graph.
2. Each neuron is represented by a single node called a *computation node*.

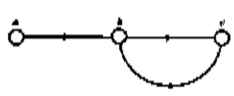


Figure 06. Single-loop feedback

3. The *communication links* interconnecting the source and computation nodes of the graph carry no weight; they merely provided the directions of signal flow in the graph.[48]

Feedback.

Feedback exists in a dynamic system when the output of an element in the system influences the input applied to that particular element, giving rise to one or more closed paths for the transmission of signals around the system. It occurs in almost every part of the nervous system of every animal, and it plays a major role in the study of the special class of neural networks known as *recurrent networks*.[49] There are various mathematical models both implementing and for analyzing feedback. The model selected may result in a stable system in which the signal is convergent, or in an unstable, divergent system. The former has the attribute of *infinite memory*, in the sense that the output of the system depends on samples

48. Haykin. *Ibid.* 15.
49. Haykin. *Ibid.* 15.

> of the input extending into the infinite past. Moreover, the memory is fading in that the influence of a past sample is reduced exponentially with time *n*.[50]

Network Architectures[51]

The structure of a neural network is intimately linked with the learning algorithm used to train the network. We may therefore speak of learning algorithms (rules) used in the design of neural networks as being structured.

We identify four different classes of network architectures:

1. Single-layer Feedforward Networks

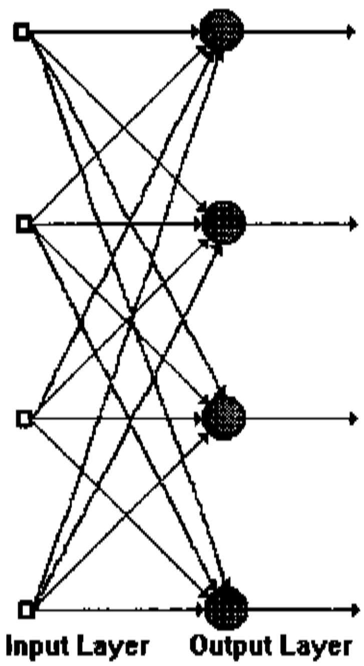


Figure 07. Single-layer Feedforward Network

A *layered* neural network is a network of neurons organized in the form of layers. In the simplest form of a layered network, we just have an *input layer* of source nodes that projects onto an *output layer* of neurons (computational nodes), but not vice versa. In other words, this network is strictly of a *feedforward* type. Such a network is called a *single-layer network* (Fig. 07), with the designation "single layer" referring to the output layer of

50. Haykin. *Ibid.* 17.
48. Haykin. *Ibid.* 18.

computational nodes (neurons). In other words, we do not count the input layer of source nodes, because no computation is performed there.

A linear associative memory is an example of a single-layer neural network. In such an application, the network associates an output pattern (vector) with an input pattern (vector), and any information is stored in the network by virtue of modifications made to the synaptic weights of the network.

2. Multilayered Feedforward Networks

The second class of a feedforward neural network distinguishes itself by the presence of one or more *hidden layers*, whose computational nodes are correspondingly called hidden neurons or hidden units (Fig. 08). The function of the hidden neurons is to intervene between the external input and the network output. Hidden layers allow a system to create its *own* internal representation. By adding one or more hidden layers, the network is enabled to extract higher-order statistics,

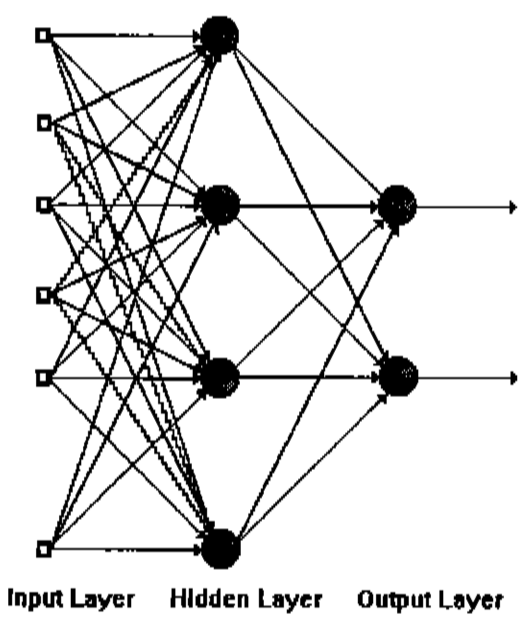
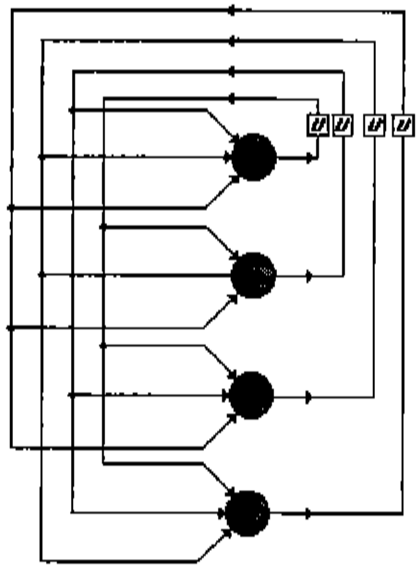


Figure 08. Feedforward Network with Hidden Layer

for (in a rather loose sense) the network acquires a global perspective despite its local connectivity by virtue of the extra set of synaptic connections and the extra dimension of neural interactions.[52] The ability of hidden neurons to extract higher-order statistics is particularly valuable when the size of the input layer is large.



u = unit-delay operators

Figure 09. Recurrent network sans self-feedback.

3. Recurrent Networks

A *recurrent neural network* has at least one *feedback* loop (Fig. 09). It may consist of no more than a single layer of neurons, with each neuron feeding its output signal back to the inputs of all the other neurons. This form may contain no self-feedback loops, and may have no hidden layers. A recurrent network with hidden neurons would combine the graphical features of figures 06 and 07, adding a self-feedback connection.

4. Lattice Structures

A *lattice* consists of a one-dimensional, two-dimensional, or higher-

52. Haykin, *Ibid.* 19.