울산대학교 경영학 연구논문집 제4권 2호 pp 157~170, 1997
Journal of Management Vol 4, No 2, pp 157~170, 1997
University of Ulsan

THE HISTORY AND BIOLOGICAL PROTOTYPE OF ARTIFICIAL NEURAL NETWORKS: FUNDAMENTALS FOR APPLICATIONS

Dae Ryong Kim
Professor of Management Information Systems School of Business Administration

<Abstract>

Research in artificial neural networks has developed in two streams The first is understanding the biological process of mental activity of human brain for developing artificial neural networks that simulate the activity of the human brain The second is developing the physical artificial neural networks to apply the net to the practical application. This paper would like to study the history of the artificial neural networks to propose the potential of the networks The research also want to figure out the biological prototype of artificial neural networks to grasp the connection and resemblance of artificial neural networks to human brain This study can help researchers on artificial neural networks understand how the human brain work and what shape the artificial neural networks must be The author hopes to give researchers the basic idea for further study on the area of artificial neural networks.

인공신경회로망의 역사와 생물학적 모형: 응용을 위한 기초

김대룡
경영학부 경영정보학 전공

<요 약>

인공신경회로망에 대한 연구는 두 흐름으로 전개되어 왔다. 첫째는, 인간 두뇌의 활동을

모방하는 인공신경회로망을 개발하기 위해 인간 두뇌의 지적 활동의 생물학적인 과정을 이해하는데 있다 둘째는, 그 신경회로망을 실제에 응용하기 위하여 물리적인 인공신경회로망을 개발하는데 있다 이 논문은 인공신경회로망의 잠재력을 보여주기 위해 인공신경회로망의 역사를 소개한다 이 논문은 또한 인공신경회로망이 인간 두뇌에 어떻게 연결되어 있고, 어떠한 유사점을 가지고 있는지를 파악하기 위해 인공신경회로망의 생물학적인 모형에 대해 묘사한다 이 연구가 인공신경회로망 연구가들에게 어떻게 인간의 두뇌가 작동하며 신경회로망이 어떤 모양이 되어야 하는지에 대해 이해하는데 도움을 줄 것이다 궁극적으로 저자는 이 분야의 연구자들에게 더 앞선 연구를 위해 기초적인 아이디어를 제공하기를 기대한다

I. INTRODUCTION

Research in artificial neural networks has pursued two streams in its quest to simulate the activity of the human brain understanding the biological process of mental activity (the software) and developing the physical artificial neural network (the hardware). The study of human mental activity includes learning, memory, decision-making, calculating, imagination, creativity, motor skills, emotions, and sensory interfaces

The basic architecture of any artificial neural network includes input elements, processing elements called nodes, and output elements The variety of artificial neural systems results from the number of layers of processing nodes, the nature of the interconnections between the processing nodes and the input and output elements, and the method of 'training' the system utilized to learn the correct output

Artificial neural networks have been successfully used for many years for classification and pattern recognition, a basic human mental activity Artificial neural networks use associative memory, a stimulus-response process, to achieve their results Such systems compare input patterns to exemplar patterns existing in memory and generate the correct solution to the problem In general, artificial neural networks exhibit the ability to generalize the real pattern of information from less than ideal input, the ability to learn or modify their behavior in response to their environment, and the ability to abstract the essence of a set of inputs Such systems are robust, correctly classifying input after the destruction of random processing nodes, and exhibit plasticity or the ability to relearn after the destruction of a great many processing nodes A basic difference between traditional statistical procedures and artificial neural networks gives these systems a particular strength in pattern recognition Traditional statistical classification techniques assume an existing functional form from the outset of problem solution Artificial neural networks are more flexible, deriving the functional form from the nature of the input data

Kohonen (1984) proposed the definition for the artificial neural networks as "Artificial neural networks are massively parallel interconnected networks of simple (usually adaptive) elements and their hierarchical organizations which are intended to interact with the objects of the real world in the same way as the biological nervous systems do." By his definition, this paper would like to study the history of the artificial neural networks to propose the potential of the networks The research also want to figure out the biological origin of artificial neural networks to grasp the connection and resemblance of artificial neural networks to human brain This study can help researchers on artificial neural networks understand how the human brain work and what shape the artificial neural networks must be The author hope to give researchers the basic idea for further study on the area of artificial neural networks This study is a part of ongoing research on a current and future deployment of artificial neural networks.

II. THE HISTORY OF ARTIFICIAL NEURAL NETWORKS

In 1943, Warren McCulloch and Walter Pitts published the first paper on artificial neural network capabilities They suggested a model based on Boolean algebra which could be effective at classifying, storing and retrieving data, simple calculation, and recursive application of logic rules resulting in effective pattern recognition Their description of a network with an artificial neuron was a simplified representation of the biological neuron in the human brain. Using symbolic logic, McCulloch and Pitts proved that such networks, if supplemented with a large memory, were equivalent to A M Turing's universal computing machine, proposed in 1936. However, the McCulloch-Pitts nets were not always reliable at the specific logic tasks they performed. John von Neumann expanded the McCulloch-Pitts net to include several artificial neurons interacting with one another This redundant information representation made the artificial neural network more reliable in its pattern recognition capabilities The human memory utilizes this concept of redundant, distributed information This early research in artificial neural networks used direct mapping of specific interconnections and patterns in the systems to achieve successful matching of input to exemplar

In 1949, D. O Hebb proposed a biological learning law that became the basis for research on artificial neural network training algorithms. Hebb proposed that the interconnections within the brain are continually changing as an organism learns new tasks. He postulated that repeated neuron activation by another neuron causes the cells to develop stronger conductances and strengthen the cell assembly Artificial neural network research began to investigate the possibility that random interconnections between network nodes could be modified through some form of learning within the system itself Such network would be adaptive to new information

In the 1950s and 1960s, researchers combined the biological insight of Hebb with advanced technology to create the first adaptive artificial neural networks for pattern recognition. In 1958, Frank Rosenblatt developed the Perceptron The network consisted of a set of sensory units connected through a single layer of artificial neurons to a set of output units The artificial neuron nodes had weight values that could be adjusted or 'trained' by the researched to correct for inaccurate output At the same time, Widrow and Hoff (1960) devised a different 'supervised' training method from Rosenblatt's and incorporated it in their artificial neural network, the Adaline (adaptive linear neuron).

W K Taylor developed the first associative network in 1956 that employed the perceptron trained method The network could produce a complete output from less than perfect input by using 'winner-take-all' circuitry The exemplar that most closely matched the fragmented input took priority over all others Such networks are referred to as associative content addressable memories (ACAM)

In 1969, research in biologically based artificial neural networks was dealt a serious setback by researchers in logic based artificial intelligence In that year, Marvin Minsky and Seymore Papert (1969) published Perceptrons Using mathematical techniques, they proved that the Perceptron and Adaline could compute only simple logical problems More complex, logical functions, such as an 'exclusive or,' were impossible to calculate with the single layer of neurons employed in these networks Thus, these networks were limited in their pattern recognition capabilities This limitation can be overcome by adding additional layers of artificial neurons However, no training algorithms were available at the time to adjust the weight matrices for the additional 'hidden' layers Minsky and Papert (1969) went so far as to suggest that training hidden layers was impossible Funding for research in the artificial neural network field became unavailable for almost 15 years Many researchers abandoned the field for research in artificial intelligence

In the 1970s and 1980s, underfunded artificial neural network researchers continued work, developing the theoretical foundation for artificial neural networks D Marr proposed his theory of the cerebellum (1969), theory of the cerebral neocortex (1970), and theory of the hippocampus (1971) These theories compared the biological functioning of the neuron system in the human brain to the associative content addressable memory of the computer Marr suggested that biological neurons have specific functions based on their physical structure and location within the brain He theorized this neuronal specificity is linked to the representation of internal classifications in memory Marr's work provided a basis for the development of artificial neural network training algorithms for multilayered networks

The development of the backpropagation training algorithm, the result of three separate research studies (P. Werbos, 1974; D B Parker, 1982, D E Rumelhart, G E Hinton, and R J Williams 1986), provided a means of training multilayer networks Thus, the limitations of artificial neural networks presented by Minsky and Papert in 1969 were overcome Backpropagation is an unsupervised training method performed by the network itself It has proved successful with small problems, but has a scaling problem, showing reduced performance with large problems. D H Ballard (1987) suggested constructing auto-associative learning module feeding back from the output to the input and then organizing the modules in a hierarchical network to overcome this scaling problem An algorithm proposed by D. Willshaw and C Malsburg (1979) can also be used with a scaling problem Their algorithm eliminates all contacts between node with un-correlated activity. The backpropagation training algorithm then adjusts only the surviving weights.

In 1986, E Mjolsness and D H. Sharp introduced a training algorithm that modified the artificial neural network rules, not the nodal weights Their research stemmed from Holland's (1975) genetic algorithm studies Genetic algorithms scale better than backpropagation when dealing with large problems

By the late 1980s, many artificial neural network architectures and capabilities had been demonstrated Besides the multilayer, backpropagation network, there are many other architectures that have been developed in the last ten years to simulate human mental activity. Although some of these models are not realistic representations of the biological system, they are often conceptually stronger in their solutions to problems than many of the direct representation models. These networks include the Hopfield Net (1984), a recurrent system feeding back from the output layer to modify the input, the Boltzmann machine (Fahlman et al., 1983) that modifies its connectivity to overcome local minima problems, the Hamming Net (1986) based on communications theory, the Carpenter/Grossberg Classifier or Adaptive Resonance Theory (1986) which creates new exemplars when no pattern can be matched, the Kohonen Self-Organizing Feature Map (1988) which organizes the nodes into neighborhoods of similarity, the counterpropagation network of Hecht-Nielsen (1987), and the bidirectional associative memory. Recent advances in optical technology have expanded the architecture of artificial neural networks into novel areas The area of pattern recognition is basic to human intelligence New pattern recognition capabilities are being explored such as vision and speech. T J Sejnowski and C R Rosenberg developed a network capable of converting English text to phonetic representations and then to speech (NETtalk 1986). NETtalk generalizes human speech well and recovers satisfactorily if damaged, much like the human brain D J Burr developed an artificial neural network proficient at recognizing handwritten characters (1987)

In the past decade, the acceptance of artificial neural networks as a viable problem solving tool can be observed in the establishment of the first artificial neural network development company, Nestor, in 1986. Another seventy four companies were formed by 1986 Neurocomputing products reached the marketplace in that year By 1987, there were two newsletters network society. Artificial neural network development appears to be bounded only by the available technology with which to construct systems and the willingness of researchers to apply the systems to all ranges of classification problems.

III. THE BIOLOGICAL PROTOTYPE

As the human brain has been compared to a flexible analog processor with a large memory capacity, it is only natural that researchers in the computer field have sought to mimic its activity. The human brain is responsible for three processes it controls autonomomic functions, it interacts with the environment through the sensory systems, and it performs all mental activity The central and peripheral nervous systems are the information systems for the brain, providing sensory input and output for its perceptions, cognitions and consciousness The sensory systems include the olfactory nerve (smell), the optic system composed of four cranial nerves, the auditory vestibular nerve (hearing), the speech system composed of three cranial nerves, and the peripheral nervous system in the skin (touch)

1. Neurons -- The Building Blocks

The central and peripheral nervous systems are composed of an estimated 10 billion individual nerve cells or neurons which send and receive messages to each other The human brain alone contains, on average, 200 meters of nerve fiber per cubic millimeter of brain tissue Neurons, are functionally specific, gathering sensory information and providing motor reaction according to their shape and location in the nervous system As shown in Ffigure 1, nerve cells consist of a cell body with a plasma membrane, the soma, a protruding element for sending information, an axon, and tree-like branches from the soma to receive information, the dendrites The soma comprises less than 5 percent of the neuron's surface area. It manufactures the proteins necessary for chemical signal processing and manages the chemical economy of the cell The axon can be form .1 millimeter to a meter in length It contains storage bladders or synaptic vesicles of signal transmitting chemicals or neurotransmitters The dendrites and axon comprise 95 percent of the cell's surface area

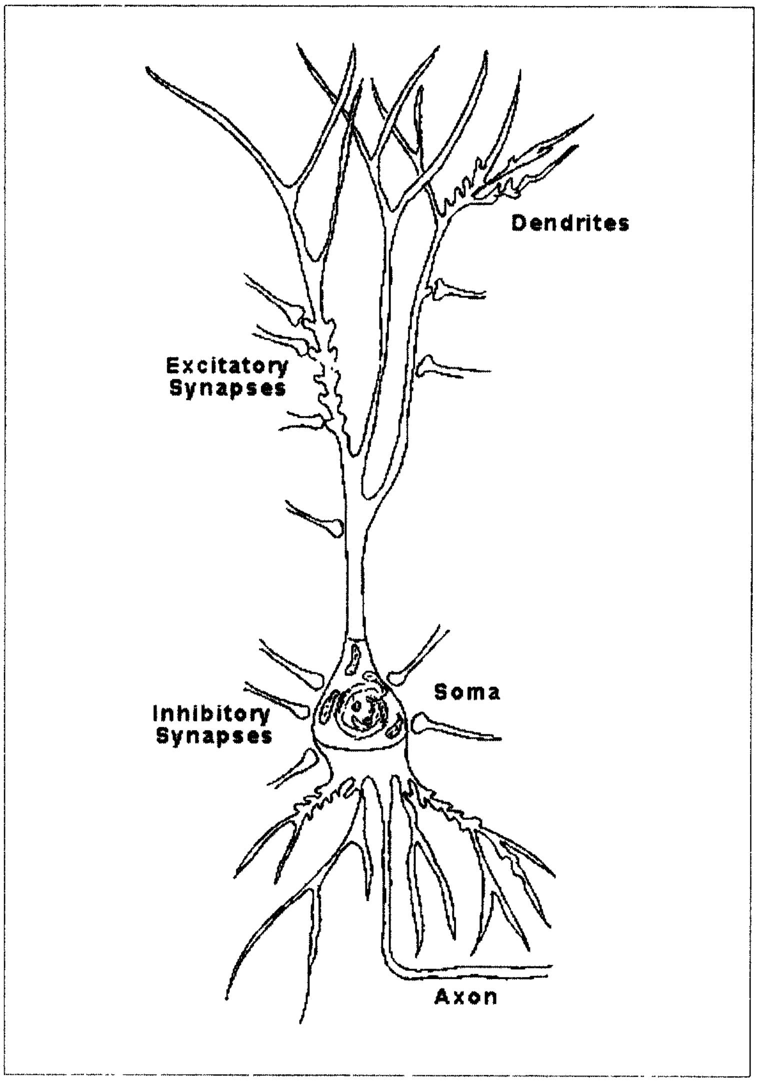

Figure 1. The Structure of Neurons

The neurons are interconnected in a variety of circuit architectures, which is shown in Figure 2. Sensory and motor pathways are constructed of long hierarchical circuits, with a local circuit neuron controlling the flow of information through the hierarchy. The hierarchical system contains both divergent circuits with one neuron transmitting to many other neurons, and convergent circuits with a single neuron receiving messages from many other neurons. Dendrites and axons are not physically connected Communication linkage is achieved at a synapse point when a neurotransmitter diffuses across a very narrow synaptic gap to a receptor site on the dendrite. Specific receptor sites exist for different neurotransmitters Over fifty different neurotransmitters have been identified in the human body, each inducing special functions in the nervous system This combination of synapse location and neurotransmitter type determines if the synapse is excitatory, communicating with other neurons, or inhibitory, not communicating with other neurons. It is believed that there are 10^{15} such interconnections between neurons in the human body This high degree of connectivity gives the human mind its computational power by allowing a parallel processing strategy

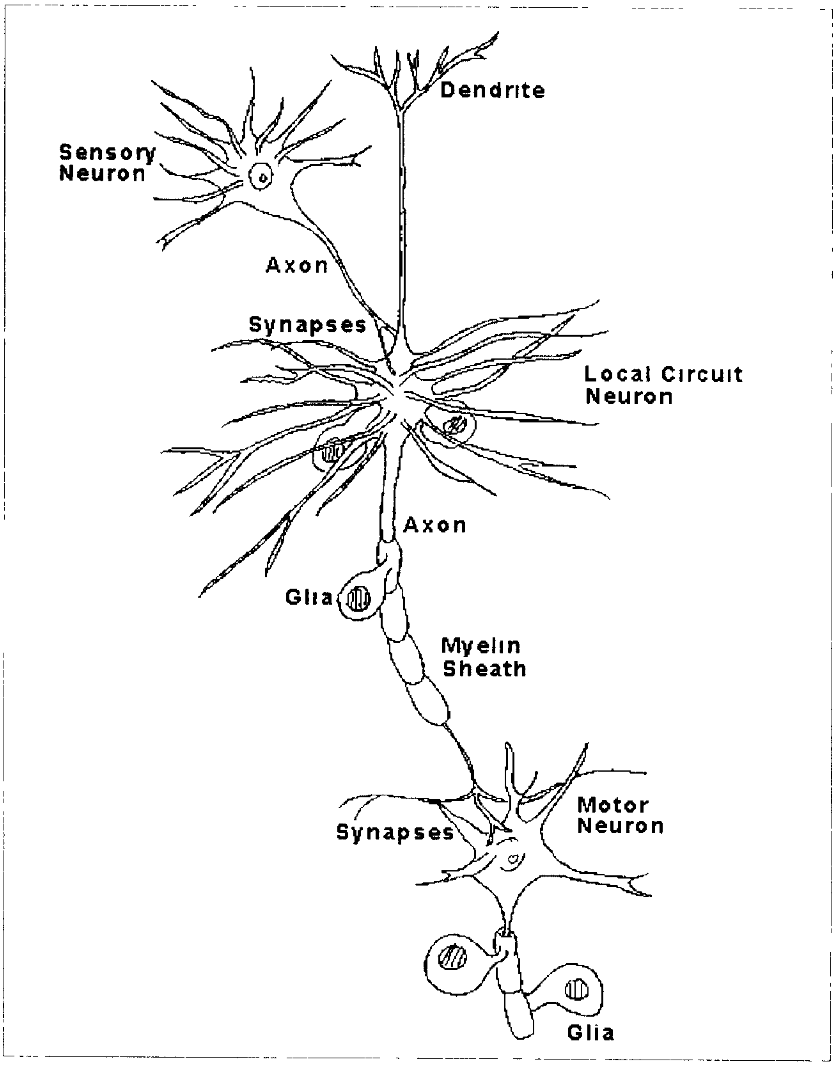

Figure 2 The Interconnections Among Neurons

2. Neuronal Activity

Neuronal activity results from two processes occurring in the cell membrane reaction to a chemical signal, and an electrical pulse modulating within the neuron. This activity is described in Figure 3. It is necessary to have a basic knowledge of chemistry to understand these neuronal activities Atoms are composed of a nucleus of protons or positively charged particles and neutrons or zero charged particles surrounded by shell of electrons or negatively charged particles. Atoms with few electrons in their outer shells try to give up the electrons and become positively charged ions Conversely, atoms with many electrons in their outer shells try to acquire additional electrons from other atoms and become negatively charged ions. This process is called reactivity

Neurons are surrounded by an equal distribution of positive ions in sodium and potassium, and negative ions in chlorine. Within the neuron, the soma manufactures five classes of proteins One class of protein serves as an ion pump on the cell membrane to maintain the electrical potential within the cell. another class works as an ion channel in the cell membrane selectively passing elements in and out of the soma. A third class of protein acts as a receptor for the different neurotransmitters. The last two classes of protein provide the physical structure to the cell and speed up chemical reactions. The structural protein has a negative ion charge of -70 millivolts of electrical potential (negative polarity). This negative polarity continually attempts to pull positive sodium and potassium ions into the cell. The plasma membrane surrounding the soma contains the protein channels to allow the sodium, potassium, and chlorine to pass in and out of the neuron. Normally these channels are closed or gated until some activity occurs to open them. As potassium passes through the plasma membrane quite easily, the membrane also contains protein based ion pumps to maintain the cell's equilibrium At rest, the neuron's soma has an internal potassium content ten times that of the external potassium level, and a sodium content ten times less than the external sodium level

Neurons that send messages to other neurons are presynaptic cells Those that receive messages are postsynaptic cells As excitatory electrical impulse travels through the axon of a presynaptic neuron, it releases a neurotransmitter from the synaptic vesicles The neurotransmitter travels across the synaptic gap between the neurons to the dendrites of the postsynaptic neuron The neurotransmitter may travel to one of two types of synapses an excitatory synapse or an inhibitory synapse The receipt of the neurotransmitter affects the electrical charge within the soma either positively or negatively The soma of the postsynaptic neuron has an electrical level or threshold of -50 millivolts to stimulate its axon to contact the next neuron in the circuit. This threshold value may require the summation or integration of the electrical potential of several presynaptic neurotransmitters to produce a change in neuron activity.

When a chemical signal is received by an excitatory synapse, the soma of the neuron becomes more positively charged than its normal negative resting state This triggers the membrane gates to allow sodium ions to enter and the cell interior depolarizes to +50 millivolts for one millisecond This charge propagates down the length of the axon until it reaches the synapse and triggers the release of a neurotransmitter to the next neuron. The signal continues through the nervous system as a propagating wave Almost simultaneously, there is an outward flow of potassium from the soma to help the cell return to its previous state or resting potential of -70 millivolts. The voltage difference between the inside of the neuron and the outside at this time has been measured at one tenth volt After firing, the ion pumps on the membrane begin to pump the sodium out of the cell body, and potassium returns to stabilize the neuron.

When a neurotransmitter is received by an inhibitory synapse, the inside of the neuron becomes more negatively charged The membrane channels allow negative chlorine ions to flow into the cell Potassium is forced outside the cell No depolarization occurs because the increase in negative internal charge inhibits the neuron from transmitting the signal any further. No other neurons are contacted

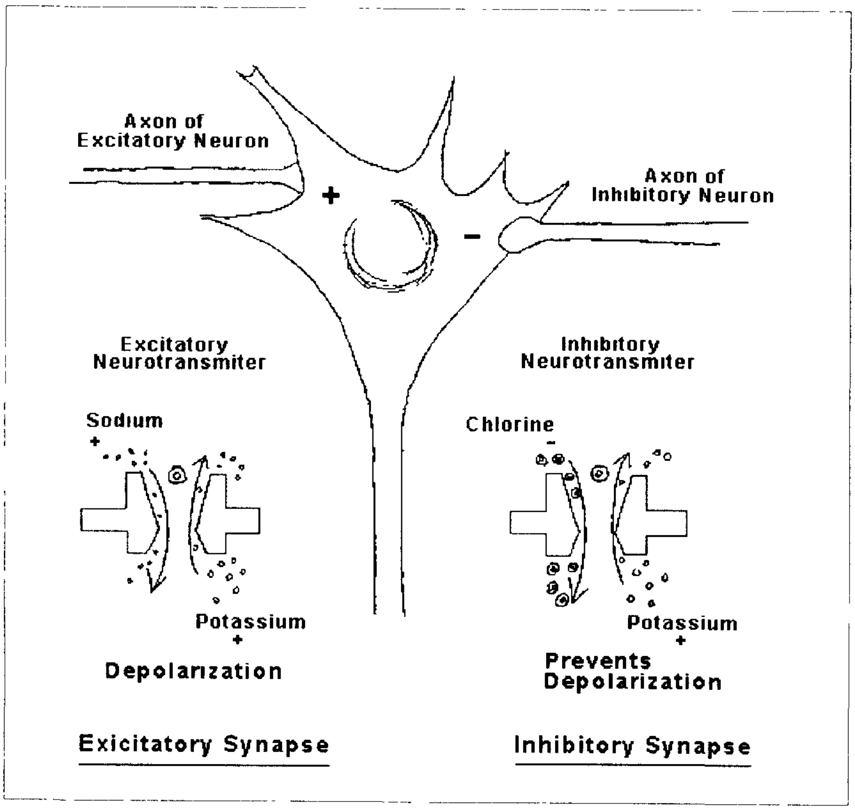

Figure 3 Neuronal Activity

The combination of chemical and electrical signals makes the single neuron slow in processing compared to conventional computers. Through the use of massive parallel processing between and among neurons, the brain is able to achieve its phenomenal processing speed. Each task is subdivided into many subtasks and processed concurrently.

3. Support Structures for Neurons

Spaces between the neurons and the circuitry of axons and dendrites are filled by glia cells These cells provide glucose to the nerve cells to produce energy to function. The glia are believed to be responsible for reorganizing the structure of neurons into various new circuity patterns when damage to the old patterns has occurred. More importantly, the glia clean away excess chemicals produced during the chemical signal transmission function By disposing of excess debris, the glia assure that no error in synaptic transmission occurs

Swann cells generate myelin. This chemical sheaths the axon much like the insulation on electrical wiring It constrains the electrical wave propagating down the length of the axon to prevent reduction in the electrical potential. The myelin sheath also works like the glia cell to guarantee no interference in signal transmission by insulating the axon from extraneous chemicals

The vascular system carries oxygenated blood to the brain to provide energy for its functioning. The brain occupies 2 percent of the body's mass, yet consumes 20 percent of the body's oxygen intake. A filtration system, the blood-brain barrier, exists between the brain and the main circulatory system. This barrier has low permeability to prevent chemical signals from the main circulatory system from affecting neuron transmissions in the brain.

The connective tissue of the brain, meninges, is filled with cerebrospinal fluid This fluid serves as a shock absorber for the neurons to prevent their damage Unlike other cells in the body, neurons cannot regenerate themselves. Nerve cells that are destroyed or damaged cannot be replaced Neurons do have the ability to grow and to form new synapses. This allows the brain to be extremely plastic, reorganizing undamaged neurons into new architectures and relearning previous knowledge.

4. Learn and Memory

Learning is a relatively permanent change in behavior that results from experience. It is an associative mental process of stimulus and response. In the biological learning process, a neuron is continually excited during the training phase The high level of nerve cell excitement generates a high level of positively charged calcium and sodium that activates certain enzymes in the soma of the cell. These enzymes attach phosphate molecules to cell proteins in a process called phosphorylation This process changes the character of the protein and the character of the ion channel in the cell membrane reducing the outflow of potassium The enzymes will continue to change the protein for a period of time after training stops These changes cause the synapses of the neuron to develop stronger conductances with the other neurons in the circuit This learning process affects memory

Memory is distributed throughout the brain cortex Short term memory is a process of limited duration neuron activation Thus, this memory has limited capacity for information manipulation and no changes occur in the neuron assemblies Long term memory is the result of physical, structural change in the cell assemblies of the nervous system Through repeated activation of neuron circuitry during the learning process, the synapses become functionally connected into a cell assembly Any activation of part of the cell assembly will access or modify long term memory These idea is parallel to Hebb's learning law, the foundation of artificial neural network training algorithms, proposed in 1949. This biological process of learning and memory has only been confirmed in the past decade

IV. CONCLUSION

Artificial neural networks are intended to respond to their environment in the same way as their biological counterparts Both artificial and biological neural networks are composed of a large number of simple elements reacting to an electrical signal They both perform computational functions, both process information in parallel, and both are capable of learning from experience As neuro-biological research increases our understanding of human mental and sensorimotor activity, it expands the theoretical base on which artificial neural network research is founded This, coupled with advances in technology, will allow researchers to develop artificial neural networks with extensive capabilities beyond the limitations of the human information processing systems and serial computer systems

REFERENCES

Anderson, J R, "Cognitive and Psychological Computation with Computation with Neural Models," IEEE Transactions in Systems, Man, and Cybernetics, 13, pp 799-815, 1983

Bloom, F E and Lazerson, R, Brain, Mind, and Behavior, New York, W H Freeman and Co, 1988

Bradford, H F, Chemical Neurobiology, New York, W H Freeman and Co, 1986

Changeux, J., Neuronal Man, Translated by Dr. Lawrence Garey, New York, Pantheon Books, 1985.

Cowan, J. D and Sharp, D. H., "Neural Nets," Quarterly Reviews of Biophysics, 21, pp 305-427, 1988.

Fahlman, S E., Hinton, G. E., and Sejnowski, T. J., "Massively Parallel Architectures for AI NETL, Thistle and Boltzmann Machines," Proceedings of te AAAI-83 Conference, 1983

Fogel, L. J., Human Information Processing, Englewood Cliffs, New Jersey., Prentice-Hall, Inc , 1967.

Hebb, D O , The Organization of Behavior, Wiley, New York, 1949

Hopfield, J J , "Neurons with Graded Response Have Collective Computational Properties like Those of Two-state Neurons," Proceedings of the National Academy of Sciences, Vol. 81, pp 3088-3092, 1984

Hawley, D D , Johnson, J. D and Raina, D., "Artificial neural Systems A New Tool for Financial Decision Making," Financial Analysts Journal, November/December, 1990

Hecht-Nielsen, R , "Counterpropagation Networks," Applied Optics, Vol. 26, No. 3, 4979-4984, Dec 1, 1987

Kohonen, T , Self-Organization and Associative Memory, New-York, Springer Verlag, 1984.

Kohonen, T , Self-Organization and Associative Memory, 2nd, Ed , Springer-Verlag, New York, 1988

Lambert, D , Bramwell, M and Lawther, G., Eds., The Brain A User's Manual, New York, Perigee Books, 1987

Lippmann, R. P., "An Introduction to Computing with Neural Nets", IEEE ASSP Magazine, 4, pp 4 22, 1987

McCulloch, W S , and Pitts, W , "A Logical Calculus of the Ideas Immanent in Nervous Activity, Bulletin of Mathematical Biophysics, Vol 5, pp 115-133, 1943.

Minsky, M L , and Papert S , Perceptrons An Introduction to Computational Geometry, M.I T. Press, Cambridge, MA, 1969; 2nd Eds., 1988.

Parker, D B , "Learning Logic," Invention Report, S81-64, File 1, Office of Technology Licensing, Standard University, Stanford, CA, 1982

Rumelhart, D. E , Hinton, G E , and Williams R. J., "Learning Representations by Back-Propagating Errors," Nature, 323, pp 533-536, 1986

Rosenblatt, F., "The Perceptron A Probabilistic Model for Information Storage and Organization in the Brain," Psychological Review, 65, 386-408, 1958

Sejnowski, T J., and Rosenberg, C R , "NETtalk A Parallel Network That Learns to Read Aloud," John Hopkins University Technical Report, 1, 1986.

Shepherd, G. M , The Synaptic Organization of the Brain, New York, Oxford University Press, 1979

Thompson, R F , The Brain An Introduction to Neurosciences, New York, W H Freeman and Co , 1985

Wasserman, P D , Neural Computing Theory and Practice, new York, Van Nostrand Reinhold, 1989

Wasserman P D and Schwartz, T , "Neural networks, part 2," IEEE Expert, pp 10-15, 1988.

Werbos, P J., "Beyond Regression New Tools for Prediction and Analysis in the Behavioral Sciences," Ph D Dissertation in Statistics, Harvard University, August, 1974

Widrow, B , and Hoff, M E , "Adaptive Switching Circuits," IRE WESCON Convention Record, Pt 4, pp 96-104, September 1960

Young, P., The Nature of Information, New York, Praeger Publishers, 1987