İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

A SYSTEMS LEVEL APPROACH TO PLANNING AND A COMPUTATIONAL MODEL OF THE TOWER OF HANOI TASK

M.Sc. Thesis By

Mete BALCI, B.Sc.

Department: Electronics and Communication Engineering

Programme: Electronics Engineering

JUNE 2007

İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

A SYSTEMS LEVEL APPROACH TO PLANNING AND A COMPUTATIONAL MODEL OF THE TOWER OF HANOI TASK

M.Sc. Thesis By

Mete BALCI

(504041213)

Date of submission: 3 May 2007
Date of defence examination: 18 June 2007

Supervisor (Chairman): Assistant Prof. Neslihan Serap Şengör

Members of the Examining Committee: Prof. İnci Çilesiz (İ.T.Ü.)

Associate Prof. Hakan Gürvit (İ.Ü.)

JUNE 2007

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

PLANLAMAYA SİSTEM SEVİYESİNDE BİR BAKIŞ VE HANOİ NİN KULELERİ PROBLEMİ İÇİN BİR HESAPLAMALI MODEL

YÜKSEK LİSANS TEZİ

Mete BALCI

(504041213)

Tezin Enstitüye Verildiği Tarih: 3 Mayıs 2007
Tezin Savunulduğu Tarih: 18 Haziran 2007

Tez Danışmanı: **Yrd.Doç.Dr. Neslihan Serap Şengör**

Diğer Jüri Üyeleri: **Prof.Dr. İnci Çilesiz (İ.T.Ü.)**

Doç.Dr. Hakan Gürvit (İ.Ü.)

HAZİRAN 2007

FOREWORD

This thesis was carried out in close collaboration with my supervisor Neslihan Serap Şengör over the last two and a half years. I would like to thank her for introducing me the field, giving advices, sharing her opinions and listening to mine. It is certain that this work will not be possible without her generous support.

I would like to thank Fatih İşbecer, my manager at Pozitron, for his unfailing support which has enabled me to complete this work.

In the first year of this research, I have attended to Behavioral Neurology course given by Hakan Gürvit at Istanbul University. I would like to thank him for giving me this opportunity and kindly supporting anyone interested in the subject.

It was a great pleasure for me to meet Yuichi Sakumura and Ricardo Carnieri at Okinawa Computational Neuroscience Course in 2006. I would like to thank Yuichi Sakumura for his ongoing support and friendship. I would also like to thank Ricardo Carnieri for his friendship and the technical talks we had during the hours we spent on the project work.

Finally, I would like to dedicate this thesis to my family. They are going to be happier than me seeing this work finished.

May 2007 **Mete Balcı**

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page No

PLANLAMAYA SİSTEM SEVİYESİNDE BİR BAKIŞ VE
HANOİ NİN KULELERİ PROBLEMİ İÇİN BİR HESAPLAMALI MODEL

ÖZET

Hayvanların önemli bir özelliği değişen bir ortamda uygun hareketleri amaca yönelik şekilde seçebilmeleridir. Amaca yönelik davranış, karar verme, planlama ve problem çözme gibi konuları içeren geniş bir terimdir. Bütün bunların ortak noktası ise beynin özellikle korteksin işlevleri ile ortaya çıkan üst düzey bilişsel işlevlere bağlı olmalarıdır. Üst düzey bilişsel işlevlere ilişkin detaylar, beynin fiziksel yapısı ve sinirsel bağlantılar ile sinir ağları ve sinirsel gösterime ilişkin hesaplamalı temeller anlaşılmadığı için tam olarak bilinmemektedir.

Bu çalışmanın amacı beynin planlamaya ilişkin olan mimarisini denek davranışları ve işlevsel anatomi ışığında anlamaktır. Bu sebeple, biyofiziksel ve hücresel yapılar özellikle düşünülmemiş, denek davranışları, beynin işlevsel mimarisi ve bunların getirdiği hesaplamalı sonuçlar göz önüne alınmıştır.

Bu çalışma amaca yönelik davranışın bir elemanı olarak düşünülen planlamanın bilişsel yönlerini ele almaktadır. Amaç, planlamayı tanımlamak ve planlama için bir hesaplamalı model önermektir. Bilişsel bilim, sinir sistemi, ilişkili öğrenme ve hücre, ağ ve öğrenmeyi içeren hesaplamalı ve teorik bulgular özetlenerek tanıtıldıktan sonra nöropsikoloji ve hesaplamalı temeller göz önüne alınarak planlamanın bir yorumu yapılmıştır. Son olarak, deneklerin planlama ödevlerindeki davranışlarını incelemek üzere planlama için bir yapı ve bu yapıdaki yürütücü işlevler için bir hesaplamalı model önerilmiştir. Modelin belirgin özellikleri, klinik çalışmalarda ve nöropsikoloji literatüründe yaygın olarak planlama ödevi olarak kullanılan Hanoi 'nin Kuleleri problemi için gerçeklenerek incelenmiştir.

Varolan literatüre şu katkılar yapılmıştır; (1) planlama sırasındaki bilgi akışını anlamak için açık bir model önerilmiştir, (2) simülasyon sonuçları ve deneklerin davranışları karşılaştırıldığında, çalışma belleğinin planlama sürecini açıklamak için tek başına yeterli olmadığı anlaşılmıştır, (3) çalışma belleği kapasitesi azaldığında, bunun modelin davranışı üzerindeki etkisinin azaldığı, diğer taraftan değerlendirme işlevinin etkisinin arttığı görülmüştür, (4) planlama ödevlerinin skorlanmasına ilişkin karşıt bir argüman belirtilmiştir. Bu bilgilerin ışığında, prefrontal hasarların, çalışma belleği dışında, değerlendirme sürecine olan etkisinin de planlama sürecinde etkili olduğu görülmüştür.

A SYSTEMS LEVEL APPROACH TO PLANNING AND A COMPUTATIONAL MODEL OF THE TOWER OF HANOI TASK

SUMMARY

An important ability of animals is to choose appropriate actions in a goal directed manner in a changing environment. Goal directed behaviour is a broad term covering decision making, planning, problem solving and so on. A common point of all these is that all of them depends on high level cognition which is an integrated function of brain especially neocortex. Details of high level cognitive functions are not understood because physical structure, neural connectivity and computational principles of neural representations and network structures are not well-known.

The purpose of this study is to understand the architecture of the brain related to planning based on the behavioral performance of subjects and functional anatomy of the brain. In this view, biophysical and cellular architecture is not considered, but the behavior of subjects, functional architecture of the brain and computational implications are taken into account.

This study focuses on cognitive aspects of planning which is considered to be a component of goal directed behavior. The aim is to set a definition and propose a computational model for planning. After cognitive science is introduced and preliminary information about basics of nervous system, associative learning and computational and theoretical findings including the models of single neurons, networks and learning are given, an interpretation of planning is established considering neuropsychological facts and computational principles of problem solving. Finally, a framework for cognitive planning and a computational model for executive control in this framework, based on the interpretation given, is presented in order to give an explanation for the behavioral patterns of subjects in cognitive planning tasks. The characteristics of the model is investigated, since it is implemented and simulated for the Tower of Hanoi (ToH) task, which is widely used as a planning task in neuropsychology literature as well as clinical studies.

The progress over existing literature is that; (1) an explicit model to understand information flow during planning is proposed, (2) based on comparative analysis between the simulation results and the performance data of subjects, it is found that the working memory capacity alone is not a sufficient factor to explain the behavior for planning e.g. in the Tower of Hanoi task, (3) when the capacity of working memory decreases, its effect on the model's behavior is also decreased which, in turn, increases the importance of evaluation process, (4) a counter-argument is proposed for scoring the planning tasks since the cognitive ability of subjects are unknown.

1. INTRODUCTION

Although it has its roots centuries before, cognitive science research has gained an enormous acceleration throughout the 20th century. Together with the behavioral studies and the advancements in experimental methods and biomedical technology, which enabled invasive and non-invasive measurement and imaging techniques that were beyond the dreams of any scientist earlier, cognitive science literature has been exploded in the second half of the 20th century. Theoretical and computational methods, which are relatively new to the field, are latest popular topics, and they have already been supported by some leading scientists in the field.

1.1 What is Cognitive Science ?

1.1.1 Origins

Throughout history, many philosophers and early scientists have been interested in understanding the nature of mind and there has been a long-running debate over the physical location of mind. Although the study of brain has a long history, the role of brain is only recently understood. Many centuries before, the heart was believed to be the home of the mind by Egyptians. In the tomb of Tutankhamen (about 1300 BC), together with his mummy, four important organs; liver, lungs, stomach and intestines are preserved in jars while the heart is preserved in its place within the body. In order to have afterlife, these organs are considered to be necessary, whereas the brain is discarded [1,2].

Around 400 BC, in ancient Greece, Hippocrates and Plato speculated that the actions of human are triggered by the brain and the brain is the locus of mind. However, Aristotle believed that the heart is the organ for cognition and the brain is important for regulating the body temperature. In the 2th century, Galen reported behavioral changes of gladiators caused by injuries to the heads, however, his ideas about the anatomy of the human brain was very inaccurate since he was refusing to do postmortem work on human bodies [1,2].

During the middle ages, the Christian church was the absolute authority of all intellectual thinking and again other organs like liver and heart was included in the

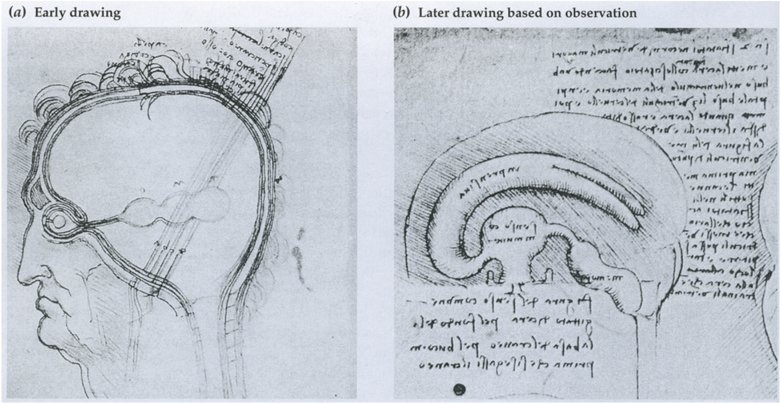

Figure 1.1: Leonardo da Vinci's drawings of the brain showing cerebral ventricles [2].

explanation of mind and actions or behaviors based on Nemesius' theory which localizes the mental ability to the ventricles in the brain. In Renaissance period, Leonardo da Vinci studied the human body including the brain (Figure 1.1) and in the 16^{th} century, Vesalius rejected the ventricular localization view. In the 17^{th} century, Descartes introduced the notion of body and mind as two separate entities (material body and nonmaterial soul) which is called mind-body dualism and as shown in Figure 1.2 by his own drawing, he proposed that the mind communicates with the body via pineal gland. Although it is not accurate in today's terms, Descartes also coined the term "Reflex" which means the reflection of external stimulation by the muscles.(Figure 1.3) In late 18^{th} century, Galvani has made experiments by stimulating a frog's leg electrically. Galvani has realized that after removing the leg from the frog's body, by electrical stimulation, the muscles in the leg still responds. This observation was against Descartes' theory, since any connection to pineal gland is not available, and the electricity is suggested as an essential element of functioning of the nervous system. Although Descartes' theory for interaction between mind and brain was put on question by Galvani's observations, his view which speculates the brain as the center of mind is confirmed by Flourens in 19^{th} century. Flourens surgically removed parts of the brain of a number of animal species and he observed if there is a change in animal's behavior after recovery. Based on observations, he concluded that the cerebrum controls voluntary movements and perception. In 18^{th} and 19^{th} centuries, Whytt, Bell, Magendie and Müller studied the nervous system including spinal cord and provided new insights to sensory and motor systems [1–3].

In the 19^{th} century, introduced by Gall, the cerebral cortex of the brain is considered to be composed of separate functional areas and this notion is called phrenology. According to phrenology, each area is responsible for a cognitive ability and as shown

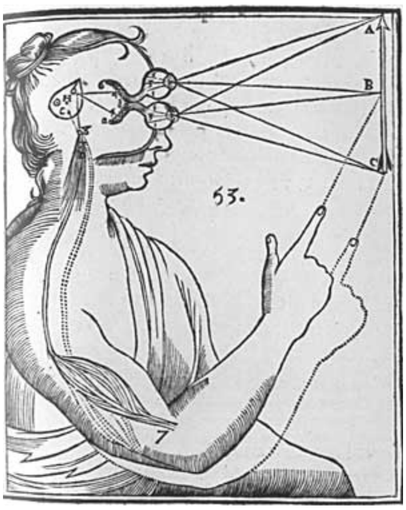

Figure 1.2: A drawing from the book De Homine by Descartes, published in 1662. In the center of the brain, pineal gland is shown.

Figure 1.3: An early view of reflexes, a drawing by Descartes [2].

in Figure 1.4 maps of the brain are created. However, many rejected this view while insisting that the brain functions as a whole. Today, it is known that some areas have peak activity when subject is performing a specific task, however, a single part is not responsible for a function alone. An important study is given by Broca in 19^{th} century, where he studied a postmortem brain of a patient who had been unable to talk. He discovered that a specific region in the brain of the patient, which is called Broca's area today, was damaged. Around the same years, Wernicke identified another region for language, which is called Wernicke's area today. A subject having damaged Broca's area cannot speak but understand the speech of others. However, a damage to

Wernicke's area does not affect the ability to speak but the speech is meaningless and the subject cannot understand the speech of others [1,2].

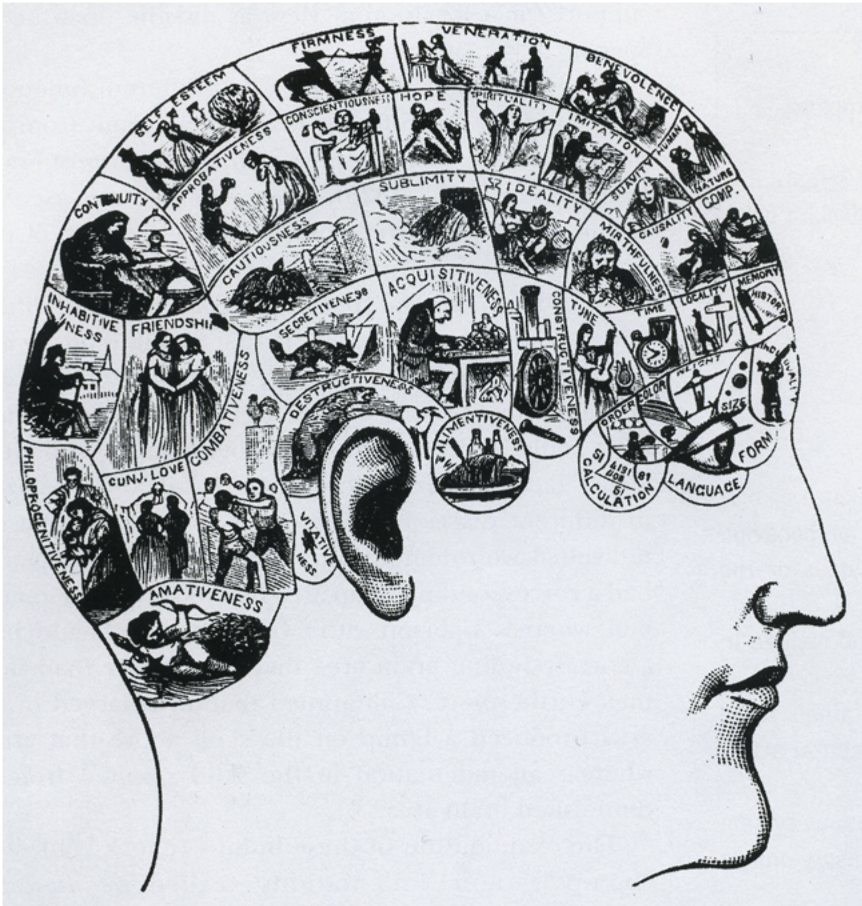

Figure 1.4: An example of phrenology map [1].

In the late 1890s, a milestone research in neuroscience has been done by Golgi and Cajal. Golgi has invented the silver staining procedure to reveal the structure of a single neuron. In this method, the tissue is first exposed to potassium dichromate and then to silver nitrate. The chemical reaction between these two chemicals creates silver chromate and silver chromate enters into a few cells of the tissue. Then, the tissue can be analyzed since the cells invaded by silver chromate looks black in the microscope. This technique was used by Cajal leading to the neuron doctrine which states that the brain is composed of independent functional units called neurons and the information is transferred between neurons across tiny gaps.(Figure 1.5) Golgi and Cajal shared the Nobel Prize in 1906 for their work on cellular properties of neurons. In following years, Sherrington studied the spinal reflex to understand the information transfer between neurons and called the tiny gaps, in the neuron doctrine, synapses which means binding (-apse) and together (syn-). In the same decade, Loewi discovered that communication between neurons is a chemical process. Sherrington and Loewi shared the Nobel Prize in 1932 and 1936 respectively [1,2,4,5].

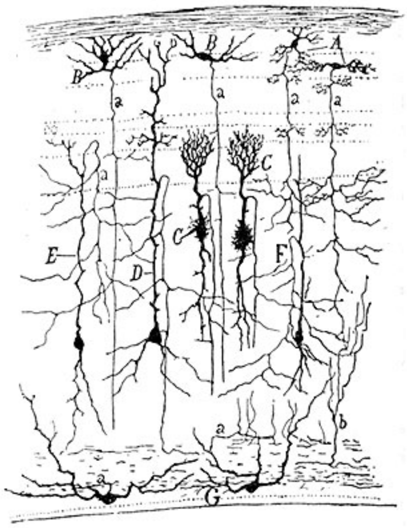

Figure 1.5: Drawing of a section through the optic tectum of a sparrow. By Santiago Ramon y Cajal, c.1900 [5].

At the end of 19$^{\text{th}}$ century and in the early 20$^{\text{th}}$ century, the study of mind has also moved to the domain of experimental psychology and behavioral studies which are leaded by Thorndike and Pavlov who has studied the learning and memory in animal subjects, have been started. Later, experimental psychology became dominated by Watson's behaviorist viewpoint that virtually denies the existence of mind. Due to this view, many studies in psychology have been restricted to examine the relation between stimulus and response until 1950s when intellectual landscape began to change. Starting by Miller's study on mental representations and procedures for decoding and encoding information, conceptual theories of mind based on complex representations and computational procedures appeared [1,2,4,5].

Pioneered by Minsky, McCarthy, Rosenblatt, Newell, Simon and others, the field of artificial intelligence had been established in 1950s. After the initial studies in AI have been completed, the foremost difficulty of learning systems was the credit-assignment problem which has been explained later in this study. Most systems were only capable of solving simple problems. As the difficulties became apparent, the interest and research in AI shifted to the fields of pattern classification and function optimization [6].

1.1.2 Definition

In the Oxford English Dictionary, the entry for the word “cognitive” is “of or relating to cognition”. Moreover, the word “cognition” is originated from latin “cognoscere” meaning “get to know” and cognition is defined as “the mental acquisition of knowledge through thought, experience, and the senses” [7].

In his commentary on Lighthill Report in 1973, Christopher Longuet-Higgins, a British scientist who have made important contributions to chemistry and artificial intelligence, coined the term “cognitive science”, as he is unsure what to call the sort of interdisciplinary research he conducted [8]. In the same decade, the Journal of Cognitive Science has began and the Cognitive Science Society was formed [4,9]. Cognitive science is the interdisciplinary study of cognition. The fields closely related to cognitive science are philosophy, psychology, neuroscience, linguistics, anthropology and artificial intelligence [9].

In general, cognitive science research is being done in three levels to give answers to following questions:

- Behavioral Level: Given an input (stimulus), what is the output (behavioral response) ?
- Functional Level: How is the output produced ?
- Physical Level: What produces the output ?

This is similar to the level of analysis suggested by Marr who introduced the notion of computational, algorithmic and implementation levels [10]. According to Marr and Poggio, central nervous system can be understood in four levels. First, computation; second, implementation of the computation by an algorithm; third, commitment of the algorithm to a mechanism; and fourth, realization of the mechanism in hardware [11].

An analogy is often used to compare the brain to a digital computer, where the physical level is the computer hardware, the behavioral level is the software, and the functional level is shared by hardware and software. However, as Churchland claims, brain is like a computer in an abstract sense, where both of them produces an output resulted from internal processing of a given input. As digital computers are only special instances of computation, “not as the defining archetype”, organism can also be regarded as a different form of computational unit where both have only basic principles in common. In addition, it is not known whether all the activities in the brain can be defined in all terms as computable functions [12,13].

Another view considers the brain as a highly parallel computer. However, parallel computation, as a term of current technology, means computing by dividing the data into multiple processes. The actual algorithms used in parallel computers are based on the algorithms used in serial computation and parallel computers are actually just a number of serial computers. On the other hand, although neural processing in the brain is slow due to biological and chemical constraints, the brain completes the computation in much less steps than digital computers.(Figure 1.6) In this sense, it is argued that today's digital computers cannot have a performance similar to brain in any real world task, so, paradigms in building machines of computation should be reconstructed based on the brain architecture [14].

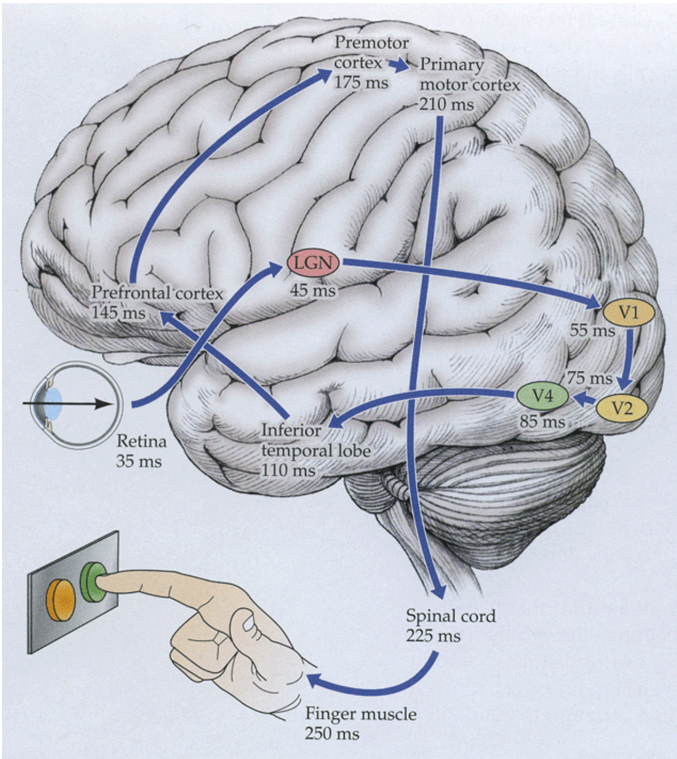

Figure 1.6: Information flow with approximate timings in the brain [2].

In 1973, in his influential report that has almost ended all of the AI research in the UK, Sir James Lighthill classified the AI research into the following three categories:

- Category A: Advanced automation or applications.
- Category B: Building robots and bridge between the other two categories. Robots mentioned here exclude unimates, e.g. industrial robots.
- Category C: Studies of the central nervous system including computer modeling

in support of both neurophysiology and psychology.

As the definition of category C points out, cognitive science and AI is quite related [15].

1.1.3 Research Methods

Being highly interdisciplinary, a wide range of research methods are used in cognitive science in different levels of study. One of them is behavioral experiments which are closely related to experimental and cognitive psychology, by which the behavioral responses of the subjects to the controlled stimuli are observed. By measuring the properties of behavioral responses such as reaction time (RT), accuracy, psychophysical judgements and eye tracking, internal processing of stimuli is being studied. Behavioral experiments provide high-level understanding of the brain [5].

Neuroimaging, being a non-invasive method, has enabled visualizing the anatomy and activity of the brain in vivo.(Figure 1.7) Specifically, fMRI (Functional Magnetic Resonance Imaging) has become a dominant research method since it gives the information gluing behavior and neural activity in the brain. fMRI measures the relative amount of oxygenated blood flowing in the brain. More oxygenated blood in a particular region in the brain is correlated with an increase in neural activity in that part. Thus, fMRI reveals the functional structures within the brain while subjects are performing controlled tasks. This makes analyzing the information flow in the brain possible. fMRI has moderate spatial and temporal resolution [2,5].

Another method called EEG (ElectroEncephaloGraphy), which is introduced in late 19^{th} century, measures the electrical activity generated by large population of neurons by electrodes placed on the scalp (in special cases electrodes are also placed inside the cortex). It has a very high temporal resolution, however, since the signal represents gross activity of many sources, only poor spatial resolution can be obtained. Although it is an ill-posed problem as there are more than one possible configuration of sources contributing to the activity measured, source localization methods with additional constraints are being used to localize the actual sources of the signals gathered by EEG electrodes in order to improve the spatial resolution. It is also possible to record the activity of a single neuron or a group of neurons [5].

TMS (Transcranial Magnetic Stimulation), which can excite neurons by inducing current in the brain using magnetic fields, is a non-invasive method. Its principles are actually known since 19^{th} century, however, the first successful study has been made by Barker et al. in 1985 [5]. It is a powerful method since a part of brain can be made work more or less for a period of time in vivo. Previously, such work can

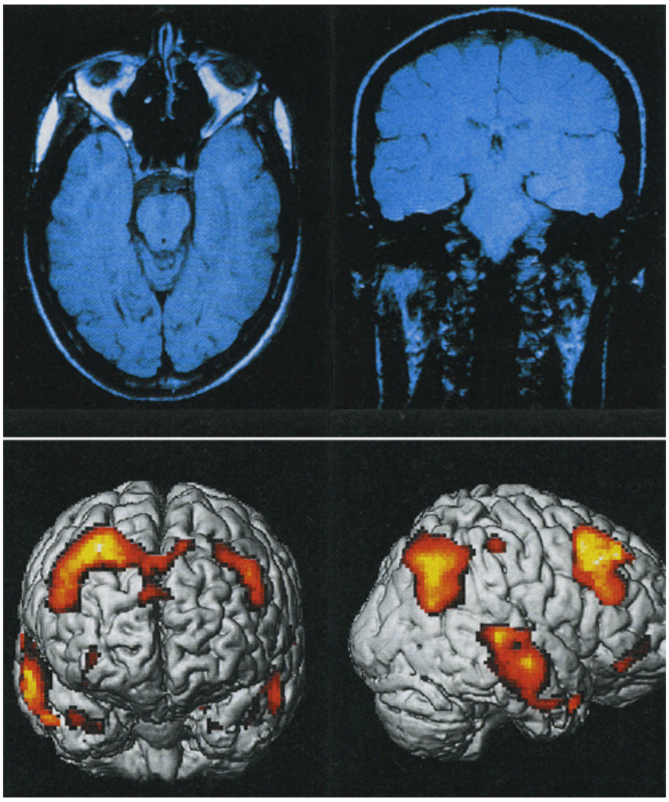

Figure 1.7: A sample of MRI (upper) and fMRI (lower) images [2].

only be done by chemical agents or genetic engineering on animal experiments with difficulties. TMS enables such studies to be done with human subjects accurately and it is believed to be without any risk [16].

An important method to study brain is to study the consequences of natural lesions or to observe the subject by removing a specific neural tissue (ablation). As in the case of Broca's study, lesions in human subjects can provide insight into the underlying neural architecture behind a mental ability. In order to investigate the brain further, a specific neural tissue can be removed or damaged in animals and the consequences of this damage can be observed. Moreover, the ablation can be done temporarily by using chemicals or electrical stimulation employing transplanted electrodes in the animal's brain [1].

In the laboratory, the methods like staining can be used to reveal the various aspects of cells. Golgi staining provides outline of the whole cell, whereas Nissl staining outline the cell bodies. Also, radioactively marked chemicals can be injected into the nervous system and the targeted neural structure can be analyzed by imaging methods after the tissue is removed. Furthermore, genetically engineered animals (usually rats) are being used to see how a specific element in cellular level affects the neural processing [2].

Computational modeling is one of the most recent approaches in cognitive science where models are often used to simulate specific aspects of intelligence to help understanding the organization of the brain for different purposes. Models can be constructed based on principles of symbolic processing, artificial neural networks or dynamic systems [5]. According to Dayan and Abbott [17], models can be categorized as follows:

- Descriptive models characterize what neurons and neural circuit do by summarizing the experimental data in a compact and accurate way. Their purpose is to describe to phenomena, thus, the descriptive models are loosely based on the realistic biophysical, anatomical and physiological findings.
- Mechanistic models explain how nervous systems operate taking into account the biophysical, anatomical and physiological findings.
- Interpretive models address the question of why nervous systems operate as they do, while exploring the behavioral and cognitive significance of nervous system function by using computational and information-theoretic principles.

1.2 Perspective

Since it is consistently increasing its popularity year by year, which results entering of more researcher from a number of fields to the cognitive science, the methods and the conclusions differ much more than before especially when some form of computation is employed. In AI and psychology, computation is usually used to construct cognitive models to study behavior, whereas, in neuroscience, computation has been applied in a broader spectrum. Since digital computers are getting much powerful, doubling their speed each 18 months (this is known as Moore's Law, named after Gordon E. Moore.), computational methods are applied in any level, even to model chemical reactions in the synapse at molecular level [18]. However, the application of computational methods differ quite much even among the scientists working on the same subject. Computation is considered more abstract in AI and psychology, whereas it is usually considered as a tool in neuroscience like a tool or a device in the laboratory. Therefore, what people understand from computational methods in cognitive science is very broad. In a session for computational methods in a cognitive science meeting, you may find studies based on signal analysis employing wavelet transforms, 3D reconstruction, approximations to the inverse problems, data fitting, future state prediction, system identification, pattern recognition, knowledge representation, production systems, neural networks, dynamical systems and so on.

In my point of view, computation can be used as a complementary setup to experimental studies in order to test and refine hypothesis about the mind. Understanding the underlying mechanisms of brain is not possible by purely bottom-up or top-down strategy. Trying to explain high-level phenomena by purely neurobiological facts does not work, so, a cognitive researcher should switch between levels. Computational models must provide a complete framework for a researcher to switch between levels, since the emergent properties of such a system cannot be constructed by individual parts as those properties are not the properties of parts but a dynamical combination of the system. A study in neurobiology by Allen Selverston (1988) is a well-known case for this dilemma, where all 28 neurons of stomatogastric ganglion of the spiny lobster is well understood, but the source of its rhythmic behavior is not known [13].

An important difference in the studies between natural sciences and cognitive science is that it is not always possible to make an experiment in cognitive science. It is even impossible when complex and large scale brain networks are taken into account. Understanding the basic behavioral and anatomical studies as well as clinical data guides us to set a hypothesis about the underlying functional network. Then, computational methods enable us to conduct experiments similar to the ones in behavioral studies. The results are two-fold; first, verification data is used to refine hypothesis and models, second, results give further clues about how to modify experiments to gather more useful information and what to expect as a result of behavioral studies.

Newell pointed out that psychology being focused on isolated issues would never lead us to understand the mind without developing integrated theories. The models of intelligence should be demonstrated in the same range of domain and tasks as humans can handle, and they should be evaluated in terms of generality and flexibility rather than success on a single application domain. Cognitive models are general but less powerful comparing to models in AI, this should change by findings from cognitive science research [19].

1.3 Orientation

This study focuses on cognitive aspects of planning which is considered to be a component of goal directed behavior. The aim is to set a definition and propose a computational model for cognitive planning. First, cognitive science is introduced and preliminary information, which is studied during this work, about basics of nervous system, associative learning and computational and theoretical findings including the models of single neurons, networks and learning are given. After that, an interpretation of planning is established considering neuropsychological facts and computational principles of problem solving. In the end, a framework for cognitive planning and a computational model for executive control in this framework, based on the interpretation given, is presented in order to give an explanation for the behavioral patterns of subjects in cognitive planning tasks. The characteristics of the model is investigated, since it is implemented and simulated for the Tower of Hanoi (ToH) task, which is widely used as a planning task in neuropsychology literature as well as clinical studies. Also, the simulation results are compared with other two models in the literature and experimental findings.

In summary, this study (1) introduces cognitive science, (2) establishes an interpretation of planning based on neuropsychological and computational findings, (3) proposes a novel model to solve the Tower of Hanoi task and (4) summarizes and discusses results with the existing studies in the literature.

As the attempt is to model behavioral performance, biophysical details of real neuronal functions are not considered in this study. Instead, neuronal substrates that are most relevant to planning are taken into account at behavioral level in order to understand the coordination and cooperation of the executive functions including memory processes during planning.

The model parameters are not optimized or fit to experimental data in order to compare temporal properties of responses between simulation results and clinical studies because, as I also believe, the real functional architecture of brain, due to evolution and maybe other reasons, has many other biological and structural features [13], that is not included and some are beyond any simulation [20], that affects the transient aspects including the temporal properties of responses during the task being evaluated. Instead, structural properties of behavioral responses are compared to other computational models as well as experimental studies.

The purpose of this study is to understand the architecture of the brain related to planning based on the behavioral performance of subjects and functional anatomy of the brain. In this view, biophysical and cellular architecture is not considered, but the behavior of subjects, functional architecture of the brain and computational implications are taken into account. Thus, the explanations are given in terms of not mechanistic but descriptive and interpretive terms. In that sense, it sounds similar to early works on artificial intelligence, whereas the crucial link between AI and understanding human cognition has been broken over the last decades [21]. Thus, this study mainly focuses on the computational and neuropsychological aspects of high-level cognition during planning.

2. PRELIMINARIES

2.1 Nervous System

2.1.1 Neuron

In contrast to reticular theory, according to the neuron doctrine introduced by Cajal in the late 19th century, neurons are the basic, independent and discrete structural and functional units of nervous system. A neuron shown in Figure 2.1, which is an electrically excitable cell in the nervous system, has a cell body (soma), an axon, and a dendritic tree. As in the case of purkinje cells (Figure 2.2), a neuron can have more than a thousand dendritic branches making connections with tens of thousands of cells. A neuron simply processes input it receives from dendtritic tree and from soma and sends output via axon by propagating electrical signals called action potentials. Three principle types of neurons are shown in Figure 2.3.

The membrane of a neuron is a lipid bilayer, 3-4nm in width, and like other cells, it keep some molecules out and let some molecules pass through. It is constructed as a lipid bilayer with embedded proteins some of which also have channels that passes a specific type of ions selectively. In contrast to other cells, the potential inside a neuron is more negative than extracellular environment when resting. Thus, the cell is said to be polarized. In addition to negatively charged large proteins which cannot go outside the cell, there are three mechanisms responsible for the cell being polarized; concentration gradient, electrostatic pressure and sodium-potassium pump. Concentration gradient refers to diffusion of ions from regions of high concentration into regions of low concentration. Electrostatic pressure is caused by distribution of electrical charges. The rule “like charges repel and opposite charges attract” applies to ions inside and outside of the cell. Since the cell is negatively charged and the cell membrane is selectively permeable to potassium, sodium and chloride ions, the potassium and sodium ions are pulled inside, whereas chloride ions are pushed outside. However, potassium concentration is higher in the cell, thus, potassium ions diffuses outside. Thus, concentration gradient and electrostatic pressure reach equilibrium for potassium ions. On the other hand, the permeability of membrane to sodium ions is much less than its permeability to potassium ions. Thus, sodium concentration is

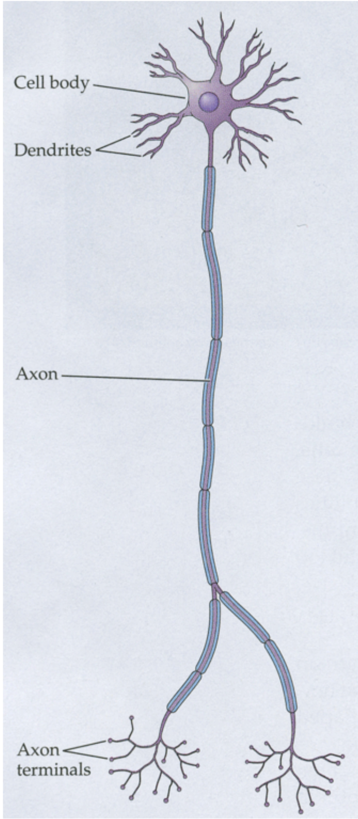

Figure 2.1: Major parts of a neuron [2].

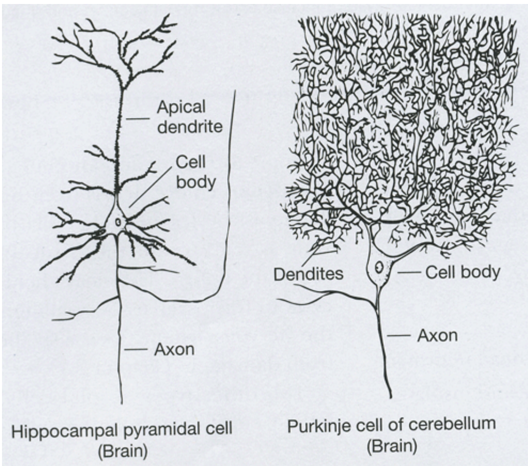

Figure 2.2: A drawing of pyramidal and purkinje neurons [1].

lower in the cell and sodium ions leak into the cell. If this leakage is not recovered, the membrane potential would eventually reach to zero. However, the sodium-potassium pump actively pushes three sodium ions out and pulls two potassium ions in. In this fashion, sodium leakage does not affect the resting potential. Since the sodium-potassium pump works against both concentration gradient and electrostatic pressure,

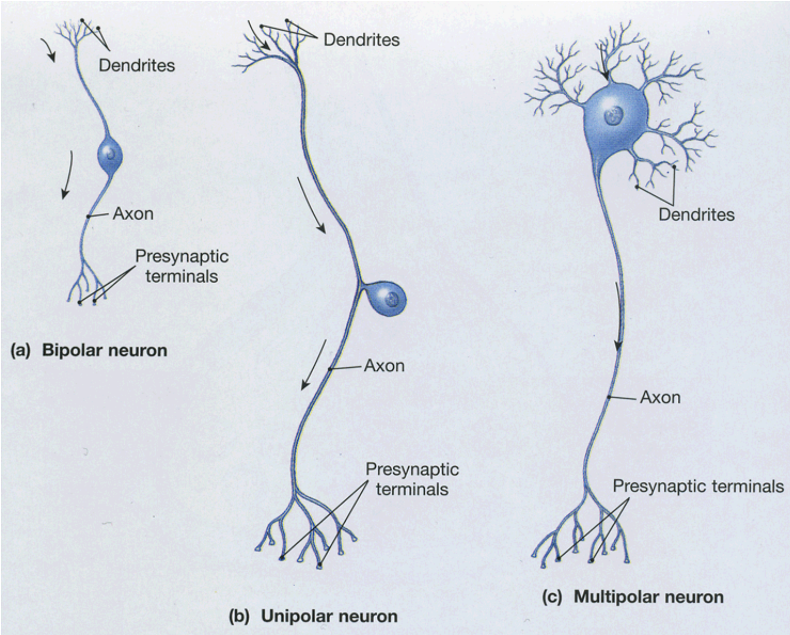

Figure 2.3: Three principal neuron types [1].

it needs energy. The ionic differences across the cell membrane is maintained by active pumps and this corresponds to a large fraction of the energy consumed by the brain [1,2].

Table 2.1: Units of concentration of ions and negatively charged large proteins inside and outside of the neuron [2].

	Na^+	K^+	Cl^-	Ca^{2+}	Protein
Outside	440	20	560	10	few
Inside	50	400	40-150	0.0001	many

If an external stimulus is applied to the membrane, a graded potential following the stimulus is observed. Moreover, if the membrane is depolarized up to a threshold (about -40mV), a brief (0.5ms-2.0ms) and large (about 40mV) change occurs. As shown in Figure 2.4, this brief change is called action potential or informally a “spike”. Unlike the graded potentials, an action potential is propagated through axon by ionic mechanisms. Also, the amplitude of action potential does not depend on the amplitude of stimulus. It is an all-or-none signal [2].

In the 1950s, Hodgkin and Huxley have studied the mechanisms of action potential and their experiments revealed that the action potential is a result of changes in the permeability of the cell membrane to ions like Na^+ and K^+ . Figure 2.5 shows that

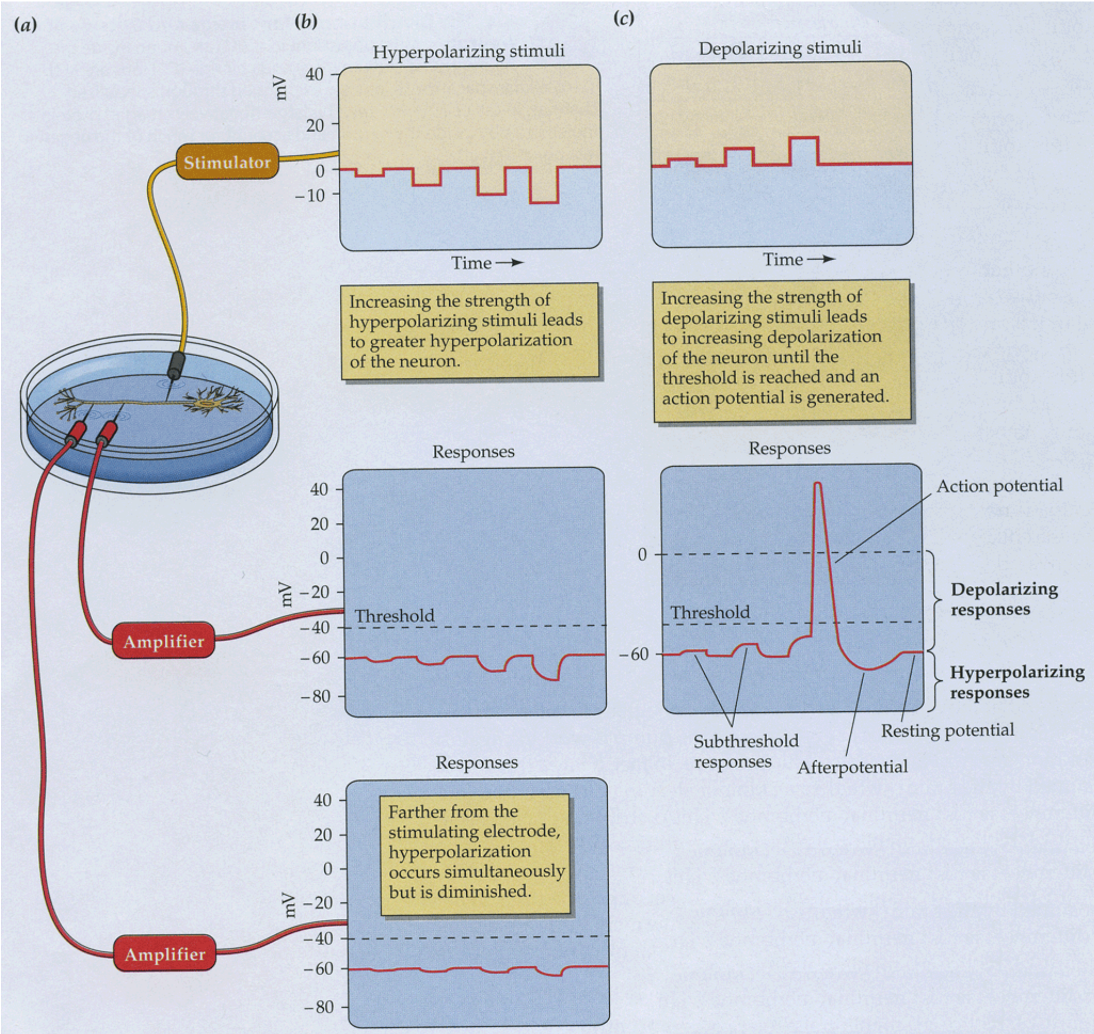

Figure 2.4: Effects of external stimulus [2].

(1) when resting, only K^+ channels are open, (2) depolarizing stimulus causes some Na^+ channels to open, (3) when threshold is reached, additional Na^+ channels open and a rapid change occurs, (4) inactivation of Na^+ channels causes repolarization or even hyperpolarization, (5) finally, the cell is returned to its resting potential [2].

An ion-channel is assembly of several proteins. The states of sub-units of this assembly represent the state of the channel. The ion-channel is either open or activated, thus channel allows a specific type of ions to pass, or, it is closed or inactivated, thus it does not allow ions to pass. The ion transfer may be gated by chemical or electrical signal as well as other physical events such as temperature. If an ion-channel is gated by a ligand which is a molecule that binds to a receptor, it is called ligand-gated. Similarly, an ion-channel is voltage-gated if it is gated by the electrical field near the channel. Ligand-gated channels are typically found at synapse where neurotransmitter or other

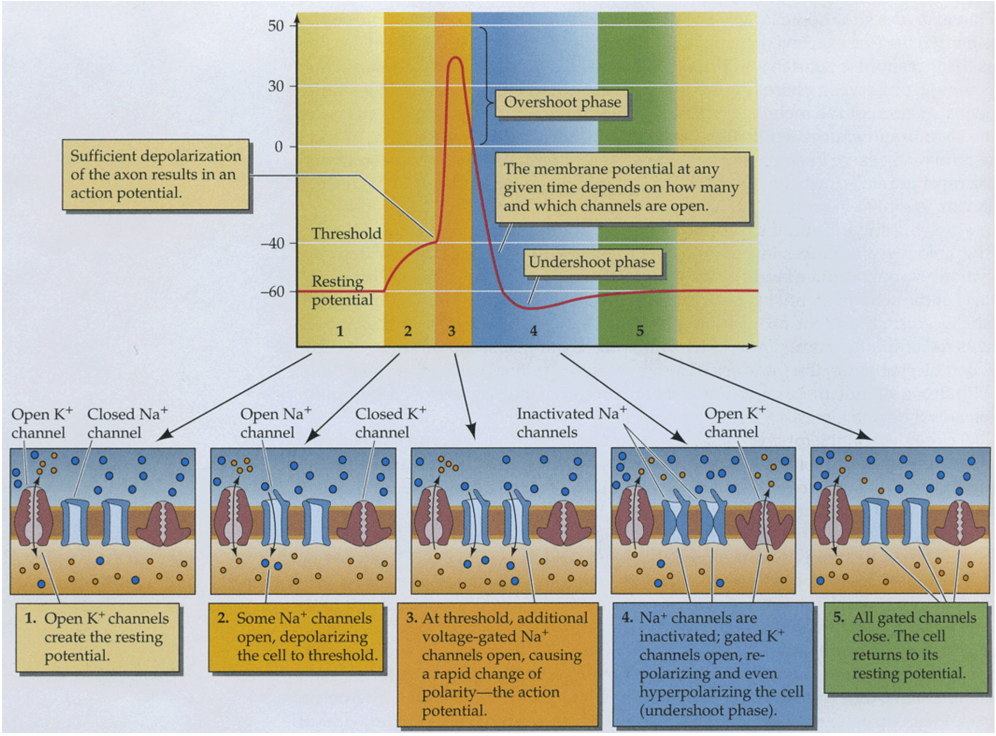

Figure 2.5: Action potential [2].

chemicals can function as ligand. On the other hand, voltage-gated channels, which include Fast Na^+ and Delayed-Rectifier K^+ channels, are typically spread over the cell's surface to generate and transmit action potentials.

The communication between two neurons occur at a junction called synapse either chemically or electrically. A chemical synapse (Figure 2.6) employs several chemical mechanisms to enable communication between the axon of pre-synaptic neuron and a dendrite or soma of post-synaptic neuron. On the other hand, in an electrical synapse (Figure 2.7), ions can freely move between cytoplasm of neighboring neurons and it is much faster than the chemical synaptic transmission. The number of chemical synapses is much more than the number of electrical synapses. A synapse usually refers to a chemical synapse.

In a synapse (Figure 2.8), (1-2) pre-synaptic action potential triggers the opening of Ca^{2+} channels, (3) the increase of Ca^{2+} inside the cell causes vesicles containing neurotransmitters to merge with the membrane, thus neurotransmitters are released to synaptic cleft, (4) neurotransmitters bind to receptors at post-synaptic neuron which directly or indirectly activate or deactivate ion-channels, (5) these ion-channels mediate post-synaptic currents, and (6-8) excess neurotranmitters are either broken down by enzymes or they are reuptaken [22]. The exact mechanism of transmitter

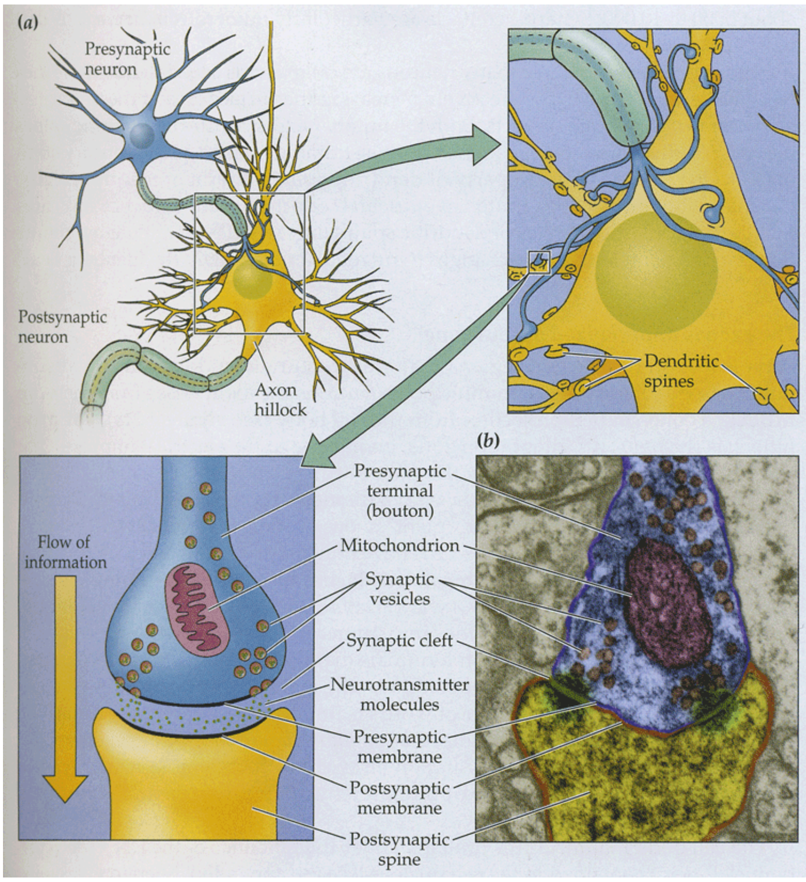

Figure 2.6: Chemical synapse [2].

release and the biochemical reactions at synapse are active research areas [2,22].

Glutamate and GABA (gamma-aminobutyric acid) are most common excitatory and inhibitory amino acid neurotransmitters found in the central nervous system. Glutamate works with AMPA (α-amino-3-hydroxy-5-methylisoxazole-4- propionic acid) and kainate receptors, which mediate fast excitatory synaptic currents, and NMDA (N-methyl d-aspartate) receptors, which is much slower and also sensitive to Mg^{2+} block which depends on the voltage. These properties make NMDA receptors a candidate for synaptic plasticity and activity-dependent development. Moreover, Mg^{2+} block's dependence on post-synaptic potential makes NMDA receptors a coincidence detector, which can be viewed as a mechanism to satisfy Hebb's learning rule. On the other hand, inhibitory neurotransmitter GABA activates fast $GABA_A$ receptors and slow $GABA_B$ receptors since additional chemical reactions are needed [22].

Receptors can be classified into two categories; ionotropic and metabotropic receptors. AMPA, kainate, NMDA and $GABA_A$ receptors are ionotropic receptors which indicate they have ion-channels directly gated by neurotransmitter (ligand). In contrast,

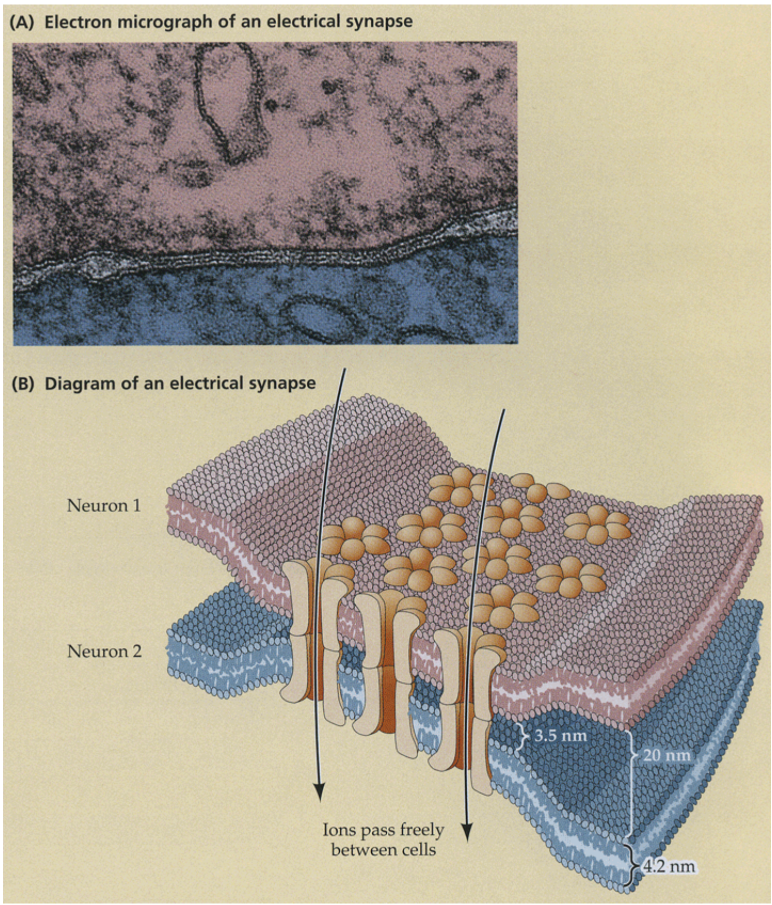

Figure 2.7: Electrical synapse [2].

metabotropic receptors do not form an ion channel and they are indirectly linked with other ion channels by intracellular messengers e.g. calcium or G-proteins.$GABA_B$ receptors are an example of metabotropic receptors where K^+ channels are not directly coupled but activated or deactivated by G-proteins which are produced after neurotransmitter binds to receptor. Synaptic interactions through second messengers are slower and this type of interaction is known as neuromodulation. GABA on $GABA_B$ receptors, adrenaline and noradrenaline on α_2 receptors, serotonin on 5-HT_1 (5-hydroxytryptamine) receptors and dopamine on D2 receptors are examples of neuromodulators. They gate ion-channels through the action of second-messengers [23].

2.1.2 Brain

Nervous system is divided into peripheral and central subsystems. Central subsystem includes brain and spinal cord. According to Maclean, human brain is composed of three interconnected brains (Figure 2.9). As evolutionary theorists suppose mammals

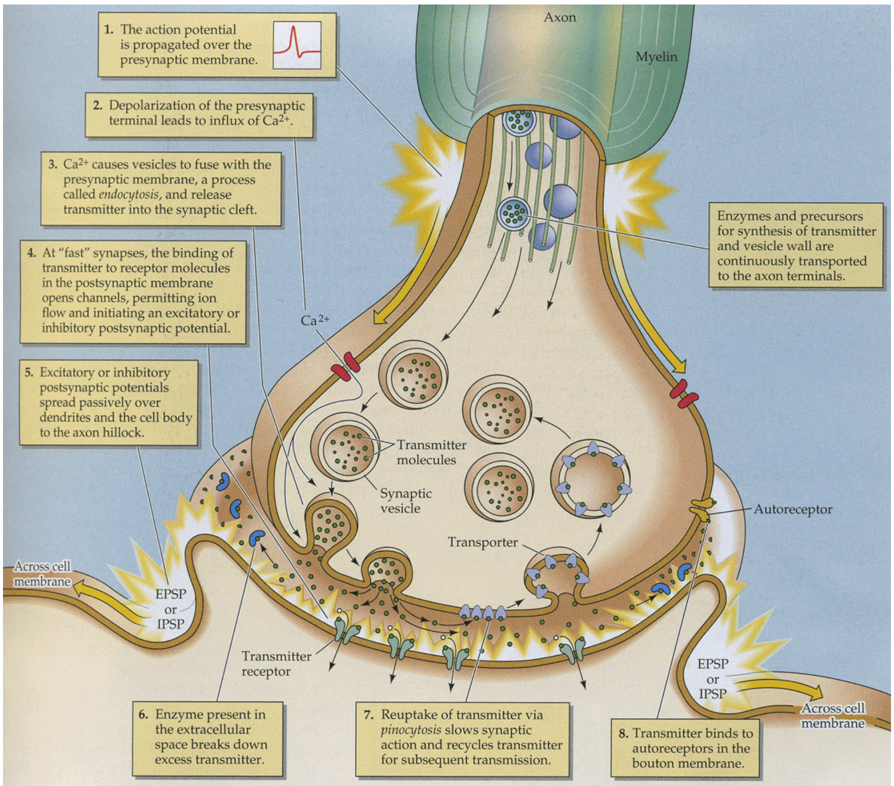

Figure 2.8: Steps in Transmission at a Chemical Synapse [2].

arose from a certain ancient reptiles, the reptilian brain can be found at the deepest. It elicits behaviors that cannot be modified by experience, e.g. reflexive behaviors. Second, an old mammalian brain is found over the reptilian brain. It is survival oriented and controls actions like feeding, fighting, fleeing and sexual behavior. Finally, above all of these, the new mammalian brain, or neocortex, is found and it is responsible for intelligent behavior [1,2].

The developmental divisions of the brain is shown in Figure 2.10 in both the embryo and the adult and an overall anatomy is shown in Figure 2.12. Hindbrain, which contains cerebellum, medulla oblongata, pons and raphe system, is responsible for coordinating motor responses, essential functions such as heart rate and respiration, passing sensory information to other brain regions and controlling sleep-wake cycle. Midbrain, which contains tectum and tegmentum, relays visual and auditory information, controls simple reflexes, affects motor system and reticular system for arousal and consciousness. Forebrain contains basal ganglia, hypothalamus, limbic system, thalamus and cerebral cortex. Basal ganglia plays an important role in integrating voluntary movements, hypothalamus generates motivation to respond internal states or external stimulus, limbic system is involved with aspects of emotions and memory storage and thalamus is the relay station of most sensory receptors on

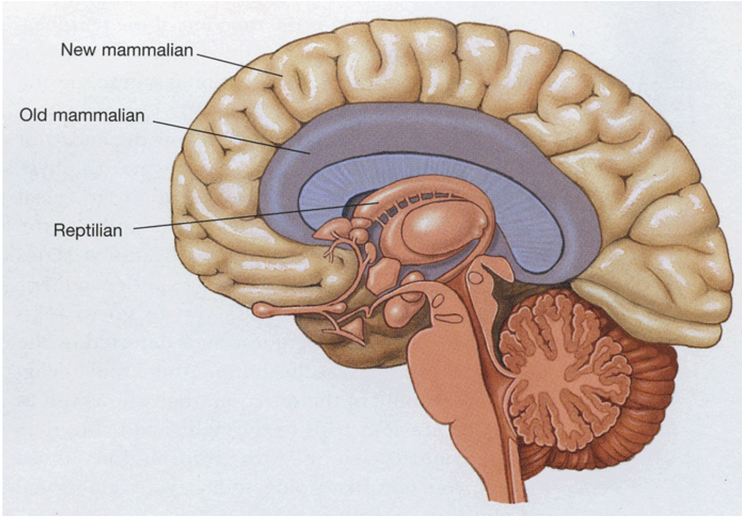

Figure 2.9: Three main divisions of brain [1].

their way to cerebral cortex [1,2].

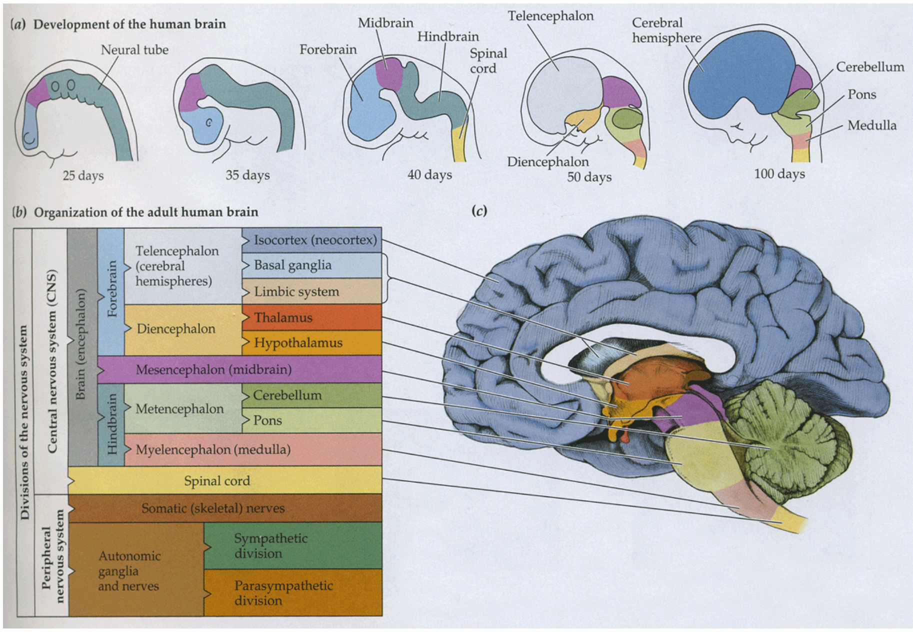

Figure 2.10: Developmental stages of the brain [2].

After Wernicke's work on language centers in the brain, Brodmann studied and created a cytoarchitectonic map of the brain consisting 52 functionally distinct areas in the human cerebral cortex. Cytoarchitectonic map is created based on the structure and characteristic arrangement of the cells. The map of the brain created by Brodmann is still widely used in the clinical studies as well as in the research.

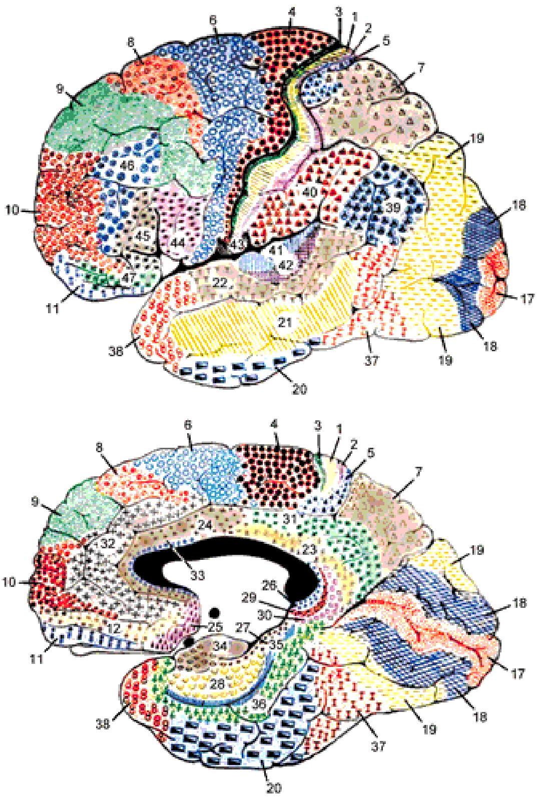

Figure 2.11: Brodmann Map [24].

Cerebral cortex, which is separated into two hemispheres, is the outmost structure of the brain. It is also the part in which language and other cognitive functions occur. Basically, the brain consists of sensory, motor and association areas. The two hemispheres are interconnected through cerebral commissures such as corpus callosum. The hemispheres processes information from and control the opposite (contralateral) side of the body. For example, a sound heard with left ear is processed in the right hemisphere. Furthermore, each hemisphere is divided into four lobes; occipital, temporal, parietal and frontal. Frontal lobe is concerned with the decision making and the control of movement. Hearing, some aspects of learning, memory and emotion occurs in temporal lobe. In the parietal lobe, the tactile sensation and the body image form. Finally, the vision occurs in occipital lobe. It should be noted that the categorization of functionality above does not mean the functions are purely localized [1,25].

Figure 2.13 shows that primary motor cortex, brainstem and spinal cord are main parts of motor processing, but basal ganglia, cerebellum and other motor cortices are also important. As shown in Figure 2.14, almost all sensory information arrives to brain through spinal cord and brainstem, then they are relayed to primary sensory cortical

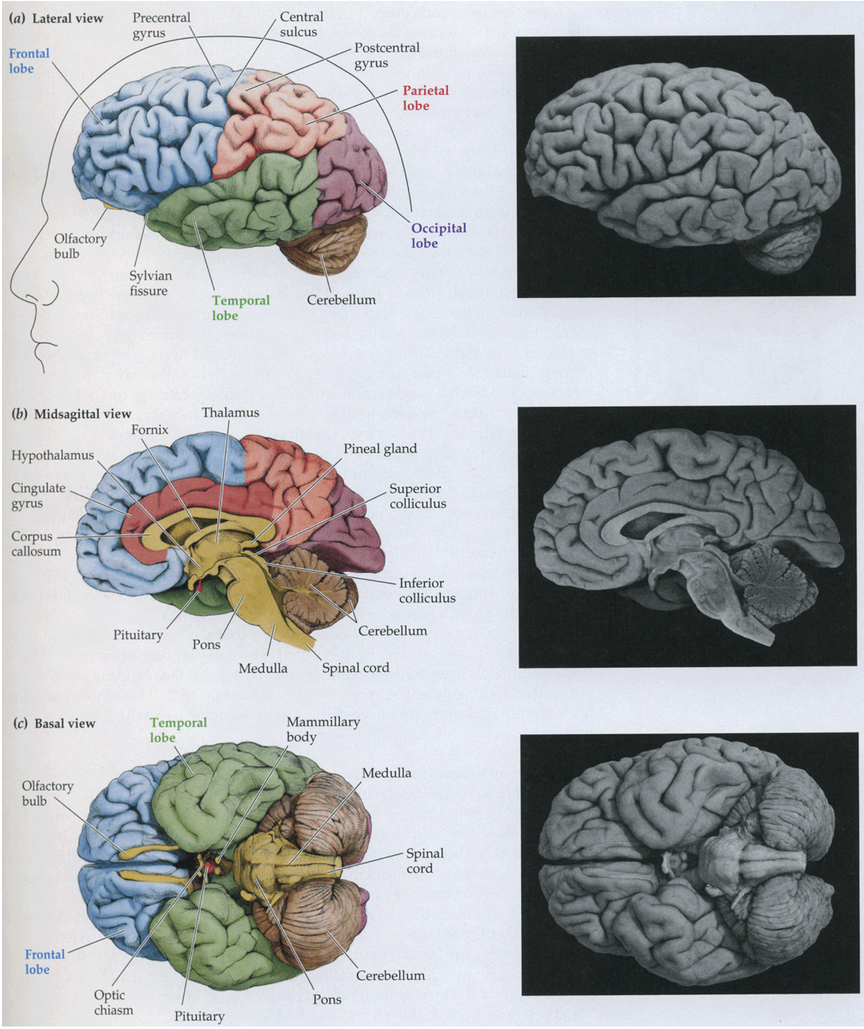

Figure 2.12: An overall anatomy of brain [2].

areas through thalamus. An important structural property of the brain is principles of convergence and divergence. (Figure 2.25) Sensory information in sensory receptors usually converges into a zone (e.g. as much as 100:1 convergence in human eye) and then diverges to various parts of cortex [2].

Cortex has some important organizational principles. A major principle is topographic mapping. As shown in Figure 2.15, sensory and motor areas are arranged topographically, which means adjacent neurons have adjacent mapping to real world. However, the maps in the cortex are distorted. For example, representation of region corresponding to fovea in primary visual cortex (V1) and representation of hands in somatosensory cortex occupies considerably large area [13].

Many brain areas not only have a topographic organization but also have laminar organization. Figure 2.16 shows laminar organization, where dark spots are stained cell

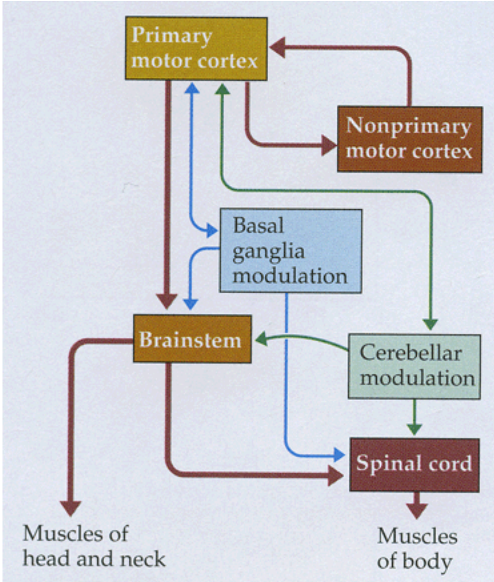

Figure 2.13: Main pathways in motor processing [2].

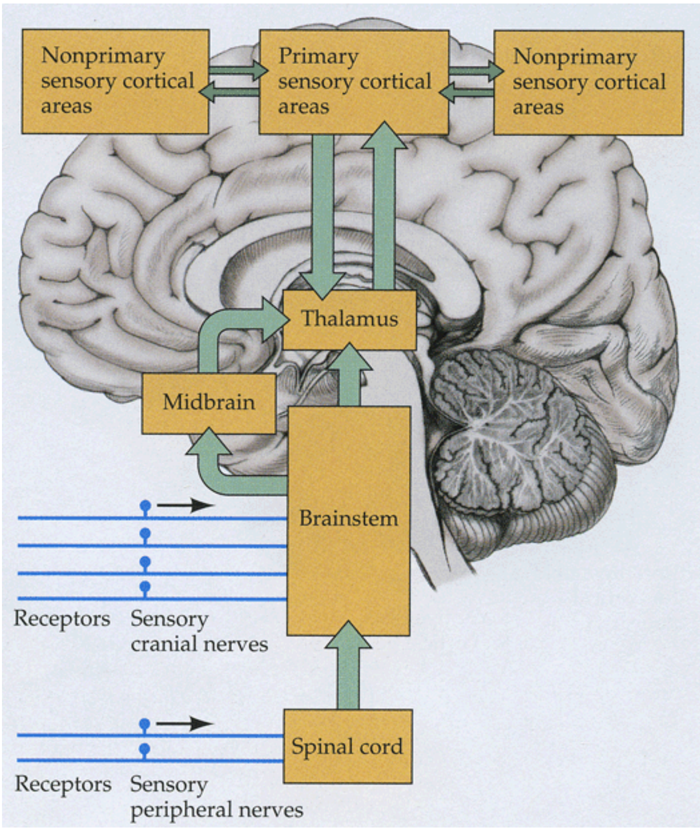

Figure 2.14: Main pathways in sensory processing [2].

bodies. As well as horizontal layers, cortex also displays vertical organization called cortical columns. In 1950s, Mountcastle discovered that cortex is composed of nearly

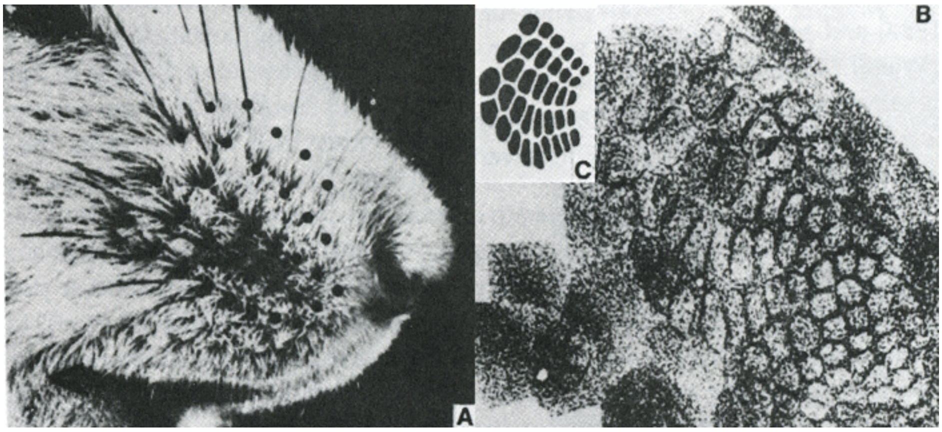

Figure 2.15: Spatial organization is preserved for whiskers [13]. From Woolsey & van der Loos, 1970.

identical cortical columns and it is speculated that cortex performs its computation by the integration of identical cortical columns [14]. Figure 2.17 illustrates both cortical layers and cortical columns for vision.

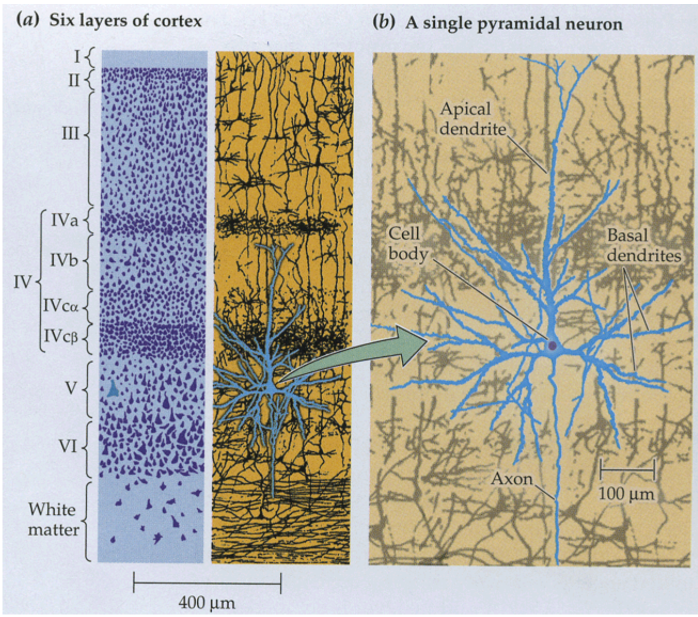

Figure 2.16: Cortical layers of brain [2].

2.2 Learning and Plasticity

Brain is a highly adaptable structure that continuously modifies and updates its functions such as perception, motor skills, regulation and reasoning by experience.

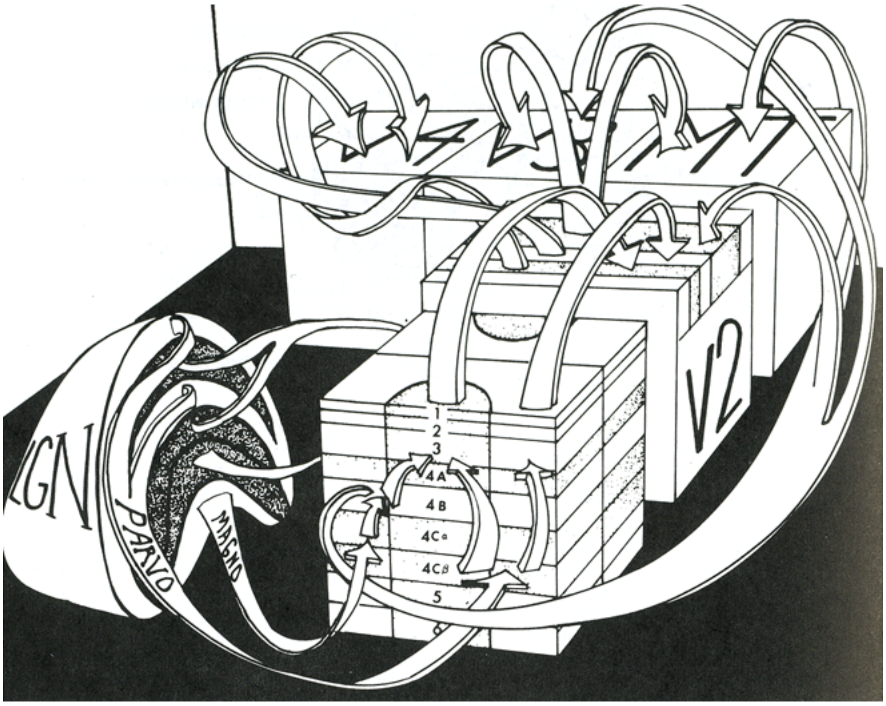

Figure 2.17: Visual pathways for one eye [13]. From LeVay & Nelson, 1991.

The basic question in learning, plasticity and memory research is how causal neuronal plasticity, which is a local property, give rise to global property of brain, the adaptable behavior of organism [13].

Learning and plasticity is being studied scientifically in four main streams; (1) neural mechanisms of relatively simple kinds of plasticity, such as classical conditioning, (2) study of temporal lobe including the hippocampus, perirhinal structures and amygdala, (3) study of visual system development, and (4) gene-development relationship [13].

The second category is where the main discussion occurs. Study of temporal lobe became a main topic after the discovery of the patient H.M. in the 1950s. To manage epilepsy, mesial temporal lobe of H.M. is bilaterally removed. Milner and her colleagues have done pioneering studies with the case of H.M. and they have showed that H.M. could not learn and retain new information (anterograde memory), even the events were important to him. However, he can remember many things happened before his surgery (retrograde memory) and his short-term memory was within normal range [13]. The changes that persist for tens of minutes or longer are called long-term potentiation (LTP) or long-term depression (LTD) [17].

During 1970s, in amnesic patients, it is discovered that a task like completion of a picture seen earlier could be performed normally, however, the patient denied having seen the picture completed. These results indicate a major division in memory system, which are reportable (explicit) and without being awareness (implicit), shown in Figure 2.18, [13].

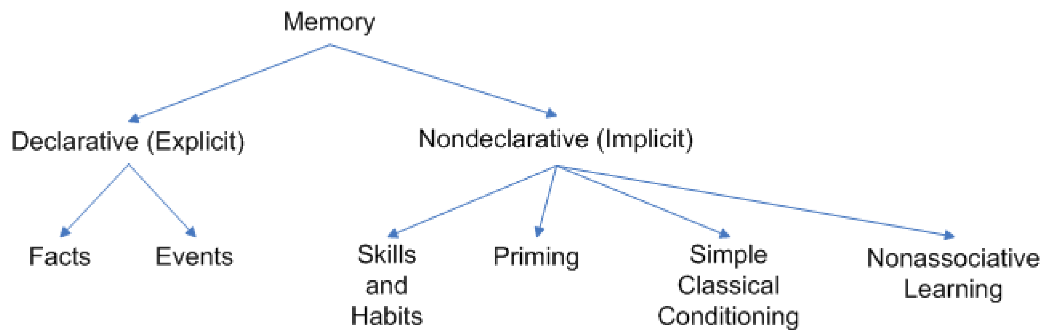

Figure 2.18: Classification of memory [13]. From Squire and Zola-Morgan, 1991.

Assuming global changes are induced by local changes, the neuronal basis of plasticity without an intelligent teacher is the next question. There are many ways for learning to occur in a neuron, e.g. new dendrites, new synapses may form or vanish or the amount of chemicals required for synaptic communication may be modified. Also, the exact temporal and spatial position of modification should be understood. In 1949, Donald Hebb addressed these questions in his book "Organization of Behavior" [13]. Although half a century before William James also mentioned a similar mechanism, the learning mechanism described below is attributed to Hebb and known as Hebbian learning [26].

> When an axon of a cell A is near enough to excite cell B or repeatedly or persistently takes part in firing it, some growth or metabolic change takes place in both cells such that A's efficiency, as one of the cells firing B, is increased (Organization of Behavior, 1949) [13].

Simplest form of Hebb rule to modify synaptic strength can be given as $\Delta w_{AB} = \varepsilon V_B V_A$, where w_{AB} is the synaptic weight, V_B and V_A are average firing rates of neuron A projecting to neuron B [13]. In this formulation, the synaptic change can only be positive since the average firing rate cannot be less than zero. If average firing rate V_B is replaced with the deviation from average firing rate $(V_B - \bar{V_B})$, where V_B is the instantaneous, $\bar{V_B}$ is the average activity of postsynaptic neuron, the synaptic weights can also be decreased [13].

Based on the neuropsychological data from the cases of H.M. and other patients, the ablation of hippocampus does not impair the memory of events occurred some time before the lesion. Therefore, hippocampus cannot be the permanent store for information, but hippocampus may take a role of a teacher for cortex to learn more detailed and categorical representations. Also, based on the studies, hippocampus is not needed for remembering events in short term period for about 60 seconds [13].

According to studies done by Baddeley and Hitch, it seems the short term memory store, which Baddeley called working memory, is able to hold as much as five or six items in several seconds. Moreover, it seems there are number of short term stores in relation for auditory, visuospatial, verbal items. The link between behavior, anatomy and physiology of short term memory has been found by Goldman-Rakic (1987). First, it is found that monkeys have a deficit in working memory when dorsalateral prefrontal cortex is lesioned. Second, cells in that region is recorded in intact animals and it is found that some cells continues to fire during the delay between the presentation of stimulus and cue [13].

Persistent activity of cells, which are attributed to working memory mechanism, suggested a circulation in a recurrent network. However, there must be a population of neurons circulating their signal for more than a hundred times for a five real seconds which is realistically unlikely to occur. A model for short term memory is proposed by Zipser (1991) by employing a gating signal as in electronic latch circuits. Other than network structure, another possibility of such a mechanism can be based on short term cellular adaptation [13].

After the introduction above, this section is devoted to associative learning.

2.2.1 Associative Learning

According to Descartes' mind-body dualism, human behavior consists of involuntary automatic responses (reflexes) elicited by sensory stimulation and voluntary behaviors which arise from the mind or soul which in turn interacts with body at the pineal gland. Animals are thought to have no mind or soul, so, they are considered as automata and capable of making only automatic responses. It is the theory of Darwin, who proposed variations in organisms are due to natural selection, that explained human and animals are not intrinsically different but evolved in order to adapt to their environments. Both Darwin and Romanes have presented many anecdotal stories related to intelligent behavior of animals. Later, Morgan stated that animals acquire the solution by trial-and-error and not by reason. Thorndike, one of the pioneers in the field, is probably the first scientist who worked on the problem of intelligence in the laboratory by studying animal behavior experimentally. According to Thorndike, animals do not learn by insight or reason. Instead, they learn by forming stimulus-response associations which are caused by the close contiguity of stimulus-response (SR) and repetition. Thorndike also introduced the law of effect. The law of effect states that the reward given after response causes the SR formation to be strengthened. Today, the law of effect is known as reinforcement, which is a term introduced by Pavlov to refer to unconditioned stimulus (such as food) in classical conditioning [3,27].

Classical (Pavlovian) Conditioning

Aristotle, centuries ago, stated that "when two things commonly occur together, the appearance of one will bring the other to mind" which is named the law of contiguity. The law of contiguity is a simple form of classical conditioning. In the late 1900s, Pavlov discovered classical conditioning phenomena when he was researching on digestive system. In the basic experiment, a light (visual stimuli) or tone (auditory stimuli), which is called conditioned stimulus (CS), preceding the food delivery, which is called unconditioned stimulus (US), is given to a dog, and, the dog salivates. The response to the US is called the unconditioned response (UR). The CS alone does not cause any salivation, so, it is said CS is neutral with respect to UR. After some trials, when the light or tone (CS) is given alone, the dog began to salivate without the actual delivery of food. The response to the CS alone is called conditioned response (CR). Thus, in classical conditioning, a neutral stimulus (light or tone, CS) is conditioned to take the role of a stimulus (food, US) which causes the response [27].

There are different types of responses studied in classical conditioning including knee jerk reflex and eye blink reflex which are involuntary automatic responses of organism, and emotional responses which are measured by the change of the conductance of the skin which is called galvanic skin response (GSR). Conditioned taste aversion is another type of response which is caused by externally making subjects ill after subject has ingested a substance with a typical taste. Therefore, subject avoids the substance which is artificially made them ill. Other than involuntary or reflexive behavior, a voluntary behavior can also be a conditioned response by a procedure called autoshaping. It is first discovered by Brown and Jenkings in their studies with pigeons' key pecking behavior [27].

It is known that the best conditioning occurs when the CS is followed by US in close temporal contiguity. Although the strength of conditioning decreases, it is possible that conditioning still occurs with different procedures as shown in Figure 2.19.

In classical conditioning, a rich set of behavioral phenomena summarized in Table 2.2 is observed. A conditioned behavior can be made extinct by repeatedly pairing the same stimulus with nothing. If stimulus is partially paired with response, the elicited

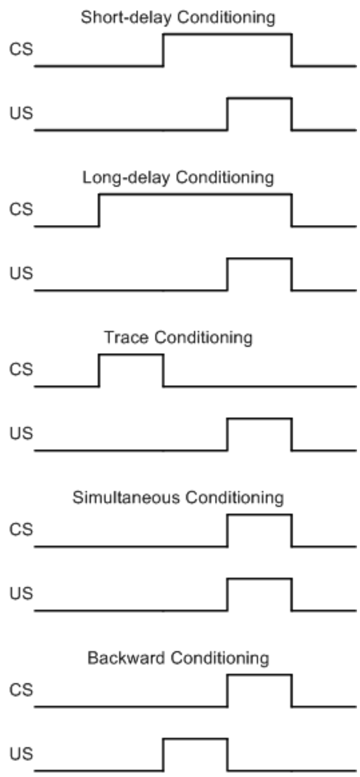

Figure 2.19: Procedures used in classical conditioning. Reproduced from [27].

conditioning is less powerful than the original procedure. If stimulus s_1 is already conditioned with a response, presenting another stimulus s_2 with the stimulus s_1 block the conditioning of s_2 with the response. If two stimuli s_1 and s_2 is conditioned together with nothing, and then s_1 alone conditioned with a response, presenting s_2 inhibits the response. If two stimuli s_1 and s_2 is conditioned together with a response, the more salient or stronger stimulus elicits more powerful response when it is presented alone. Finally, if stimulus s_1 is conditioned with a response, conditioning another stimulus s_2 with the stimulus s_1 results conditioning of the stimulus s_2 with the response [17].

Instrumental (Operant) Conditioning

In classical conditioning, a stimulus (CS) is conditioned to take the role of another stimulus (US) that elicits a response (UR). In instrumental conditioning, the response is conditioned to occur because the response precedes the reinforcer. Thus, response should occur for reinforcer to appear. In contrast to classical conditioning where the experimenter is responsible to deliver reinforcer, the reinforcer is delivered according to the response of organism in instrumental conditioning. An animal can be trained for complex behavior such as rolling over or pressing a bar through a process of shaping or successive approximations [27].

Table 2.2: Classical conditioning paradigms. Reproduced from [17]. s, r and $\cdot$ represents stimulus, reward and absence of reward respectively. 'r' and '$\cdot$' is expectation of reward and absence of reward. α denotes a partial or weakened expectation.

Paradigm	Pre-Train	Train	Result
Pavlovian		$s \to r$	$s \to$ 'r'
Extinction	$s \to r$	$s \to \cdot$	$s \to$ '$\cdot$'
Partial		$s \to r$, $s \to$ '$\cdot$'	$s \to \alpha$'r'
Blocking	$s_1 \to r$	$s_1 + s_2 \to r$	$s_1 \to$ 'r' $s_2 \to$ '$\cdot$'
Inhibitory		$s_1 + s_2 \to \cdot$ $s_1 \to r$	$s_1 \to$ 'r' $s_2 \to$ -'r'
Overshadow		$s_1 + s_2 \to r$	$s_1 \to \alpha_1$'r' $s_2 \to \alpha_2$'r'
Secondary	$s_1 \to r$	$s_2 \to s_1$	$s_2 \to$ 'r'

Instrumental conditioning can be done in two settings. The experimenter allows the subject to perform in discrete trials or the subject can perform freely. The latter is also called free-operant responses. Discrete trials approach was done by Thorndike in his puzzle box mainly with cats. Free-operant responses approach was done by Skinner in so called Skinner box or operant chamber mainly with rats. Many types of behavior, such as a child's cry to take attention of parents, can be viewed as a response in instrumental conditioning [27].

Instrumental conditioning can also be done with aversive outcomes instead of positive reinforcement. In escape learning, an animal learns to relocate between two parts of the place he stays when a harmless electrical shock is given. This type of reinforcement that reduces or terminates the aversive stimuli is called negative reinforcement. In avoidance learning, a warning stimulus is given before the aversive stimuli. After trials, although the aversive stimuli is absent, the animal learns to relocate when the warning stimuli is presented. Avoidance learning created a theoretical problem since it is not obvious what is reinforced the avoidance behavior. The first explanation is given by Miller (1948) and Mowrer (1947) which is named as the two-factor theory of avoidance learning. According to the theory, both classical conditioning of fear and instrumental conditioning of escape response from fear exist in the avoidance learning. Later, a cognitive explanation by Seligman and Johnston (1973) is proposed to overcome the unanswered questions by the two-factor theory. They proposed response-outcome expectancy is actually learned in avoidance learning and the animal choice the response which is paired with the absence of shock. The final type of aversion is due to punishment. Both positive and negative reinforcement increases the possibility of response to occur, however, punishment decreases that possibility [27].

During instrumental conditioning, the reinforcer, either reward or punishment, can be given immediately or can be delayed partly or wholly until a certain sequence of actions are completed. Thus, future expectations, but not immediate receipt of reinforcer, play a role learning the appropriate action in instrumental conditioning with delayed rewards [17]. Instrumental conditioning alone is used to refer instrumental conditioning with immediate rewards in this study.

Classical conditioning may also be explained in terms of instrumental conditioning in some experiments. It is not known that a simple experiment with monkeys, where monkey learns to take a food when an audible signal presented, involves classical or instrumental conditioning [28]. However, it is known by the neat experiment designed by Browne (1976) that classical conditioning is a different paradigm, so, it cannot be reduced to a form of instrumental conditioning generally. Surprisingly, Mackintosh notes that instrumental conditioning may be explained as a form of classical conditioning, however, as he argues, it is not possible to explain all instrumental learning in his way [29].

2.3 Computational Models

2.3.1 Neuron Models

An electronically compact neuron has only a few long segments and all the segments are narrow. Thus, in an electronically compact neuron, membrane potential is uniform across the surface of the neuron. Such neurons can be modeled by simpler models. In the following sections, models of electronically compact single neurons; Hudgkin-Huxley, Connor-Stevens and Integrate-and-Fire models is briefly explained. A model for electronically incompact neuron should employ compartmental models in order to consider spatial distribution and propagation of membrane potential and those models are not covered in this study.

Hodgkin-Huxley Model

The first complete mathematical explanation of dynamics of a single neuron is given by Hodgkin, Huxley and Katz (1952). They have conducted experiments on the giant axon of the Longfin Inshore Squid (Atlantic Squid, Loligo Pealeii), which is part of the squid's water jet propulsion system, and the large diameter of axon enabled inserting electrodes inside. Based on experimental data, they have constructed an empirical model to describe generation of action potentials. According to their study, current is carried through membrane; (1) by charging the membrane capacitance, and (2) by the movement of ions. Na^{+} and K^{+} are found to be main components of ionic current, whereas Cl^{-} and other ions are considered as leakage current. Figure 2.20 shows the Hodgkin-Huxley model as an electrical circuit.

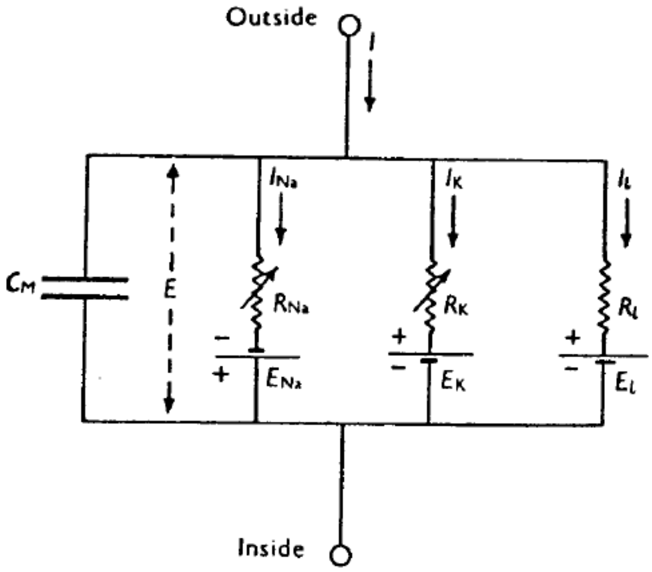

Fig. 1. Electrical circuit representing membrane. $R_{Na} = 1/g_{Na}$; $R_K = 1/g_K$; $R_l = 1/\bar{g}_l$. R_{Na} and R_K vary with time and membrane potential; the other components are constant.

Figure 2.20: Hodgkin-Huxley Model as an electric circuit [30].

Total membrane current is given by:

$$I = C_M \frac{dV}{dt} + I_i \tag{2.1}$$

where I is the total membrane current density (inward current positive), I_i is the ionic current density (inward current positive), V is displacement of the membrane potential from its resting value (depolarization negative), C_M is the membrane capacitance per unit area (assumed constant) and t is time [30].

The ionic current I_i is sum of the sodium, potassium and leakage currents, which are given by:

$$I_i = I_{Na} + I_K + I_l \tag{2.2}$$

$$I_{Na} = g_{Na}(E - E_{Na}) \tag{2.3}$$

$$I_K = g_K(E - E_K) \tag{2.4}$$

$$I_l = \bar{g_l}(E - E_l) \tag{2.5}$$

where E_{Na} and E_K are the equilibrium potentials for the sodium and potassium ions which is given by Nernst equation shown in Equation 2.6. E_l is the potential when leakage current due to chloride and other ions is zero [30].

$$E_X = \frac{V_T}{z} ln \frac{[X]_{outside}}{[X]_{inside}} \tag{2.6}$$

K^+ conductance is given by:

$$\begin{aligned} g_K &= \bar{g_K} n^4 && (2.7) \\ \frac{dn}{dt} &= \alpha_n(1-n) - \beta_n n && (2.8) \end{aligned}$$

where $\bar{g_K}$ is total conductance per unit area, α_n and β_n are rate constants which depend on voltage. (refer to [30])

Na^+ conductance is given by:

$$\begin{aligned} g_{Na} &= m^3 h \bar{g_{Na}} && (2.9) \\ \frac{dm}{dt} &= \alpha_m(1-m) - \beta_m m && (2.10) \\ \frac{dh}{dt} &= \alpha_h(1-h) - \beta_h h && (2.11) \end{aligned}$$

where, similarly, $\bar{g_{Na}}$ is total conductance per unit area, α_m, β_m, α_h and β_h are rate constants which depend on voltage. (refer to [30])

Connor-Stevens Model

There are two issues with Hudgkin-Huxley model. The first is, when a spike occurs, the membrane potential jumps discontinuously. This is called type II behavior, whereas membrane potential is continuous when a spike occurs in type I behavior. The second is that Hudgkin-Huxley model does not explain spike-rate adaptation. These issues are solved with the Connor-Stevens model. First, A-current transient K^+ channel is included into the model for type I behavior. Second, slow Ca^{2+} dependent K^+ channel is included to model after-hyperpolarization (AHP) conductance, which builds up

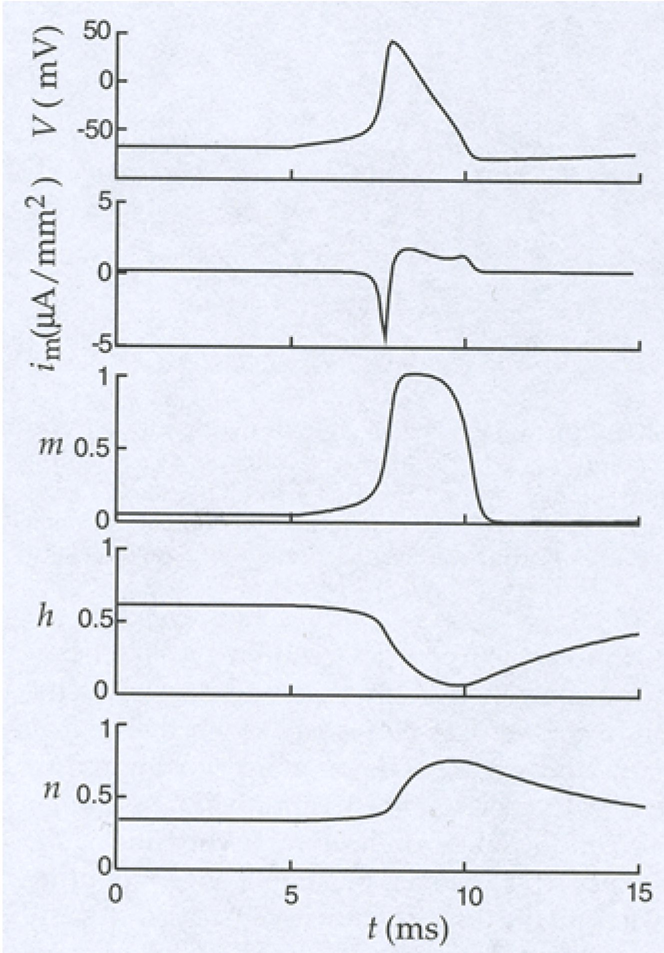

Figure 2.21: Dynamics of V, m, h and n in Hodgkin-Huxley model [17].

during sequence of action potentials, which contributes to spike-rate adaptation [17].

Integrate-and-Fire Model

During simulations, especially when large group of neurons are involved, solving differential equations of Hodgkin-Huxley model is exhausting even in very powerful computers. The simplest model that can be used for spiking neurons is called Integrate-and-Fire model. In the simplest case, this model ignores all the active ion-channels and synapses, it is called passive or leaky integrate-and-fire model. Therefore, there exist only leakage current and current injection, and the equation is simplified into: [17]

$$C_M \frac{dV}{dt} = -\bar{g}_L(V - E_L) + \frac{I_e}{A} \quad \textbf{(2.12)}$$

In passive integrate-and-fire model, when membrane potential is greater than spike threshold potential ($V_t h$), a spike is generated and V is set to resting potential (E_L), which is smaller than spike threshold potential.

Stochastic Models

All the models above are deterministic in their nature. However, as shown in Figure 2.22, real neurons do behave unpredictably probably due to noise.

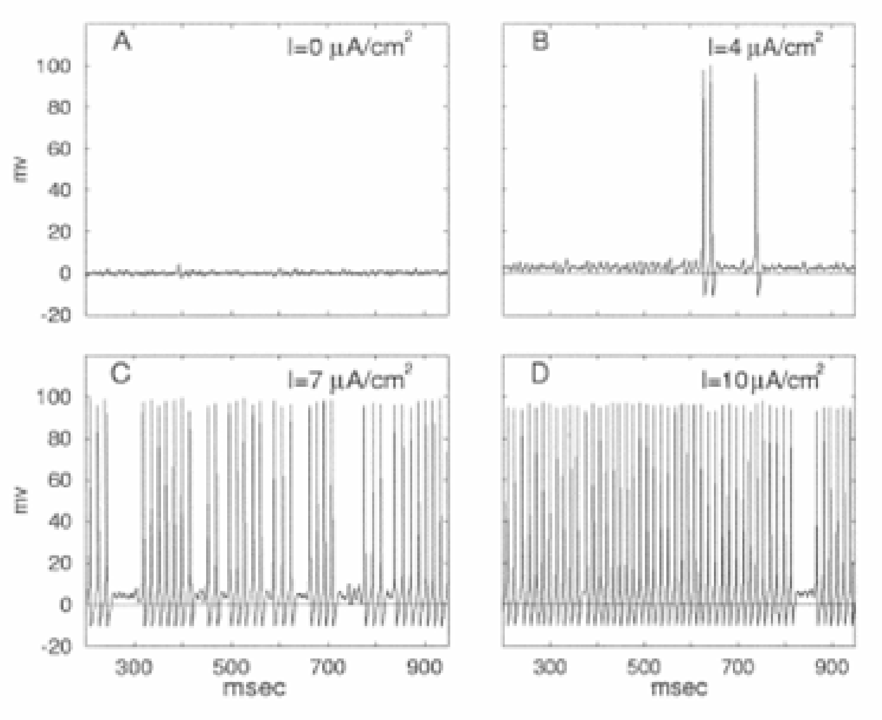

Figure 2.22: Response of stochastic model [31].

In order to model stochasticity in a neuron, the deterministic equations of sodium and potassium conductances can be replaced by stochastic markov chains in the Hodgkin-Huxley model. Thus, potassium conductance can be modeled by a five states and sodium conductance can be modeled by an eight states markov chain as shown in Figure 2.23 and 2.24 and Hodgkin-Huxley equation can be modified accordingly.

$$[n_0] \underset{\beta_n}{\overset{4\alpha_n}{\rightleftharpoons}} [n_1] \underset{2\beta_n}{\overset{3\alpha_n}{\rightleftharpoons}} [n_2] \underset{3\beta_n}{\overset{2\alpha_n}{\rightleftharpoons}} [n_3] \underset{4\beta_n}{\overset{\alpha_n}{\rightleftharpoons}} [n_4]\,,$$

Figure 2.23: Markov chain of potassium conductance [31].

$$
\begin{array}{ccccccc}
[\mathrm{m_0h_1}] & \underset{\beta_m}{\overset{3\alpha_m}{\rightleftharpoons}} & [\mathrm{m_1h_1}] & \underset{2\beta_m}{\overset{2\alpha_m}{\rightleftharpoons}} & [\mathrm{m_2h_1}] & \underset{3\beta_m}{\overset{\alpha_m}{\rightleftharpoons}} & [\mathrm{m_3h_1}] \\
\alpha_h \uparrow\downarrow \beta_h & & \alpha_h \uparrow\downarrow \beta_h & & \alpha_h \uparrow\downarrow \beta_h & & \alpha_h \uparrow\downarrow \beta_h \\
[\mathrm{m_0h_0}] & \underset{\beta_m}{\overset{3\alpha_m}{\rightleftharpoons}} & [\mathrm{m_1h_0}] & \underset{2\beta_m}{\overset{2\alpha_m}{\rightleftharpoons}} & [\mathrm{m_2h_0}] & \underset{3\beta_m}{\overset{\alpha_m}{\rightleftharpoons}} & [\mathrm{m_3h_0}],
\end{array}
$$

Figure 2.24: Markov chain of sodium conductance [31].

Multi Compartment Models

In single compartment models mentioned above, the membrane is assumed to be equipotential. However, mainly due to the length of axon, the membrane can be long and spatial effects must be taken into account. Therefore, instead of having a single compartment for each neuron, a neuron is modeled having multi compartments. A single Purkinje cell is modeled with 4500 compartments by De Schutter and Bower (1994) and De Schutter (1998). Since the spatial structure is important for such models, 3D reconstructions of real neurons are usually used in order to find the position, length and other parameters of the compartments.

Neural Coding

Individual spikes of a neuron can either independently encode an information or encode the information as a correlation among spikes. Former is called independent spike coding and the latter is called correlation coding. Independent spike codes are much easier to analyze than correlation codes, however, it is known correlation codes also exist in the brain [17].

Theoretically, the probability distribution of spike times as a function of the stimulus completely characterizes the neural response. Sometimes, an inhomogeneous Poisson process, which is computed from the time dependent firing rate $r(t)$, is sufficient to describe the spike generation. If this is the case, the time dependent $r(t)$ contains all the information encoded in the spike train. Such a neural code is called a rate code [17].

2.3.2 Network Models

Similar to biological architecture, a network consists of some neurons and interconnections between them. The neurons can be based on any model mentioned above or simple firing rate models, which are also extensively used in engineering, can be used. The decision of selecting the neuron model to be used in the network is made according to the theoretical findings and the availability of computational resources.(See [17])

Based on the patterns of connectivity in brain especially in cortex, the connections between the neurons can be; bottom-up or feedforward, top-down and recurrent, which are shown in Figure 2.25. The neurons at the input (thalamus) are connected in a bottom-up fashion to layers above (cortex). Also, there exists enourmous feedback connections from upper layers (e.g. cortex) back to lower layers (e.g. thalamus) in brain which is called top-down. On the other hand, recurrent connections exist between the neurons which stays at the same stage in the processing pathways. Although bottom-up and top-down connections can be comparable, the recurrent connections typically outperform the number of other two [17].

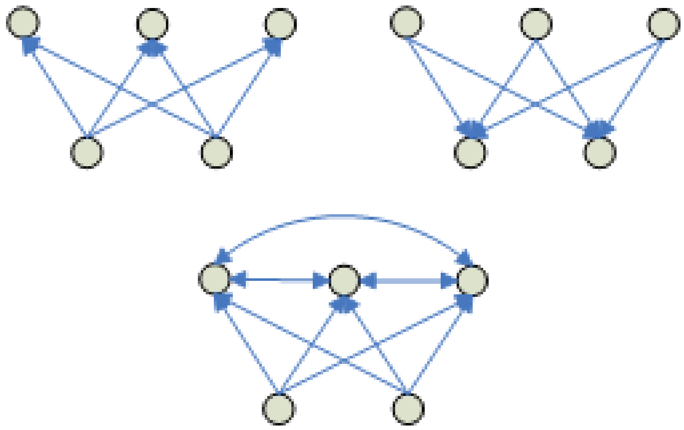

Figure 2.25: Three types of networks are shown. Top left figure shows only bottom-up or feedforward connections, top right figure shows only top-down connections. Figure at bottom is the same as top left with additional recurrent connections in upper layer. These figures also show convergence and divergence principles between upper and lower layers with a ratio of 2:3.

Similar to neural coding, networks of neurons can employ two types of coding. One is called independent neuron coding and it means each neuron independently code some parts of an information. Like independent spike codes, the networks based on independent neuron codes are easy to construct and analyze. However, it is also possible to encode the information by correlations among the activity of individual neurons. An important example of correlation codes in a network is found in hippocampal place cells. In rats, it is found and also experimentally verified that place cells in hippocampus code the spatial location. Apart from the place field an individual cell responds, the phase difference between the rhythmic pattern of spiking of a neuron and collective firing is shown to encode more information [17].

As an example of network models, models based on parallel distributed processing and adaptive resonance theory will be described briefly.

Parallel Distributed Processing (PDP)

In 1980s, Rumelhart and McClelland suggested a class of models called Parallel Distributed Processing (PDP). In these models, it is assumed that information processing occurs by the interaction of a large number of simple processing elements called units, where each units send excitatory and inhibitory signals to others. A unit may represent an hypothesis, a goal or an action and interconnections among the units relate hypotheses to constraints, goals to subgoals, subgoals to actions, actions to muscles etc. PDP models are proposed as an alternative to serial or sequential models of cognition [32]. An example PDP model for typing the word "very" is shown in Figure 2.26.

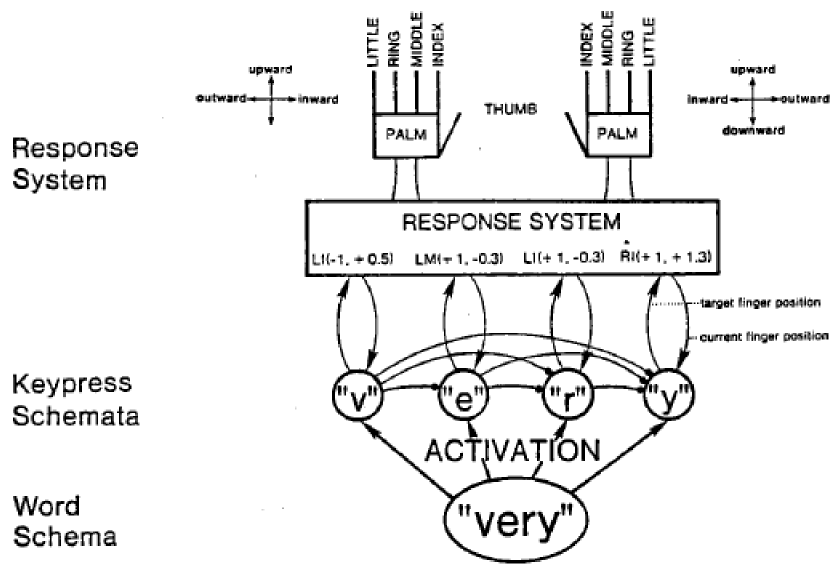

Figure 2.26: The interaction of activations in typing the word very. The very unit is activated from outside the model [32]. From Rumelhart and Norman, 1982.

Adaptive Resonance Theory (ART)

Adaptive Resonance Theory (ART) is proposed by Stephen Grossberg in late 60s, and the first explicit study appeared in Biological Cybernetics at 1976. ART is based on classical conditioning paradigms and it has a solid mathematical background based on dynamical systems. However, the dynamical system has analytical solutions; thus, it is possible to implement the system to work in real time without solving the differential equations [33,34].

Grossberg considers the fundamental problem of perception and cognition in the following questions; “How humans discover, learn and recognize invariant properties of the environments to which they are exposed ?” and “How do we continue to learn throughout life ?”. We still do not know any general artificial system capable of discovering invariant principles of the environment and learning continously without forgetting the past [33–36].

Grossberg reinterprets these questions as stability-plasticity dilemma. In this dilemma, plasticity is the learning ability of the system, whereas stability is the ability to buffer learned information from irrelevant inputs. ART is a self organizing neural network where the invariant properties are emerged as recognition codes while the codes are also stabilized and scaled accordingly. It is proposed as a model to work in real environment with real inputs. Thus, it can work in response to an arbitrary temporal sequences of arbitrarily many input patterns of variable complexity [33–36].

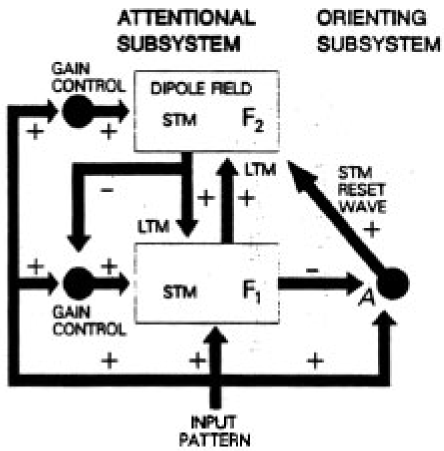

Figure 2.27: Anatomy of ART I network [36].

The first type of ART nework, which is called ART I, is shown in Figure 2.27. It is composed of two layers; the input layer F1 and the output layer F2. F1 and F2 is composed of discrete neurons, and F1 neurons are connected to the input. Also, each F1 neuron, in a bottom up fashion, is connected to all neurons in F2. F2 is a type of winner take all network where each neuron have recurrent inhibitory connections to the other neurons in F2 and top down connections to all neurons in F1. There are also two control mechanisms; gain control and orienting subsystem. The gain control decides if it is time to compare the features of the representation in F2 with the features of the input in F1, and the orienting subsystem decides if the representation in F2 is correct

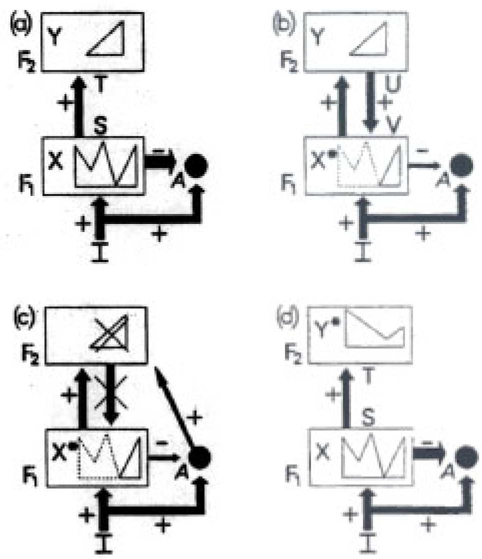

Figure 2.28: Search process in ART I network [36].

or another search process should be invoked.

The working steps of ART I is shown in Figure 2.28. First, input I is presented to network and it causes an activation in F1 called X, and F1 activity causes an F2 activation called Y. Then, F2 activation is send back to F1. Thus, F1 activation can be interpreted as a metric of matching between the activation caused by input I and the activation caused by representation Y in F2. Then, orienting subsystem checks if this match is compatible enough. As shown in (c) in Figure 2.28, the orienting subsystem can reset the F2 representation in order to reinitiate another search process. If a compatible match is found, the representation of input I is the activity exist in F2.

3. PLANNING

3.1 Overview

An important ability of animals is to choose appropriate actions in a goal directed manner in a changing environment. This ability also depends on the evolution since the higher an organism ranks in the evolution chain the more complicated high level cognitive ability it has. Goal directed behaviour is actually a broad term covering decision making, planning, problem solving and so on. A common point of all these is that all of them depends on high level cognition which is an integrated function of brain especially involving prefrontal cortex. Details of high level cognitive functions are not understood because physical structure, neural connectivity and computational principles of neural representations and network structures are not well-known. Studying high level cognitive processing using invasive and non-invasive research techniques is not easy because of integrated nature of high level processes involving prefrontal cortex which is considered to be silent in hundred years before.

A famous anecdotal story about planning is reported by Penfield about the behavior of his sister several months after the removal of her right frontal lobes. She had planned to prepare a meal for a guest and her family. Although everything she need was ready in the kitchen, she was not able to prepare the meal and confused by her long running effort. After the surgery, there was no change in her personality or other cognitive functions, however, she had a defect for planned behavior [37].

Basically, decision making is the activity of deciding on a choice and taking an action which is satisfactory to the organism in a way. The context of decision can be as simple as pressing a button or emotionally complex such as relocating to another city. However, a decision is made to achieve or satisfy the goal in either case. In this view, any paradigm involving a goal, which is either given to or self-generated by organism, and an action, which is taken by an organism, can be regarded as decision making. Therefore, it can be said that the simplest decisions are made according to instrumental conditioning. As previously explained, in instrumental conditioning, a specific action among a set of actions is chosen since the most satisfactory reward is expected to follow that action. The goal in this simple behavior is to get the most valuable reward.

In this perspective, instrumental conditioning can be considered as an action choice problem, where the aim is to maximize the reward by performing the most valuable action. The value can be an external reward such as a drop of juice or an internally generated state, e.g. an emotional state, the action brings. There are two types of action choice problems [17]. One is like instrumental conditioning where the reward immediately follows the action and it is called "static action choice problem". An example of static action choice is foraging. The other type of action choice problem is called "sequential action choice problem" and it is like instrumental conditioning with delayed rewards where the reward is partly or wholly delayed until a specific sequence of actions is completed. Thus, in sequential action choice problems, action selection must be based on future expectations rather than immediate reward.

In classical conditioning, animal responds to a stimulus, so the response is stimulus driven. On the other hand, in instrumental conditioning and in many real world settings, animal self-generates a respond to get a reward from environment. It seems it is possible to think the process of self-generating response as a response to internal stimulus or state. A reward, as it is mentioned, can be a concrete entity or something conceptual like optimality argument. The optimality argument is a term in behavioural ecology which argues that animals behave as efficient as possible to survive and breed that leads them to adopt behavioural strategies which ensure maximal reproductive success [29]. However, optimality argument is in controversy (More information can be found in Gould and Lewontin (1970) and Stephens and Krebs (1986)).

In this chapter; (1) a few decision problems in engineering and neuropsychological framework is summarized, (2) computational, psychological and neurological aspects of decision making is covered, and then, (3) an approach to cognitive planning is established.

3.2 Principles

3.2.1 Decision Problems

Five well-known decision problems; the pole-balancing, foraging, maze task, the Tower of London task and the Iowa Gambling task are presented here. The problem which is the focus of this study, the Tower of Hanoi, is presented in the next section.

The Pole-Balancing Problem

Studied by Michie (1967) and Barto et al. (1983), the pole-balancing problem (Figure 3.1), or the cart-pole problem, is a well-known example of reinforcement learning. As shown in Figure 3.1, the cart is free to move forward (or right) and backward (or left) on a one dimensional track between fixed two end points and the pole is attached to the top of the cart where it can freely move in the vertical plane aligned with the track. There are two possible actions, pushing the cart either to right or left by constant force. There are also two aims; (1) keeping the pole balanced vertically, (2) keeping the cart from crashing to the ends of the track. Since there is more than one action to satisfy the goal, the pole-balancing problem is an action choice problem.

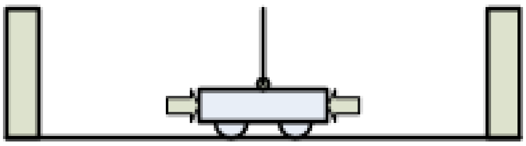

Figure 3.1: The pole-balancing problem.

Foraging

Animals obtain food by searching, which is called foraging. In real world, food is distributed non-uniformly among patches. Therefore, an animal should search for a patch of food, exploit the patch until the food is getting hard to find, and then search for a new patch. When a new patch is found, it is easy to find food in the patch. However, it gets harder when food is exhausted in that patch. At a point, animal should make a decision of leaving the patch and searching a new one. This is called patch-leaving problem. Since the animal gains energy from food and spends energy while searching, it is accepted by many that the long term average of energy gain should be maximised [29].

An example is bumblebees foraging on flowers. A neuron, which is identified as VUMmx1, is projecting widely to brain regions involved in odor processing in the honeybee suboesophogeal ganglion and it is shown that this neuron delivers information about reward. A model of bee foraging has been proposed based on neuromodulation [38].

Maze Task

A maze is a tour puzzle having branches with choices of path and direction, whereas a labyrinth has an unambiguous route and is not designed to be difficult to navigate. Mazes are often used in experiments to study spatial memory and learning. A well known example to explore hippocampus in the formation of spatial memories is water maze task developed by Morris in 1984 [5].

A maze can be constructed such that end points contain different amounts of rewards. Such a maze is suited for studying sequential action choice behavior since the end points can be reached by a sequence of actions and the reward is fully delayed. In Figure 3.2, animal should decide to turn either left or right at decision points A, B and C. Then, it receives reward (or no reward) when the end point is reached after two successive decisions are made (A and B or A and C).

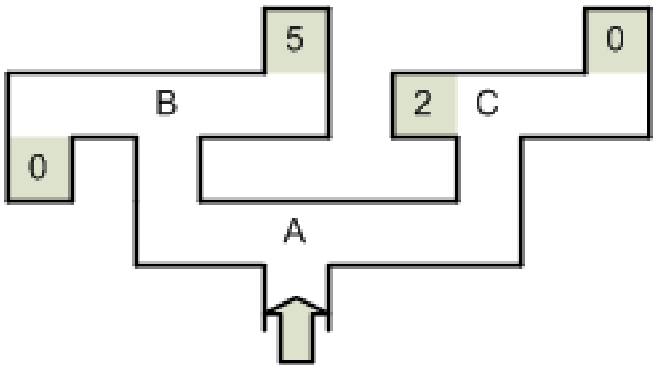

Figure 3.2: A sample maze. There are three decision points (A,B,C) and four end points (shaded) with rewards (0,5,2,0). Reproduced from [17].

The Tower of London Task

In order to study planning, Shallice (1982) developed the Tower of London (TOL) task (Figure 3.3), which is a variant of the Tower of Hanoi puzzle. The task consists of three colored discs (or balls), three pegs (or bins) each having different capacity (either 1,2 or 3), a start state and a goal state. Participants are required to achieve to the goal state from start state by planning how to move the balls in the minimum number of moves [39,40].

Although it is a variant of the Tower of Hanoi task, they are not isomorphic because the Tower of Hanoi task has another constraint on arrangement of discs. Moreover, an exact value function can be constructed for the Tower of London task that measures

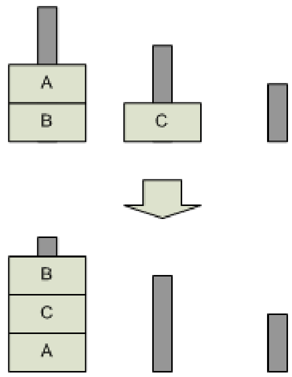

Figure 3.3: The three-disc Tower of London task. Reproduced from [39]. From Shallice, 1982.

the distance from the goal by counting the discs not in the goal position [41]. However, a simple value function cannot be constructed for the Tower of Hanoi task, because, in order to complete the task, a move that conflicts with the goal must be made.

The Iowa Gambling Task

The Iowa Gambling Task is designed by Bechara and refined by Damasio and Anderson. It is especially used by Damasio in support of Somatic-Marker hypothesis. In this experiment, the subject is given a loan of $2,000 and the subject is told to lose as little as possible of the loan and try to make as much extra money as possible. The task is performed by selecting cards from four decks labeled A, B, C and D. The subject does not know when the task ends. By selecting and turning the cards from decks, the subject always earns money and sometimes pays some money indicated on the card to the experimenter. The task is constructed such that cards selected from deck A and B pays $100, however, unpredictably, some cards require the subject to pay high amounts as much as $1250. The cards in deck C and D pays $50, however, the subject is required to pay less on average [42].

During the task, the normal population starts with sampling each deck, choose deck A and B at start, but move to deck C and D and stays on this decision until task ends. People with high-risk taking behavior occasionally sample a card from deck A and B. However, patients with ventromedial frontal lesions typically perform just opposite of normal population. Although they describe their selves as low-risk person and they know which decks were bad when asked, they turn more and more card from deck A and B. Moreover, this behavior seems to be persistent when the task is repeated after some months [42].

3.2.2 Computational Aspects

Markov Decision Process (MDP)

The Markov decision process is the formal model of action choice or decision problems. A Markov decision process consists of four elements:

- State space: S
- Action function: A
- Transition function: T
- Reward function: R

In a finite, discrete-time Markov decision process; S is a finite set of discrete states, A is a finite set of discrete actions for all x in S and all events occur in discrete-time. The state at time $t+1$ depends only on the state and the action taken at time t. This is described in the transition function T, which transforms the system from $s(t)$ to $s(t+1)$. A reward is also received at each time step depending on the state and the action taken and it is defined by the reward function R [29].

A system is said to have the Markov property, if the next state of the system depends only on the current state and the action taken. In a Markov decision process, both transitions and rewards can be probabilistic, however, they depend only on the current state and the current action. Therefore, Markov decision processes hold the Markov Property. Markov property of the system do not have to be an intrinsic part of the system, it can be the part of the state-space model of the real process [29].

Policy and Return

A living animal should decide what to do in each state. A policy is a mapping between states and actions such that there is an action for each state. Thus, following a policy an action can be selected for a state. If a policy always maps a state to the same action, it is said to be stationary. If an action is chosen according to a probability distribution each time a state is visited, the policy is said to be stochastic [29].

According to a policy, an action is taken and a reward is received. However, a natural behavior requires that rewards over a period of time should be maximized instead of an immediate reward received. In this view, return value can be defined according to; (1) total reward, (2) average reward, (3) total discounted reward [29].

Credit Assignment Problem

As noted by Minsky in [43], some problems have a definite success criterion such as a game is won or lost as in chess or checkers, and a credit or blame should be properly assigned to each of the internal decisions or actions performed which have contributed to the success or failure of the problem. Since internal decisions affect the actions, and actions affect the environment, hence outcome, the problem is two-fold. The first subproblem, called temporal credit-assignment problem, is to assign credits for actions which are responsible for the outcomes, and the other, called the structural credit-assignment problem, is to assign credits to decisions which are responsible for selecting the actions [6].

In classical reinforcement learning, the system learns by rewards, which is a feedback from the environment, and assigns credit to actions performed, however, the internal decisions are not credited or blamed [6].

Due to the problem of credit assignment, finding an optimal policy is not easy. Since there are a number of different possible policies available, the behavior, according to the optimal or near optimal policy, cannot be based on solely trial-error mechanisms [29].

Dynamic Programming

Invented by Bellman in 1940s, dynamic programming is a method of finding the best decisions to solve a problem. With the help of constructing an evaluation function, which is also known as value or return function, dynamic programming is used to solve the credit assignment problem. The steps of computation in dynamic programming is; (1) computing the evaluation function for the current policy, (2) computing action-values with respect to the current evaluation function, (3) improving the current policy by choosing actions with optimal current action-values. This procedure can be done successively as in policy-improvement method, or concurrently as in value-iteration method. Independent of the method used the optimal solutions for the policy and the evaluation function can be found since there is no local maxima. However, in many real world problems, the state-space is either large or only a part of it is available and in order to find the optimal policy, the entire state-space should be examined [29].

Reinforcement Learning

The reinforcement learning is a term introduced to AI and engineering by early works of Minksy (1954, 1961). Although this term is not used in psychology, the idea stays on the same ground with the study of reinforcement and instrumental conditioning. The reinforcement learning is typically considered associative, thus, stimulus-action associations are learned observing the stimuli and reinforcement [6]. A typical learning system shown in Figure 3.4 performs while receiving stimuli and a reinforcement signal from the environment. The learning system consists of a critic, which credits or blames actions, and an actor to perform actions. Since the critic can only judge actions, but cannot provide alternative/better actions or guide the decision-making process, it only contributes to the temporal credit assignment. On the other hand, the actor can refine its decisions, thus, it contributes to the structural credit assignment.

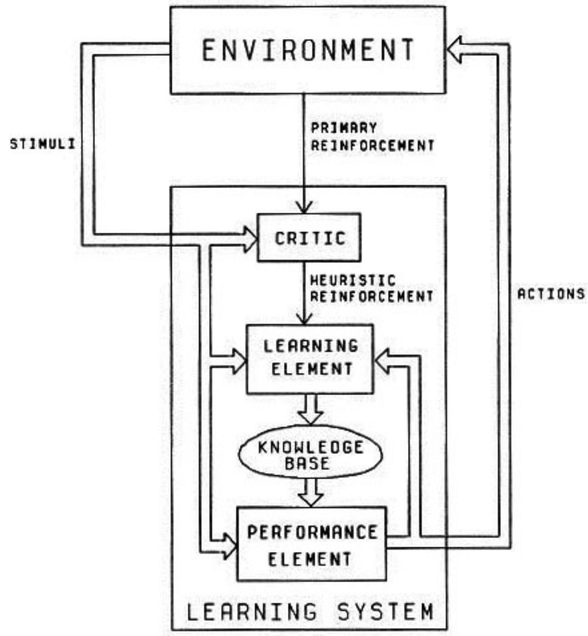

Figure 3.4: Information flow among components of a learning process [6].

Associative Learning

The goal directed behavior is to perform the most suitable and appropriate action to achieve the goal. In the most straightforward manner, stimuli and actions are learned in pairs which is called associative learning, thus, the goal can be reached by a single action. An important issue in associative learning is constructing a metric for similarity, which is required generalization and discrimination. Generalization is the transfer of association from one stimuli to another when they are similar, and discrimination refers to showing different actions to dissimilar stimulus [6].

There are different approaches to generate associations. In the independent-associations approach, where the similarity, generalization and discrimination is ignored, a distinct association is made between each stimuli and an action. The second type of approach is the stimulus-sampling-theory (SST), which include similarity and generalization issues. In SST, each stimulus is a subset of independent stimulus elements and each stimulus element is associated with a particular action. An action associated with a stimulus element is selected if the stimulus element is the one selected randomly in all stimulus elements composing stimuli. The third is the linear-mapping approach which is widely used in engineering. In this approach, stimuli are represented as vectors and each element of the vector codes the presence or absence of a particular stimulus element in an all-or-none fashion. The inner product of the vectors representing different stimulus is the degree of similarity of them. The fourth approach selects action of the prototype that best matches to the stimuli presented. The linear-mapping approach is better than either SST or the independent-associations approach in some respects, however, it is not possible to form nonlinear associations or arbitrary categorization of linearly-dependent stimulus vectors [6]. The problem of all these approaches are emphasized as knowledge-representation problem in AI.

3.2.3 Psycological Aspects

Problem Solving

The standard theory of problem solving (Newell, Shaw & Simon, 1958) involves the mental inspection and manipulation of representative structures. Newell & Simon (1976) proposed physical symbol system hypothesis and stated that symbolic processing is a necessary and sufficient condition for intelligent behavior, which precedes another claim namely problem space hypothesis which involves search through a space of candidate states generated by operators [44].

Newell & Simon (1972) in their information processing theory define the key notions of cognitive system, task environment and problem space. The cognitive system includes sensory receptors, motor effectors, memory systems and central processing unit, the task environment includes the problem itself, and the problem space includes the states, operators, evaluation functions and search strategies [44].

Newell & Simon's (1972) means-end analysis is consistent with this view. Means-end analysis is;

- Calculation of the difference between current and desired state
- Selection of an operator to reduce the difference

- Application of the operator or creation of a sub-problem to transform current state into ones in which operators can apply

Subjects in novel environments usually apply backward chaining, whereas experiment results forward chaining [44].

Control of Behavior

According to Watkins, the behavior is controlled in two modes; (1) look-ahead, (2) primitive. In look-ahead mode, an internal model of the world is used and consequences of actions are mentally simulated before performing in real world. He summarized the essential abilities for look-ahead control as; (1) a transition model, (2) a reward model, (3) an ability to consider states other than current. In principle, having these abilities, the best course of actions can be computed. In addition, if an approximate value function is available, the computational requirements of look-ahead control decreases and the possibility of finding the best course of actions increases. With an exact value function, it is possible to find the best course with a one step look-ahead [29].

When future states are not considered, behavior can occur by primitive control. Although looking-ahead is better to decide actions than primitive control, even higher animals can behave primitively due to stochasticity of environment, ambiguous reward schedule (value function) or a need for quick action. Primitive control can happen considering one of the followings: (1) policy, (2) action-values, (3) value function. It is obvious that the computational requirements for primitive control are much less than look-ahead since there is no need for a model of transition or reward. If a policy f is considered, an action $f(x)$ is performed. Thus, similar to classical conditioning, there is a memory for stimulus-response pairs. If the primitive control is based on action-values $Q(x,a)$, the action a with the maximal action-value is selected. Action-values are similar to instrumental conditioning (with immediate rewards), where the outcome (reward) of an action is stored with the action and it is selected accordingly. Finally, the action can be chosen according to a value function but not immediate reward. Here, the primitive control is based on the value function and it can be a strategy for instrumental conditioning with delayed rewards where the value function can be improved according to the course of actions selected while reaching the reward. Instead of value function, a function monotonically related to value function can be used in primitive control based on action-values or value function [29].

Since the look-ahead control is slow, the behavior can be based on primitive control but the primitive mechanism can be guided by an imitation teacher looking-ahead. Also, Laird et al. (1986) proposed “chunking” as a general learning mechanism that caches the action sequences and treats them as single actions [29].

3.2.4 Neurological Aspects

It is argued that the deficits in prefrontal cortex (PFC) affects novelty awareness, working memory capacity, subgoal selection strategy and conflict resolution between long-term goal and short-term subgoals; thus directly affects the subjects’ progression toward the goal [37,41,45,46]. Lesion and functional imaging studies show that dorsolateral (DL) PFC is important for mental act of temporal sequencing demand of planning [47]. Recent neuropsychological evidence also indicates that neurons in PFC are involved in prospective coding of anticipated reward and expected objects [47–49]. Furthermore, functional imaging has suggested that anterior DLPFC may have a unique role in maintaining goals in working memory while proceeding with subgoals [47].

As explained in the previous section, there is a need of creating representation of the environment for cognitive planning. While ventrolateral prefrontal cortex (VLPFC) takes part in forming associations between visual cues and the actions and furthermore maintain such representations in an active state until the goal is achieved, dorsolateral prefrontal cortex (DLPFC) has a role in choosing the most task relevant internal representation and in linking the short-term memory (STM) representations to goal-directed motor behavior [50,51]. Thus maintenance in working memory i.e., creating a mental scratchpad where the "inner world" is created and evaluated has been achieved through these neural substrates.

In order to identify, to process or prepare actions, there is need for information about what occurred over time, superior temporal gyrus (STG) provides this contextual information [52]. Dorsolateral prefrontal cortex and posterior medial prefrontal cortex are active in action selection while ventromedial prefrontal cortex compares ongoing actions and performance outcomes with internal goals and standards [50]. Monitoring unfavorable outcomes, response errors and conflicts and decision uncertainty activates rostral cingulate zone, meanwhile anterior cingulate cortex detects conflict between plans of action and reinforces lateral prefrontal cortex for greater cognitive control [50,53].

Left inferior frontal gyrus and anterior lateral regions of prefrontal cortex are active in building subgoals and in monitoring and management of subgoals, respectively [47, 54].

Even though, sub-cortical areas such as basal ganglia and midbrain dopamine neurons are also effective in planning here only cortical substrates are reviewed. As can be followed from above information, prefrontal cortex is the most important component of cortex during planning.

3.3 An Interpretation of Planning

Associative learning and instrumental conditioning is not an explanation when an action alone is neither sufficient to complete the task nor responsible for the result. Therefore, planning is a need when a complex “mental goal” in terms of brain resources must be reached through a sequence of actions with the help of cognitive processes. Here the term “mental goal” is emphasized since a goal such as holding a cup of coffee is probably independent of an executive mental activity. Although holding a cup of coffee is a quite complex task for motor system, it has nothing to do with high-level, executive control of the brain which is the focus of this study.

Previously, instrumental conditioning with delayed rewards and sequential action choice are mentioned to have similar properties with planning. In both of them, the action to be performed is based on the current state and the outcomes of past actions. However, there is no need to think about the outcome of an action in advance as the statistical and temporal properties of the reward schedule are learned during trials. So, the behavior is purely exploratory at start and shifts to exploitory-exploratory stage by learning since the responses are being tuned to increase the reward received. Though, it is never purely exploitory since pure exploitory behavior prevents adaptation to changes in the environment. The exploitation-exploration paradigm is well-known but not understood in terms of the brain networks and neural processing.

In contrast, planning is an ability to help decision making by conscious inner experiences which help to solve a task although there is no associative learning similar to one in instrumental conditioning. These inner experiences of the external world occur by forming the representations of the environment, generating the actions and applying these actions to the internal representation of the world in mind. For example, a maze task is a sequential action choice problem [17], which can also be viewed as an instance of instrumental conditioning with delayed rewards, where subject can neither see the maze nor any cue for the exit. Planning attributes no advantage for this task. There can be no inner experience which is helpful to decide between actions by simulating outer world in such a situation. The only way to escape from the maze is trial-and-error exploration which later becomes alternating between explotation and exploration when learning happens. Now, assume that the subject has a limited ability to see through the walls and there are cues for the exit spread across the maze. Then, the subject in the modified maze can carefully plan his next action by looking through the walls, searching for cues and finding the way to reach the cue.

By analogy, looking through the walls is super-hero replacement of inner experience we have. Instead of directly seeing the future states, we generate the actions in our mind, as if we will apply them in reality, and apply them to our inner world. Then, we sense our inner world as in reality. However, the inner world is just dreaming without the cues for the exit in the modified maze. Such cues, which are put by an intellect for us in the modified maze, are not available to us in many real life situations. Therefore, a complementing ability to simulation, evaluation is needed. So, we can compare our inner world with the goal and can say we are on the right or wrong track. So, planning helps goal directed behavior by exploiting the existing regularities in the task by performing simulation and evaluation. Based on the Watkins' classification on control of behavior, which is explained in Section 3.2.3, planning in this study is the control of behavior by looking ahead, where the task rules corresponds to the transition model and the evaluation strategy constructs a reward model.

A well-known daily problem of planning is playing table games such as chess, checkers and backgammon which require big and instantaneous effort to simulate feature state of the board in mind. These are almost only studied in terms of AI [55–59] since it is highly complex to be analyzed and traced in the brain with the current measuring and the brain imaging technology. Instead of table games, the Tower of Hanoi and the Tower of London for judging planning ability and the Iowa Gambling Task for judging decision making ability is widely used in neuropsychology literature as well as in clinical studies.

Problem solving and decision making processes are similar in their nature and both essentially refers to the same thing for goal directed behavior. The difference between them is due to the nature of the problem, time span and emotional factors attributed to them. Problem solving usually refers to solving artificial problems or small concrete problems of real life, having a short time span and not affected by emotional factors. Decision making, on the other hand, is deciding the choices in real life situations and it may span quite a long time, maybe years. Decision making is also thought to be affected by emotional factors. In either problem solving or decision making, planning is a crucial component required to solve the problems or to make decisions or generally to behave in a goal directed manner.

Only very simple tasks, those having a small state space, can be completed in a direct fashion. The others have a state space that cannot be completely analyzed as a whole with the brain which has limited resources. The limitation of resources of the brain is overcome by biological facilities and computational architectures. The brain is bombarded with many inputs coming from the environment, but there are only limited higher-order processes. That means only a subset of stimuli can be processed at a time which is selected by an attention mechanism which is out of scope in this study. Also, processing in the brain is slow due to chemical and mechanical processes required such as synaptic communication, but it features neocortical computing paradigm [14]. Another limitation is active memory capacity. Although the brain outperforms any computer we have today regarding to memory capacity, it is almost certain that the long-term memory traces are not being used in instantaneous operation. Instead, short-term or working memory is shown to have roles in memory tasks. On the other hand, decision making may span quite a long time and the previous knowledge may transfer among tasks. These indicate that the long-term memory also plays an important role in the context of goal directed behavior.

Planning occurs in an abstract problem space, where the motor skills such as moving, grasping etc. are not important. We can view sensory receptors as operators mapping the real world into internal world and motor effectors as operators mapping the internal world into real world.

When the subject faces the problem in the environment (external world, external reality), through the sensory receptors, he constructs internal representations of the external world and stores the generated representation of the world (internal world) in his memory and holds it in an active state. Internal world corresponds to the states in the problem space, so we map task environment to the problem space through sensory receptors. At this point, the executive mental process starts and manipulates the representations by applying transformations and validating it at the same time (simulation and evaluation). Selection of transformations requires evaluating the current state of the internal world according to the goal.

Once the “inner world” is formed and the subject holds it in an active state, the executive mental processes originate and manipulate the representations by applying transformations and validating it at the same time which are named “simulation” and “evaluation” respectively. The selection of transformations requires evaluating the current state of the inner world according to the goal. Moreover, a complex task can be simplified if its goal can be divided into subgoals. Then, each subgoal represents a new problem. In this way, when possible, a task can be structurally decomposed into others. Accomplishing the subgoals brings the accomplishment of the goal. Subgoals may have a temporal order in which they should be done accordingly. Online storage required for planning, to hold generated states, subgoals etc., is a part of working memory, thus executive processes run on working memory. For some tasks, the memory requirements are relaxed since the external world can function as an auxiliary memory [44].

During planning various settings are tried out, "looked-ahead", if any of them satisfies the goal or subgoal it is immediately applied. If no goal or subgoal can be satisfied, the action that transforms the environment to the situation where it is probably closer to the goal or subgoal is selected. Thus execution and reasoning is interleaved during planning [44].

There seems to be a consensus on the definition of decision making, but such a consensus is not available for planning. Goel & Grafman define planning as charting a course from point A to point B in a modeling space without bumping into the world [37]. Changeux & Dehaene define planning as "the goal directed, trial-and-error exploration of a tree of alternative moves." [41]. Some other authors failed to differentiate between constructing and executing or following a plan [37]. Knight & Grabowecky explain generating internal models of external reality, which is called "Simulation", and processes that monitor sources of information to develop accurate models of the form and context of internally and externally generated information, which is called "Reality Checking". Simulation and reality checking can be considered supervisory, executive or organizational functions [60].

In the light of all above, I define planning as a high level cognitive process, which is a part of decision making, and as an improvement over primitive control which helps decisions by making considering feature expectations possible. Thus, to construct a plan, one needs an ability to foresee or simulate a future state of our world by employing a mental set of simulated actions which are also applicable to real world and select the sequence of actions that is valuable which means they bring world to a more beneficial state. At this point, it can be speculated that the abstract processing capability required for planning is made possible by the large size of human cortex when compared to other primates. Moreover, Descartes' argument can be modified as follows; animal behavior is solely based on current or past experiences whereas human behavior is also shaped according to future expectations which can be the reason behind the cultural and social history of man.

The types of behavior, in this context, can be categorized as follows. First, The uncontrolled (involuntary) behavior occurs in the form of a reflex. Reflexive behavior is probably first observed by Fernel and characterized by an act occurring without the action of will or any other directive of the mind [3]. A reflex can be conditioned to occur in response to a stimulus (classical conditioning). In some types of classical conditioning, a response, which is not a reflex, is elicited in response to a stimulus given.

Second, a response can be based on the current stimulus considering the past results of outcomes. This is similar to above, however, the stimulus-response pairs are learned according to the outcomes of previous responses as in instrumental conditioning (with immediate rewards).

Third, a different behavior is required for goals which can only be achieved by performing a sequence of actions. For example, in a maze task, the exit can only be found by performing a correct sequence of steps. Since there is no cue or reward after each step, a response like turning right or left cannot be associated with the stimulus. Instead, a sequence of steps, e.g. turn left, then right, then left, is associated with the task. This category includes instrumental conditioning with delayed rewards.

Fourth, the computational requirements of the third task can be relaxed if subject can wholly or partially see its environment and an evaluation process is available. If so, instead of trying all possibilities, the planning process including the simulation and evaluation can guide the behavior. This category includes the tasks that can be planned. For some tasks in this category, the goal is too far and it can be divided into subgoals. Then, each subgoal represents a new problem. In this way, the original task can be accomplished by accomplishing sub-tasks.

Based on the interpretation of planning, a framework in Figure 3.5 is considered for a coarse model of planning. Environment is the external world with which subject is interacting and it includes task related objects and information. In this study, neither learning the task rules and representations nor transfer of knowledge from one type of task to another is taken into account.

In this framework, behavior is displayed through a series of interactions between sensory, motor and executive processes. Sensory and motor processes are considered to be passive such that they do neither modify the information nor take active role during planning. So, neither sensory representations nor motor commands has a specific, implicit role, thus planning is considered to occur in an abstract domain [44]. The invariant properties of the task and the domain is used during planning [14].

According to the serial interaction pattern, sensory (low-level) processes gathers information from the environment and transfers it to (high-level) executive processes as a neural code in a domain independent, invariant form. During sensing, it is

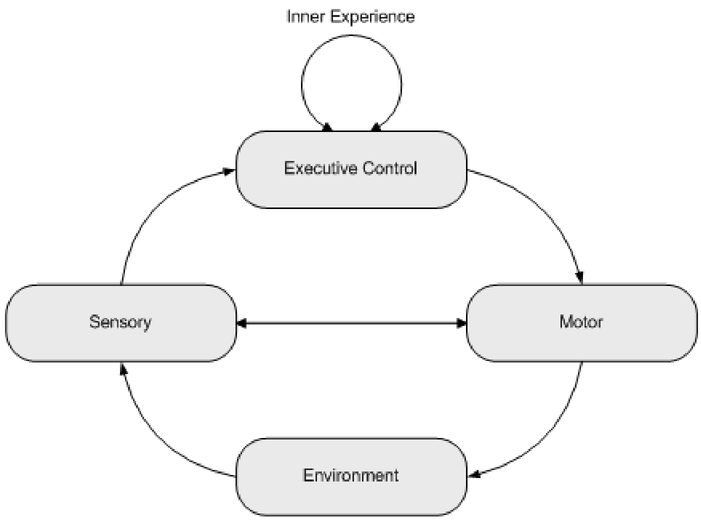

Figure 3.5: Framework

assumed that an attention mechanism is initiated by executive processes such that only a subset of stimulus coming from external world reaches to higher order sensory areas. Executive process accomplishes the planning by simulating and evaluating the current environment. It then commands to motor process in order to perform the action or actions. Motor process applies the actions requested by executive process through motor effectors and the behavior occurs physically.

The focus of this study is the executive control in the framework. The proposed model, which represents the executive control in Figure 3.6, compromise of spatially and temporally organized memory units and processing units. Memory units hold the goal, subgoal, the simulated state representing the inner world and sequence of actions planned. Processing units include subgoal generation, action generation, action selection and evaluation (reinforcement generation).

The current state of the world is assumed to be always available to the subject and the goal is learned as a long term objective. The subject has only one aim that is to transform current state of the world into the one that is specified as the goal. If the problem is complex enough, subject accomplishes the goal by searching the problem space partially by dividing the goal into subgoals. Subgoal generation is accomplished at most obvious and explicit terms. In this work it is proposed that subjects follow a very simple strategy that is enough to derive the subgoals at any time as long as the goal is available. Not like [37,45], generation of a subgoal must be such a coarse milestone that its generation is not be affected by conflict resolution problems.

The discrete steps, called actions, required to solve the task is generated by action generation unit. Those actions are selected by action selection unit according to the

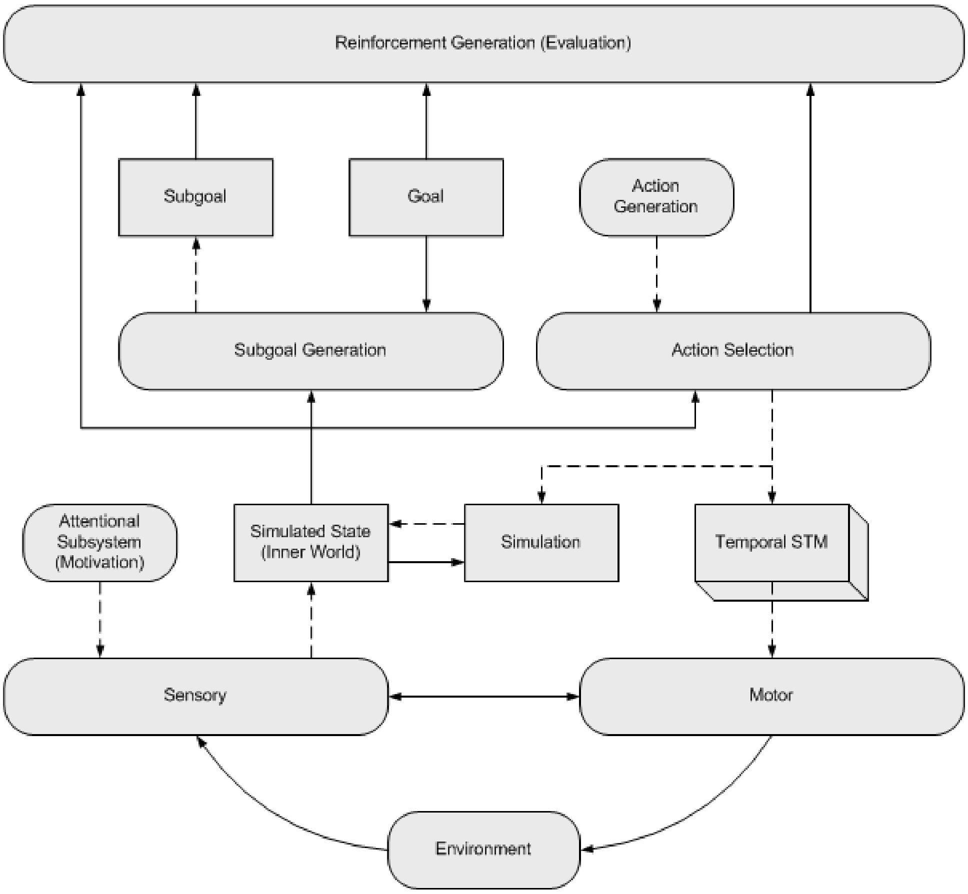

Figure 3.6: Proposed Model for executive control

task constraints.

Evaluation or reinforcement generation is accomplished by monitoring the overall activity of executive processes. The selected actions are validated or deleted according to the task demands. Simulated state is reinstalled to external world when the subgoal is not satisfied with the current action sequence. This triggers subgoal generation if the goal is not satisfied and triggers application of the plan if subgoal is satisfied. When a subgoal is not satisfied, a set of actions are selected according to the subjective evaluation criteria depending on the task. The dashed arrows in Figure 3.6 represents pathways affected by evaluation process.

Conflict resolution exists at simulation stages, e.g. biasing toward a group of actions, due to deficits or due to implicit or explicit conditioning. How learning takes place during planning and performing actions according to plans is not considered. Deriving high-level concepts for an encountered problem is a specialized learning process and it is acquired by the subject's experience during the task evaluation and/or by subject's previous experience in various related events. So the proposed model does not include the learning phase but it illustrates how solving non-specialized problems is actually possible by planning as long as domain knowledge about the problem is satisfactory.

4. THE TOWER OF HANOI TASK

4.1 Overview

In the previous section, an interpretation of planning is given and a framework with a model of executive control is presented. In order to investigate the characteristics of the model, it is implemented and simulated for the Tower of Hanoi (ToH) task, which is widely used as a planning task in neuropsychology literature as well as clinical studies. Then, the simulation results are given in comparison with the other two models in the literature and with experimental findings.

There exist some mature frameworks for human cognitive architecture like Soar [61] and ACT-R [62] which are based on production systems. Instead of cognitive architectures, the proposed model for executive control in this framework is designed from the beginning to understand information flow to trace the process of planning.

In recent years, two models of planning based on an hierarchical neural network [41] and a production system [45] are proposed. In the first study, planning is defined as "the goal directed, trial-and-error exploration of a tree of alternative moves" and as a planning problem, they have chosen the Tower of London task [41]. The hierarchical neural network model proposed in that study includes three levels; (1) gesture level, (2) operation level, and (3) plan level. In contrast to this study, gesture level provides sensory-motor coordination by pointing to the beads on the pegs. Similar to this study, the current state and the goal state is represented as a model of the task. Since the Tower of London task allows a simpler evaluation model e.g. sum of remaining goals, the model's decisions are continuously monitored according to this evaluation. As a conclusion, it is shown that the model shows increase in error rate and solution time as problem gets harder, like human subjects. Also, to simulate prefrontal lesions, they prevent activation of plan units and show similar behavioral patterns as human subjects with prefrontal deficits [41].

The second study proposes a computational model based on CAPS production system for frontal lobe dysfunction based on the Tower of Hanoi task and its relation with working memory [45]. They have calibrated and fit the model parameters to the performance of 20 normal subjects and 20 frontal lobe patients. Thus, they directly compare their result with the patient data. In contrast to this study, they allow subjects to undo their moves meaning they do not count undo moves. They also assume there are two subgoals; clearing a disc and clearing a peg. These are, however, subgoals, and there is no evaluation concept similar to this study. They argue if there is more disks than pegs there is a need for interrupting current subgoal and do another to complete the first. So, to complete second subgoal, a backward move that invalidates the first subgoal should be done. They call such a situation as goal-subgoal conflict. Goel & Grafman concluded that patients' poor performance cannot be attributed to the patients' inability to plan or look ahead but it can be explained in terms of working memory capacity and goal-subgoal conflict resolution. During simulations, they have manipulated two parameters, move times and working memory activation. They assign time units to the steps of the program implemented and scale to fit for the patient simulations. Working memory activity constraint is simulated by decay on working memory holding moves and subgoal where goals have quite a longer decay. Finally, they have found similar results comparing to data based on performance of normal subjects and subjects having prefrontal deficits [37,45]. They do not believe that the Tower of Hanoi task is an ideal problem to study planning but an interesting problem to study working memory capacity in complex problems.

Both studies above seem to be solely descriptive models. Due to that, both calibrate and fit model parameters according to actual performance data. The first model is a neural network, however, it is not constructed on a biological basis but it only uses neurons as a computation units.

Since planning is considered to occur in an abstract domain, neither sensory representations nor motor commands play a role in the proposed model. The implementation, at the moment, is based on symbolic processing of domain invariant codes and it is done in computer software written by the author. Since it is at systems level, it does not include computation in a biological context using neural networks. The next step of this research is to implement the framework based on a biological architecture employing learning and using functional units similar to neocortical organization.

As shown in Figure 3.6, sensory processes derive information from external world. For example, in the Tower of Hanoi task, the spatial properties of pegs and disks, such as relative position of disks and pegs and size of disks are assumed to be constructed in an invariant form at the sensory level [14]. Similar to sensory processes, motor processes are also passive during planning. They transform the environment according to the conceptual, domain invariant, state using primitive operations such as moving a disk between two pegs to motor commands such as reaching, grasping, holding, moving, putting etc. in coordination with sensory system [14].

In this section, (1) detailed information about the Tower of Hanoi task is given, (2) implementation of the model is explained, and (3) simulation results and conclusion is presented.

4.2 About the Task

The Tower of Hanoi is a well known puzzle especially in computer science, since it is a perfect problem to explain recursive programming. Historically, the Tower of Hanoi is invented and marketed by a French mathematician François Édouard Anatole Lucas under the nickname N. Claus de Siam, an anagram of Lucas d'Amiens, since Amiens is the town he was born. He is also famous due to the formula he gave for finding the n^{th} number in the Fibonacci sequence [5]. His interest in recreational mathematics resulted in a four volume work "Récréations Mathématiques" which has become a classic [63].

Figure 4.1: From the box cover of the original Tower of Hanoi puzzle [64].

The classic puzzle consists of three pegs and at least three discs each of which is different in size. All of the discs are initially stacked on random pegs and the purpose is to stack all the discs on a specific peg in descending order in size by moving each disc from a peg to another. The rules of the puzzle are quite simple. One can only move one disc at a time without holding any other discs, and it is illegal to put the larger disc on top of a smaller one. For the recreational use the puzzle consists of 5 to 8 discs.

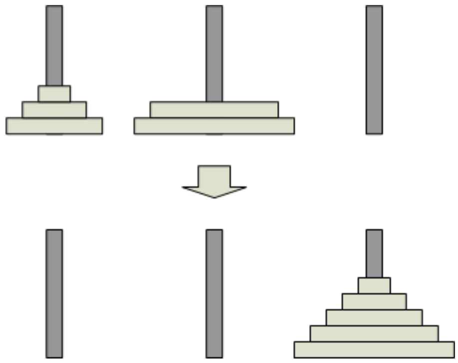

Figure 4.2: An example of Tower of Hanoi Puzzle

There is also a legend about an Indian temple with a large room containing three posts surrounded with 64 golden discs. According to the legend, when the last move of this puzzle completed, the world will end. Since the count of optimal moves to solve the Tower of Hanoi is exponential on the disc count, billions of years is required to solve the legendary 64 discs puzzle even with a speed of one move per second [5].

The optimal solution of the Tower of Hanoi takes $2^n - 1$ move where n is the number of discs. The puzzle with any number of discs can be solved with a simple recursive implementation below in Java. However, since the complexity of algorithm is exponential, finding a solution for a puzzle containing lots of discs is computationally impossible even on super computers [5]. Also, the recursive algorithm needs a huge amount of memory for such a puzzle.

```
/**
 * Prints the moves required to solve regular tower of hanoi puzzle
 *
 * @param n number of disks
 * @param src initial peg of disks
 * @param dst target peg of disks
 * @param tmp other peg
 */
public static void hanoi(int n, int src, int dst, int tmp) {
```

```
    if (n == 0) return;
    hanoi(n-1, src, tmp, dst);
    System.out.println("Disk " + n + ": " + src + "->" + dst);
    hanoi(n-1, tmp, dst, src);

}
```

The optimal solution to the Tower of Hanoi with 4 or more pegs is still an open problem [5].

4.2.1 The Tower of Hanoi as a Neuropsyhological Test

The Tower of Hanoi (ToH) puzzle is a planning task that is used to assess the planning ability of the prefrontal cortex (PFC) [37,45,47]. To successfully complete the task, the subject should be able to evaluate a correct temporal sequence of moves, which transforms the initial state to the goal state. To construct correct temporal sequence of moves, subjects need to "look ahead" and solve the problem in their heads, before physically moving a disc. The correct temporal sequence is constructed by generating internal representations of possible moves according to the rules of the task and organizing the moves in a temporal sequence by selecting some moves. Generating the internal representations of possible moves and sequences requires working memory, whereas organizing the correct temporal order and the strategy of selection of moves toward the goal requires planning ability. It is argued that the deficits in PFC affects novelty awareness, working memory capacity, subgoal selection strategy and conflict resolution between long-term goal and short-term subgoals; thus directly affects the subjects' progression toward the goal in ToH [37,45,65,66].

Moreover, there is a controversy if the ToH task actually measures the planning ability or it solely depends on the working memory. Goel & Grafman (1995) argued that "...the ability to look ahead is neither necessary nor sufficient to solve the Tower of Hanoi puzzle." "...because you can look ahead all you like, but unless you see the trick, the counterintuitive backward move, you won't solve the problem." [37] The controversy and their argument is discussed later in conclusion.

4.3 Implementation

The proposed model for executive control is adapted to investigate the behavior of normal subjects and subjects with prefrontal involvement during the ToH task. This is done by implementing subgoal and action generation units, action selection and evaluation units, and representations for the ToH task, while leaving the general model untouched. Using this realization, generating possible moves, imagining moves mentally, holding these moves in memory have been fulfilled and simulations are carried out.

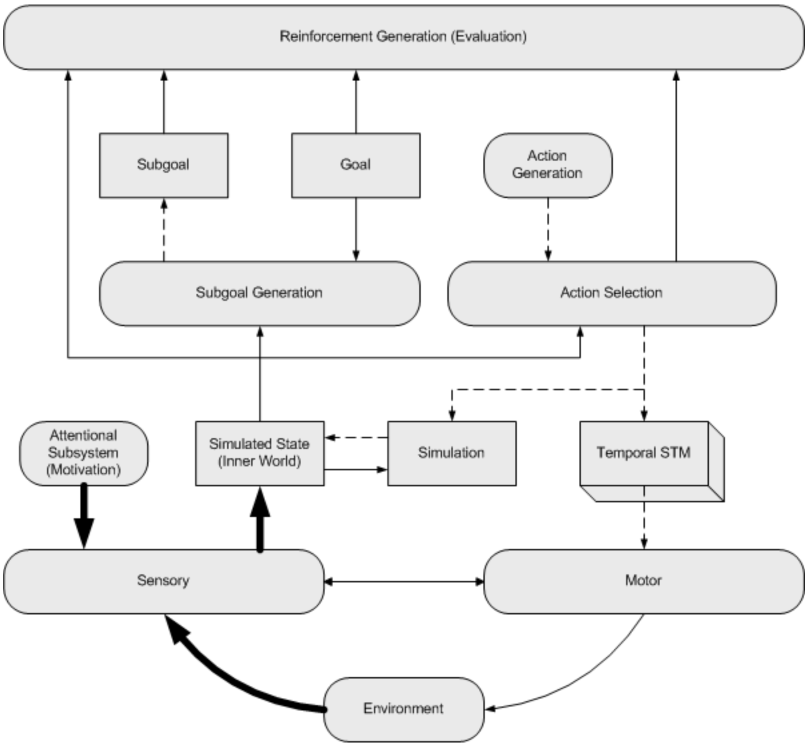

Figure 4.3: Sensing step.

Specifically, pegs are represented with number identifiers (1, 2, 3) and discs are represented with letter identifiers (A, B, C, D, E). Goal is stored as a target peg identifier, which is a number (one of 1, 2 or 3). Subgoal is represented as a target disc identifier which must be put on the target peg, which is a letter (one of A, B, C, D or E). Moves are represented as disc and peg identifier pairs, which consists of a letter and a number, e.g. (A,2) which means moving disc A to peg 2.

The model is implemented based on symbolic operations like string matching and stack operations in Java language. There is no reason to implement the model as a neural network model since the operations are quite simple. However, a meaningful model for cognitive science research should be based on the organization of cortical networks, memory and learning mechanisms as briefly summarized in previous sections.

The implementation includes; (1) an initialization phase where the goal of the task is

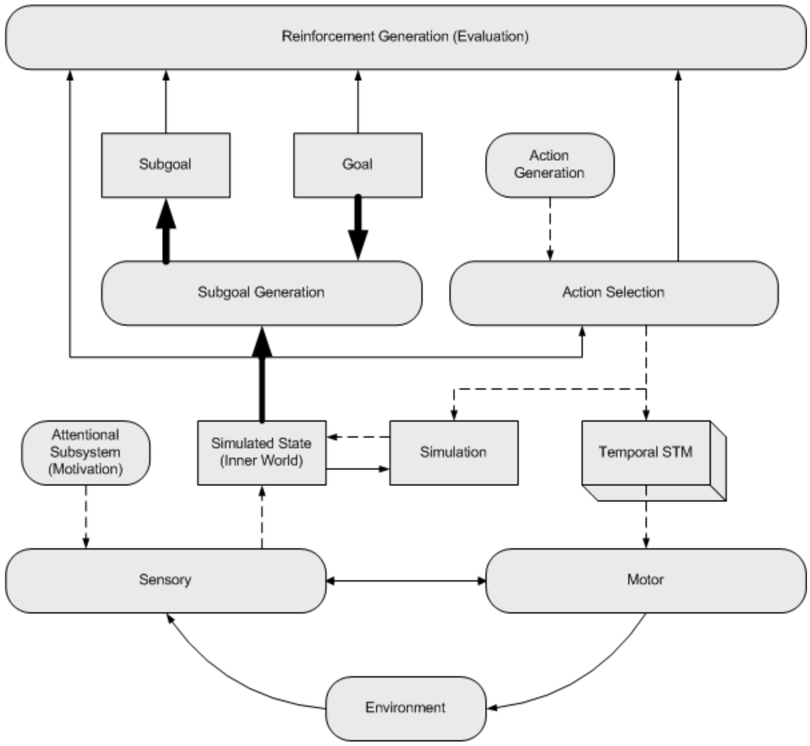

Figure 4.4: Subgoal generation step.

stored in the goal memory unit, and (2) a running phase where the Sensory-Executive-Motor loop is executed until the goal is partially or fully satisfied or the model is exhausted due to time or maximum move count constraints. The Sensory-Executive-Motor loop is realized in six steps:

1-Sensing: A sample of environment is used to derive an invariant spatial structure of task. Although it is important, top-down attentional modulation is not implemented and the model is assumed to be working in an isolated environment without any distractors. Spatial information about the task is stored in the simulated state. (Figure 4.3)

2-Subgoal Generation: A subgoal is generated and stored in the subgoal memory unit with respect to the goal and the simulated state. (Figure 4.4)

3-Action Generation and Selection: Actions are generated, selected and being validated on simulated state. (Figure 4.5)

4-Simulation: The current state is advanced to the next state by simulating the action selected. (Figure 4.6)

5-Evaluation: The current state is evaluated by comparing it with the state which satisfies the subgoal. (Figure 4.7)

6-Acting: Planned sequence of actions are applied to the environment. (Figure 4.8)

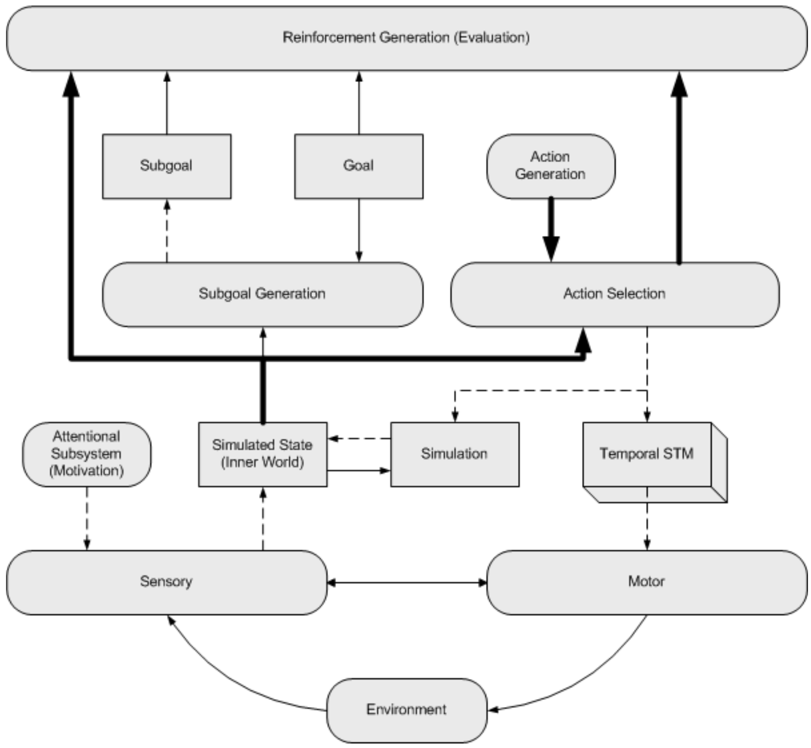

Figure 4.5: Action generation, selection and validation step.

The model interacts with the environment only in the first and the last steps. An internal model of the task is constructed in the first step (sensing), after that, a subgoal is generated according to the goal and the current state of the internal model of the task (subgoal generation). The subgoal is also represented as a model of the task in a unique state where subgoal is accomplished. Then, a repeating procedure starts and continues until the subgoal is satisfied. In this procedure, actions are generated and selected, then validated according to the task rules. If the action is valid, it is stored in memory, otherwise, another action is tried. Then, the internal model of the task is modified (or simulated) based on the last action selected. Then, it is compared with the state stored as a subgoal (evaluation). If subgoal is reached, two states are equal (or equivalent), then as a last step, the stored sequence of actions are actually applied to task to convert the external model to internal model. When the last subgoal is accomplished in this fashion, the goal is also satisfied and task is completed successfully.

There are two more outcomes of the evaluation step other than reaching the subgoal. First, subgoal may be a state that cannot be satisfied with any sequence of actions. If this is the situation, evaluation strategy is changed to have more relaxed requirements

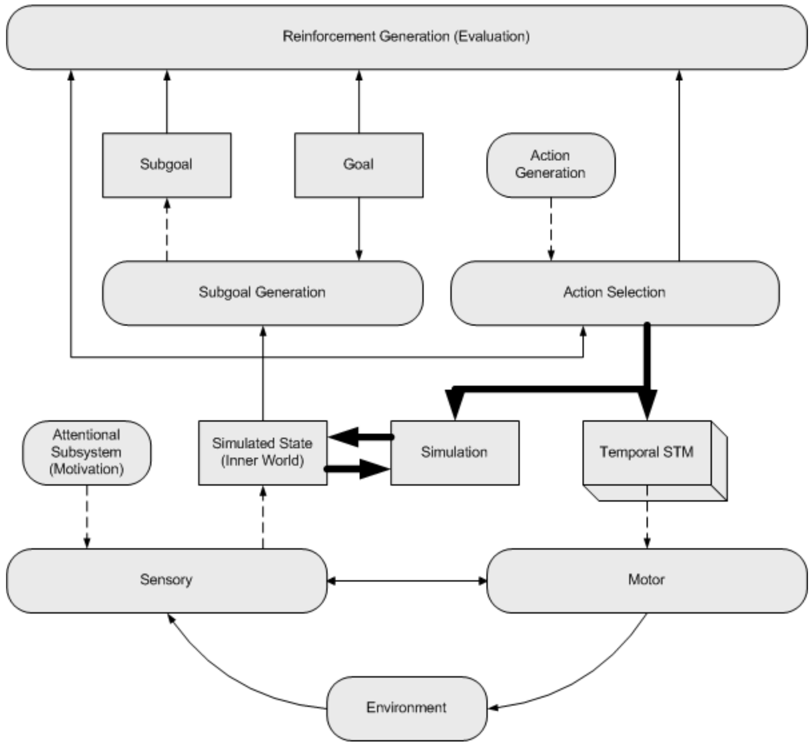

Figure 4.6: Simulation step.

such as clearing the target peg first. Second, subgoal cannot still be reached by any (modified) strategy in a requested time or by less than a specific move count. In that case, the task is left unsolved, the model is exhausted. It can be said that subjects' lost their motivation if they cannot solve the problem by spending too much time or by making lots of moves.

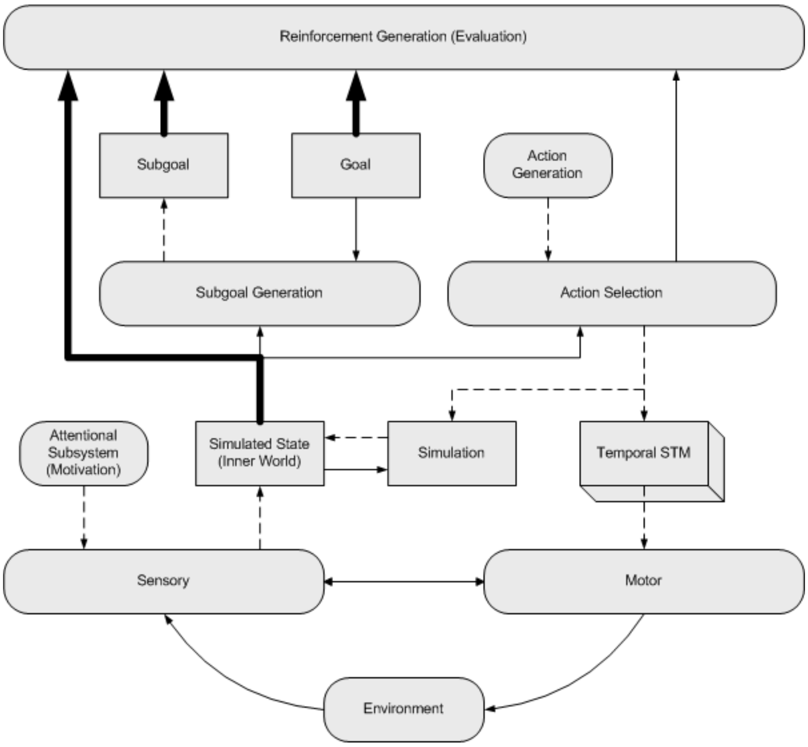

Figure 4.7: Evalution step.

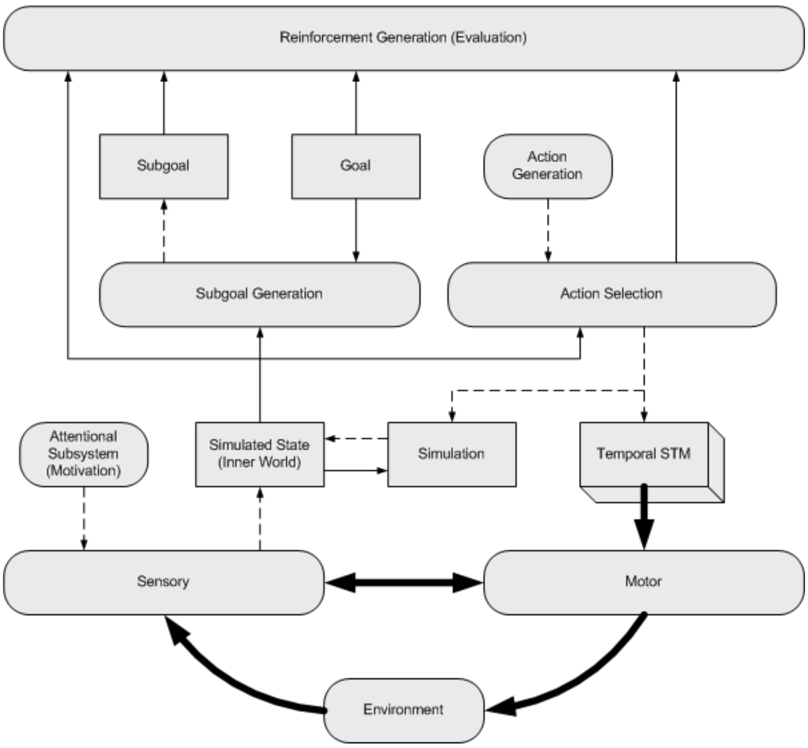

Figure 4.8: Acting step.

4.4 Simulation and Results

The simulations are carried out for nine initial configurations shown in Figure 4.9 given in [45] by changing specific properties of the model. The goal, which is same for all the initial configurations, is to stack the discs on the middle peg. These nine problems are divided into three categories according to the move count required to solve them.

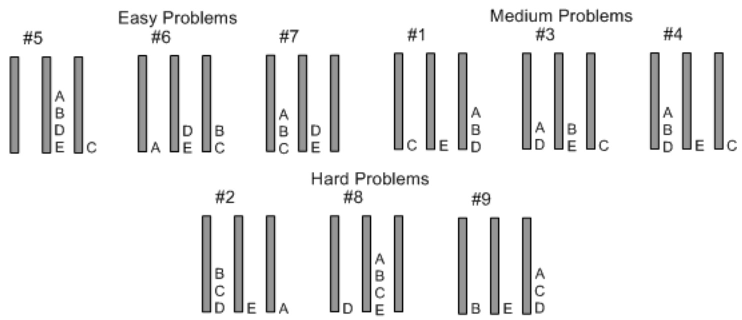

Figure 4.9: The Tower of Hanoi configurations used for simulations [45].
(A=smallest, E=biggest)

First, the capacity of working memory holding the planned actions is tested by changing the capacity between two and five. Working memory capacity to hold about five items is an experimental finding in psychology. The memory for other parts such as goal and subgoal is not considered, because goal and subgoal have other cues to help their retrieval, whereas the planned sequence of actions does not. Second, the evaluation strategy used, in case subgoal cannot be reached, is tested by employing three strategies; (1) clearing the target disc (removing all the discs on top of the disc to be moved), (2) clearing the target peg (removing all the discs on top of the disc in goal position in the target peg), and (3) both together.

The simulation results are given in Table 4.1. The results give the number of moves the model performed for each problem using each evaluation strategy (both, clearing only target disc, clearing only target peg) and each memory capacity (two to five). X represents an unsolved problem.

4.5 Conclusion

The proposed model succeeds to solve the Tower of Hanoi task by generating the temporal sequence of actions and selecting the correct one considering an evaluation process when needed. As it can be followed in Table 4.1, by decreasing the

Table 4.1: Simulation results.

Problem	Evaluation Strategy											
	Both				Only Disc				Only Peg			
	5	4	3	2	5	4	3	2	5	4	3	2
1	11	11	11	11	11	11	11	X	11	11	X	X
2	14	15	14	X	14	14	X	X	X	X	X	X
3	10	10	10	10	10	10	10	X	10	10	X	X
4	11	11	11	11	11	11	11	X	11	11	X	X
5	7	7	7	7	7	7	X	X	7	7	7	X
6	7	7	7	7	7	7	X	X	7	7	X	X
7	7	7	7	7	7	7	7	X	7	7	X	X
8	16	15	X	X	X	X	X	X	15	15	X	X
9	14	15	14	X	X	X	X	X	X	X	X	X

memory capacity, mostly an increase in the probability of leaving the task unsolved is experienced where a change in the move counts are only observed in three setups. For example, problem 4 can be solved with first two strategies with a memory capacity of two, however, it is left unsolved when the last strategy is used with the same memory capacity. Memory capacity affects how far or deep the planning process can clearly foresee in the problem space. Therefore, the decrease in the memory capacity forces the planning process to be completed in a limited span in the problem space. Such a decrease in memory capacity increases the dependence of task solution on the evaluation process; thus the evaluation process plays an important role when working memory capacity for the task is less than the required amount. Thus the prefrontal deficits can be responsible for the faults in evaluation process.

In the literature, such tasks are usually classified according to the move counts required to solve them [45]. However, according to the simulation results, it seems this may not be a correct way to judge subjects' performance, because the difficulty level of a problem depends on the subject's limitations on both working memory and evaluation strategy, thus high level cognition [65,66]. For example, following the last strategy, the problems 7 and 8 can be solved only with memory capacity of five or four. However, according to categorization based on move counts required to solve the task, problem 7 is an easy one but problem 8 is an hard one. Thus, the memory capacity alone cannot be a sufficient argument to explain the planning phenomena.

It also seems that models of planning including the one in this study cannot be a framework for a comparative study of subjects' performance since learning is missing but absolutely important. During trials, subjects obviously adapt their strategy and they use the knowledge gathered from other tasks on a problem they face. Neither generating nor adapting the strategy nor transfer of knowledge is modeled in the literature as well as in this study.

Moreover, it is also certain that the evaluation strategy must be adaptable in long-term, since we learn some invariant properties on some trials and apply it to others. Subgoal generation may be one of such an invariant property. These comments based on the simulation results are partially in agreement with [45]. The results on the effects of memory capacity is compatible with [45], but we do think that goal-subgoal conflict is a side effect which depends on the evaluation process.

5. CONCLUSION

5.1 Achievements

In this thesis, first, an introduction to cognitive science including its origin and definition, basics of nervous system, learning and computational models are given. Then, the goal directed behavior is investigated to reveal the required control of behavior which is planning. Based on the literature in psychology, artificial intelligence and neuroscience, planning is reinterpreted to set a definition and construct a framework to study behavior of subjects during planning tasks. Finally, a novel model in order to solve the Tower of Hanoi task is proposed. The proposed model is implemented in computer software and simulations are done for nine problems given in [45].

The progress over existing literature is that; (1) an explicit model to understand information flow during planning is proposed, (2) based on comparative analysis between the simulation results and the performance data of subjects, it is found that the working memory capacity alone is not a sufficient factor to explain the behavior for planning e.g. in the Tower of Hanoi task, (3) when the capacity of working memory decreases, its effect on the model's behavior is also decreased which, in turn, increases the importance of evaluation process, (4) a counter-argument is proposed for scoring the planning tasks since the cognitive ability of subjects are unknown. Based on the evidence given in this study, the effects of prefrontal lesions on the planning ability are not only restricted to the capacity of working memory but also evaluation process.

According to the interpretation of planning given in this thesis, it is speculated that the reason behind high level cognitive difference between human and animals can be the ability of human cortex to construct detailed representations of external world and simulate possible future events. Thus, the evalution and reward systems in human considers not only previous or current states but the future states he may encounter. Existence of such an architecture by evolution can be the reason why only human questions his being by problems of philosophy in order to foresee his future.

5.2 Further Work

The proposed model in this thesis lacks two important phenomena; learning and neocortical computation. First, learning in related components should be investigated to understand how subjects learn the invariants of the task by repeated trials. Next, the question of how such knowledge is applicable to different initial conditions as well as different domains should be studied.

Finally, the proposed model should be implemented on a computational framework of neocortex. I think neocortex, which is quite larger in human, can give rise to planning by uniform computational principles defined in the literature e.g. prediction and memory.

BIBLIOGRAPHY

[1] **Klein, S.B.**, 2000. Biological Psychology. Prentice Hall.

[2] **Rosenzweig, M.R., Breedlove, S.M. and Watson, N.V.**, 2005. Biological Psychology. Sinauer Associates.

[3] **Bennett, M.R. and Hacker, P.M.S.**, 2003. Philosophical Foundations of Neuroscience. Blackwell Publishing.

[4] `http://plato.stanford.edu`

[5] `http://en.wikipedia.org`

[6] **Sutton, R.S.**, 1984. Temporal credit assignment in reinforcement learning, *PhD Thesis*, Department of Computer Science, University of Massachusetts.

[7] `http://www.askoxford.com`

[8] `http://www.inf.ed.ac.uk/events/christopherlh.html`

[9] `http://www.cognitivesciencesociety.org`

[10] **O'Reilly, R.C. and Munakata, Y.**, 2000. Computational Explorations in Cognitive Neuroscience. The MIT Press.

[11] **Marr, D. and Poggio, T.**, 1976. From Understanding Computation to Understanding Neural Circuitry, Massachusetts Institute of Technology, Cambridge, MA, USA.

[12] **Churchland, P.**, 1989. Neurophilosophy: Toward a Unified Science of the Mind-Brain. The MIT Press.

[13] **Churchland, P. and Sejnowski, T.J.**, 1992. Computational Brain. The MIT Press.

[14] **Hawkins, J. and Blakeslee, S.**, 2004. On Intelligence. Times Books.

[15] **Lighthill, J.**, 1973. Artificial Intelligence: A General Survey in Artificial Intelligence: a paper symposium, Science Research Council, .

[16] `http://www1.elsevier.com/homepage/sah/ifcn/doc/rtms-inf.htm`

[17] **Dayan, P. and Abbott, L.**, 2001. Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. The MIT Press.

[18] **Bhalla, U.S.**, 2006. Signaling computations at the synapse: modeling molecular computation, *Okinawa Computational Neuroscience Course*, Okinawa, Japan, June-July 2006.

[19] **Langley, P.**, 2004. Architectures for Intelligent Systems, *The AAAI Fall Symposium on Achieving Human-Level Intelligence through Integrated Systems and Research*, , 2004.

[20] **Markram, H.**, 2006. The Blue Brain Project, *Nature Reviews Neuroscience*, **7**, 153-160.

[21] **Langley, P.**, 2006. Intelligent behavior in humans and machines, Computational Learning Laboratory, CSLI, Stanford University.

[22] **Koch, C. and Segev, I.**, 1998. Methods in Neuronal Modeling. CRC Press.

[23] **Arbib, M.A.**, 2002. The Handbook of Brain Theory and Neural Networks. The MIT Press.

[24] `http://spot.colorado.edu/~dubin/talks/brodmann/brodmann.html`

[25] **Kandel, E.R., Schwartz, J.H. and Jessell, T.M.**, 1995. Essentials of Neural Science and Behavior. Appleton & Lange.

[26] **Anderson, J. and Rosenfeld, E.**, 1989. Neurocomputing. The MIT Press.

[27] **Roberts, W.**, 1997. Principles of Animal Cognition. McGraw-Hill.

[28] **Suri, R.E.**, 2002. TD models of reward predictive responses in dopamine neurons, *Neural Networks*, **15(4-6)**, 523-533.

[29] **Watkins, C.**, 1989. Learning from Delayed Rewards, *PhD Thesis*, Psychology Department, University of Cambridge.

[30] **Hodgkin, A.L. and Huxley, A.F.**, 1952. A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve, *Journal of Physiology*, **117**, 500-544.

[31] **Schneidman, E., Freedman, B. and Segev, I.**, 1998. Ion Channel Stochasticity May Be Critical in Determining the Reliability and Precision of Spike Timing, *Neural Computation*, **10(7)**, 1679-1703.

[32] **Rumelhart, D.E., McClelland, J.L. and the PDP research group.**, 1986. Parallel distributed processing: Explorations in the microstructure of cognition. The MIT Press.

[33] **Grossberg, S.**, 1976. Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors, *Biological Cybernetics*, **23**, 121-134.

[34] **Grossberg, S.**, 1976. Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions, *Biological Cybernetics*, **23**, 187-202.

[35] **Grossberg, S.**, 1980. How does a brain build a cognitive code ?, *Psychological Review*, **87**, 1-51.

[36] **Carpenter, G.A. and Grossberg, S.**, 1987. A massively parallel architecture for a self-organizing neural pattern recognition machine, *Computer Vision*, **37**, 54-115.

[37] **Goel, V. and Grafman, J.**, 1995. Are Frontal Lobes Implicated in Planning Functions: Re-examining the Data from the Tower of Hanoi, *Neuropsychologia*, **33(5)**, 623-642.

[38] **Montague, P.R., Dayan, P., Person, C. and Sejnowski, T.J.**, 2002. Bee foraging in uncertain environments using predictive hebbian learning, *Nature*, **377**, 725-728.

[39] **Morris, R. and Ward, G.**, 2005. The Cognitive Psychology of Planning. Psychology Press.

[40] **Newman, S.D., Carpenter, P.A., Varma, S. and Just, M.A.**, 2003. Frontal and parietal participation in problem solving in the Tower of London: fMRI and computational modelling of planning and high-level perception, *Neuropsychologia*, **41**, 1668-1682.

[41] **Dehaene, S. and Changeux, J.P.**, 1997. A hierarchical neuronal network for planning behavior, *PNAS*, **94**, 13293-13298.

[42] **Damasio, A.R.**, 1994. Descartes' Error. Vintage Books.

[43] **Minsky, M.**, 1961. Steps toward artificial intelligence, *Proceedings IRE*, **49**, 8-30.

[44] **Langley, P. and Rogers, S.**, 2005. An extended theory of human problem solving, *27. Annual Meeting of the Cognitive Science Society*, , 2006.

[45] **Goel, V. Pullara, D. and Grafman, J.**, 2001. A Computational Model of Frontal Lobe Dysfunction: Working Memory and the Tower of Hanoi, *Cognitive Science*, **25(2)**, 287-313.

[46] **Owen, A., Downes, J., Sahakian, B., Poley, C. and Robbins, T.**, 1990.

Planning and spatial working memory following frontal lobe lesions in man, *Neuropsychologia*, **28**, 1021-1034.

[47] **Fincham, J.M., Carter, C.S., van Veen, V., Stenger, V.A. and Anderson, J.R.**, 2002. Neural mechanisms of planning: A computational analysis using event-related fMRI, *PNAS*, **99(5)**, 3346-3351.

[48] **Rainer, G., Rao, S.C. and Miller, E.K.**, 1999. Prospective coding for object in primate prefrontal cortex, *The Journal of Neuroscience*, **19**, 5493-5505.

[49] **Koechlin, E., Basso, G., Peitriniand, P., Panzer, S. and Grafman, J.**, 1999. The role of the anterior prefrontal cortex in human cognition, *Nature*, **399**, 148-151.

[50] **Ridderinkhof, K.R., van den Wildenberg, W.P.M., Segalowitz, S.J. and Carter, C.S.**, 2004. Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring and reward-based learning, *Brain and Cognition*, **56**, 129-140.

[51] **Owen, A.M.**, 1997. Cognitive Planning in Humans: Neuropsychological, Neuroanatomical and Neuropharmacological Perspectives, *Progress in neurobiology*, **53**, 431-450.

[52] **Paulus, M.P., Feinstein, J.S., Leland, D. and Simmons, A.N.**, 2005. Superior temporal gyrus and insula provide response and outcome-dependent information during assessment and action selection in a decision-making situation, *Neuroimage*, **25**, 607-615.

[53] **Matsumoto, K. and Tanaka, K.**, 2004. Conflict and Cognitive Control, *Science*, **303(13)**, 969-970.

[54] **Braver, T.S. and Bongiolatti, S.R.**, 2002. The Role of Frontopolar Cortex in Subgoal Processing during Working Memory, *Neuroimage*, **15**, 523-536.

[55] **Shannon, C.E.**, 1950. Programming a Computer for Playing Chess, *Phil.Mag.*, **41**, .

[56] **Newell, A., Shaw, J.C. and Simon, H.A.**, 1958. Chess-Playing Programs and the Problem of Complexity, *IBM Journal*, , 320-335.

[57] **Bernstein, A.**, 1958. A Chess Playing Program for the IBM 704, *Chess Review*, , 208-209.

[58] **Samuel, A.L.**, 1959. Some Studies in Machine Learning Using the Game of Checkers, *IBM Journal*, , 211-229.

[59] **Tesauro, G. and Sejnowski, T.J.**, 1989. A Parallel Network that Learns to Play Backgammon, *Artificial Intelligence*, **39**, 357-390.

[60] **Gazzaniga, M.S.**, 1995. The Cognitive Neurosciences. The MIT Press.

[61] **Laird, J.E. Rosenbloom, P.S. and Newell, A.**, 1986. Chunking in SOAR: The anatomy of a general learning mechanism, *Machine Learning*, **1**, 11-46.

[62] **Anderson, J.**, 1993. Rules of the Mind. Lawrence Erlbaum Associates.

[63] `http://www-history.mcs.st-andrews.ac.uk/Biographies/Lucas.html`

[64] `http://www.cs.wm.edu/~pkstoc/toh.html`

[65] **Balci, M., Kaplan, G.B., Sengor, N.S. and Gurvit, H.**, 2005. A Model of Working Memory and Planning in the Tower of Hanoi Task, *IX. International Conference on Cognitive Neuroscience*, Havana, Cuba, September 2005.

[66] **Balci, M. and Sengor, N.S.**, 2005. A Connectionist Model of Planning: A Behavioural Approach, *Dynamical Neuroscience XIII: Computational Cognitive Neuroscience*, Washington D.C., USA, November 2005.

[67] **Destexhe, A. Mainen, Z.F. and Sejnowski, T.J.**, 1994. Synthesis of models for excitable membranes, synaptic transmission, and neuromodulation using a common kinetic framework, *Journal of Computational Neuroscience*, **1**, 195-230.

[68] **Barto, A.G., Sutton, R.S. and Anderson, C.**, 1983. Neuron-like adaptive elements that can solve difficult learning control problems, *IEEE Transactions on Systems, Man, and Cybernetics*, **13**, 834-846.

[69] **Heil, J.**, 2004. Philosophy of Mind. Routledge.

[70] **Minsky, M.**, 1986. The Society of Mind. Simon and Schuster.

[71] **Searle, J.**, 1984. Akıllar, Beyinler ve Bilim. Say Yayınları.

[72] **Searle, J.**, 1992. The Rediscovery of the Mind. The MIT Press.

[73] **Sutton, R.S. and Barto, A.G.**, 1998. Reinforcement Learning. The MIT Press.

[74] **De Schutter, E.**, 2000. Computational Neuroscience. CRC Press.

BIOGRAPHY

Mete Balcı was born in 1980 in Istanbul, Turkey. He graduated from Istanbul Ataturk High School of Sciences in 1997 and attended Istanbul Technical University. He received B.Sc. degrees in both Electronics & Communication Engineering and Computer Engineering in 2003 and 2004 respectively. In 2007, he received M.Sc. degree in Electronics Engineering also from Istanbul Technical University. He had presented the parts of this research at the IX. International Conference on Cognitive Neuroscience (ICON) held in Havana, Cuba and at the Dynamical Neuroscience XIII: Computational Cognitive Neuroscience Conference (CCNC) held in Washington D.C., USA. He is employed at Pozitron Ltd. as Director of Research and Development.