Neurocomputational model explaining role of Dopamine in reward based learning

A thesis submitted in partial fulfillment of the requirements
for the award of the degree of

Master of Technology

in

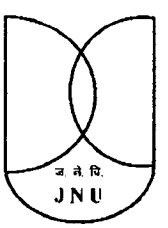
Computational and Systems Biology

Submitted by

Sandeep Singh
E. No.: 07/75/MT/16

Under the guidance of

Dr. Lovekesh Vig
Asst. Professor



School of Information Technology
Jawaharlal Nehru University, New Delhi

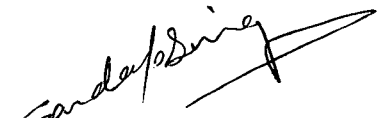
May, 2009

Acknowledgement

I would like to dedicate this work to the memory of my mother, Mrs. Malti Devi, who encouraged me for opting science stream in secondary education, later for engineering education and constantly guided me at each juncture of the life till her death.

I am indebted to my supervisor, Dr. Lovekesh Vig, for his outstanding instruction, encouragement, good advice and patience over the course of my candidature. In addition, I am grateful for his helpful comments on drafts of this project report and comment on project status presentation preparations held time-to-time during my course work. I would like to thank Dr. Ashish Gupta for giving me the initial problem, helping me to create models and review the project report. I am thankful to all the lecturers of our school for the courses they taught and their valuable comments during project status presentations. I am also thankful to entire classmates for their support throughout the course work.

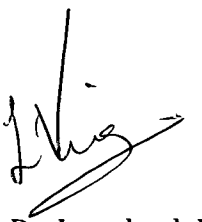
Finally, I would like to thank to my family members for constant emotional support especially to my wife Savita, whose initial decision for joining this course, constant encouragement, advice and helpful discussion made this work possible.



Sandeep Singh

Certificate

This is to certify that the thesis entitled **Neurocomputational Model Explaining Role of Dopamine in Reward Based Learning** submitted by **Sandeep Singh** to the **School of Information Technology, Jawaharlal Nehru University, New Delhi** in partial fulfillment of the requirements for the award of the degree of **Master of Technology** is a bonafide record of research work carried out by him under my supervision. The contents of this project work, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.



New Delhi
Date: 25-May-2009

Dr. Lovekesh Vig
Asst. Professor
School of Information Technology
Jawaharlal Nehru University
New Delhi

Prof. Indira Ghosh
Professor and Dean
School of Information Technology
Jawaharlal Nehru University
New Delhi

Table of Contents

Acknowledgement ii

Certificate iii

Table of Contents iv

List of Figures vii

List of Abbreviations and Symbols ix

Chapter 1. Introduction and Motivation 1

Chapter 2. Literature Review 4

2.1. Classical Conditioning 4

2.2. Classical Conditioning in Real World Scenario 4

2.3. Operant Conditioning: 5

2.3.1. Positive Reinforcement 5

2.3.2. Negative Reinforcement 5

2.3.3. Positive Punishment 5

2.3.4. Negative Punishment 6

2.4. Mesolimbic Dopaminergic System 6

2.4.1. Reward Mechanism in VTA 6

2.4.2. Dopamine Receptor Subtypes 7

2.4.3. Learning in Mesolimbic Dopaminergic Pathway 7

2.4.4. Role of Dopamine in Reinforcement Learning 8

2.4.5. Neurocomputational Models for Tonic and Phasic Activities of Dopamine 9

2.5. Neuron and its Computational Representation 10

2.6. Psychological Experiments for Acquisition and Extinction Exhibition 15

2.7. Computational Models for Exhibition of Acquisition and Extinction Processes 17

2.8. Neurocomputational Model without Catastrophic Interference 20

2.9. Neurocomputational Model to Exhibit the Dual-Pathway Neural Network 21

Chapter 3. Cognitive Modeling Using PDP++ 22

3.1. PDP++ 22

3.2. Leabra Framework: 24

3.3. Neurocomputational Model Designing in Leabra Modeling Framework 25

Chapter 4. Mathematical Model Explaining Relation Between Hunger, Dopamine And Reward Rate 32

4.1. Background 32

4.2. The Model 35

4.3. Model's Behavior in Different Scenarios 37

4.3.1. Chronic Starvation 37

4.3.2. Spontaneous Recovery 38

4.4. Neurocomputational Model Exhibiting Relation Between Reward and Hunger .41

4.5. Conclusion 42

Chapter 5. Neurocomputational Model 43

5.1. Computational Model 43

5.2. Dopaminergic Pathways in Basal Ganglia 44

5.3. Phasic Burst and Dopamine Mediated Learning 45

5.4. Deficits Induced by Medication in Parkinson's Disease 46

5.5. Neuroanatomy and Biochemistry of Basal Ganglia 46

5.6. Cellular Mechanisms of Dopamine in Basal Ganglia 48

5.6.1. D1 in Support of Go Cells 49

5.6.2. D2 Inhibits No-Go Cells 49

5.6.3. Dopamine in Basal Ganglia: Effects on Synaptic Plasticity 50

5.7. Neural Model of Basal Ganglia and Dopamine 51

5.7.1. Mechanics of the Model 52

5.7.2. Overall Network Division of Labor 52

5.7.3. Simulated Effects of Dopamine 53

5.7.3.1. Dopamine Modulates Learning 54

5.8. Neurocomputational Model for Examining the Role of Dopamine Influenced by Reward Rate 54

5.8.1. The Model 55

5.8.1.1. Satiation Layer 55

5.8.1.2. Connection Between Satiation Layer to Striatum Layer 55

5.8.1.3. Script Controlled Manipulation 56

5.9. Conclusion 56

Chapter 6. Simulations and results 58

6.1. Change in Action Frequency within a Session 58

6.1.1. Behavioral Experiment Results 58

6.1.2. Mathematical Model Results 59
6.1.3. Neurocomputational Model Results 60
6.2. Effect of Dopamine Level Manipulation 61
6.2.1. Behavioral Experiment Results 61
6.2.2. Mathematical Model Results 62
6.3. Effect of Reward Rate Manipulation 63
6.3.1. Behavioral ExperimentRresults 63
6.3.2. Mathematical Model Results 63
6.3.3. Neurocomputational Model Results 64
6.4. Relation between Reward Rate and Action Rate 66
6.4.1. Behavioral Experiment Results 66
6.4.2. Mathematical Model Results 66
6.4.3. Neurocomputational Model Results 68
6.5. Effect of Reward Ratio Over Controlled and Over-weight Animals 68

Chapter 7. Discussion 70

7.1. PDP++ 70
7.2. Models and Simulations 71
7.2.1. Mathematical Model 71
7.2.2. Neurocomputational Model 72
7.3. Scope for Future Work 73

References 74

List of Figures

Figure No.	Title	Page No.
2.1	The relation between rate of responding and rate of reinforcement on a random interval schedule	10
2.2	The Model simulation data showing the similarity with behavioral data	10
2.3	The sigmoid activation function	13
2.4	Speed of Acquisition and Extinction	17
2.5	Rescorla Experiment	18
2.6	Result from simulation	18
2.7	Result from Rescorla's experiment	19
2.8	Result from simulation	19
2.9	Result from Rescorla experiment	19
2.10	Result from simulation	19
3.1	Root Window	25
3.2	Project Window	25
3.3	Network Window, after creating the network of neurons	26
3.4	Environment window, after creating two events	27
3.5	Project window, after adding the all essential objects for cognitive Neurocomputational model	28
3.6	Control Panel for Trial Process	29
3.7	Graph Log View showing the 'sum square error' and 'average last cycle	29
3.8	Leabra Layer Specification Window showing kWTA parameter	30
3.9	Emergent Package version 4.0.17	31
4.1	Chronic hunger results in a decrease in the amount of extra-cellular dopamine.	33
4.2	Action rate is directly proportional to the dopamine level.	34

4.3	Neurocomputational model created on PDP++ showing the relation between reward and hunger and context layer	41
5.1	The corticostriato-thalamocortical loops, including the direct and indirect pathways of the Basal Ganglia	48
5.2	Neurocomputational model showing the basal ganglia and modulation of dopamine as suggested by Frank	51
5.3	Neurocomputational model showing the satiation layer and connection between satiation layer and Striatum	56
6.1	Behavioral results. In control rats and in dopamine depleted	58
6.2	Mathematical model results	59
6.3	Neurocomputational model results	60
6.4	Behavirol experiment results. In controlled and dopamine depleted rats	61
6.5	Mathematical model results	62
6.6	Behavirol experiment results. In control rats, and in pre-fed rats.	63
6.7	Mathematical model results: In conrol and pre-fed case	64
6.8	Neurocomputational simulation results	64
6.9	Behavirol experiment results showing elation between reinforcemnet ratio and response rate	66
6.10	Mathematical model results	67
6.11	Neurocomputational model results	68
6.12	Mathematical simulation results showing effect reward ratio on controlled and over-weight case of rats	69

List of Abbreviations and Symbols

ActionFreq	Action frequency
CR	Conditioned response
CS	Conditioned stimulus
DA	Dopamine
DAR	Dopamine responsivity
EDA	Extra cellular dopamine
GPe	External segment of globus pallidus
GPi	Internal segment of globus pallidus
LEABRA	Local, Error-driven and Associative, Biologically Realistic Algorithm
LP Count	Lever press count
PDP++	Parallel distributed processing (software name)
PMC	Pre-motor cortex
SNc	Substantia nigra pars compacta
SNr	Substantia nigra pars reticulate
UR	Unconditioned response
US	Unconditioned stimulus
VTA	Ventral tegmental area

Chapter 1

Introduction and Motivation

The signals received by the sensory organs from the environment bind with the nervous system to adapt to the environmental constrains based on the benefits associated with this adaptation. This adaptation while having obvious evolutionary benefits also results in what we refer to as learning. In biological terms, this behavioral adaptation is the modification of the interconnected synaptic weights. These synaptic weights and interconnectivity are reformed at recalling of the previously learned behavior. Some different environmental pressure and situation may cause the synaptic weights to learn or adopt a new behavior and this phenomenon is called unlearning or extinction.

Earlier there was a misconception that learning and unlearning are the reverse assignment of synaptic weights for the same association. Strong evidences has shown that the unlearning is not only the reversal of synaptic weights but it also involves some separate decremental processes which suppresses the weights acquired in the learning process and leaves most of the acquired weights unaffected.

Acquisition, which is the initial learning behavior, can be explained in terms of synaptic plasticity i.e. the reformation of interconnection of neurons and the assignment of synaptic weights for this association. Extinction, which is the adaptation of new behavior to a changing environment, can be explained in terms of rewiring between neurons and reassignment of new synaptic weights for the interconnection. In one model (McClelland and Rumelhart 1985; McCloskey and Cohen 1989), the extinction process involves the destruction of what was originally learned. However, now there are various evidences available to disprove this model (Bouton, 2004); thus much of the synaptic weights and neuron interconnections are available which are formed during acquisition process.

The acquisition and extinction process in our daily life can be explained by this example: when someone learns to play lawn tennis then actually the network of neurons work together to develop the reflex about the shots required to play the game and that is the acquisition phenomena. But when the same person learns to play the table tennis, the person starts forget the playing techniques of lawn tennis and this is the extinction phenomena.

Researchers use animals for the behavioral experiments and exploring the internal biochemical processes in the brain. First they create an environment so that the animal can learn a particular body movement. After that, they insert thin electrode in the brain for electrophysiological experiments to understand the brain parts which are active to perform a particular task. Researchers usually provide food or fruit juice as reward for any correct response of the animal.

The animals are food deprived animals and their body weights are 10 – 20 percent less than the normal body weight so that they become more responsive for a learning task. This food deprivation increases their tonic dopamine level in reward pathway and they become vigorous responsive for receiving the reward. For receiving the reward, animal presses the correct lever among the set of levers. If the animal presses the correct lever, it receives the reward otherwise it does not receive the reward. Number of lever presses in a specified time is counted which gives the action rate of the animal. It has been observed that a hungry animal press more correct lever during the initial stage of the experiments and later on the action rate get decreases. This might be the animal get satiated after receiving rewards as soon as it satiated, the action rate gets decreases. In our work we tried to answer these type of questions and prove it with the help of mathematical and neurocomputational models.

Behavioral results for acquisition and extinction processes can be explained by Neurocomputational models based on Leabra Modeling Framework (O'Reilly and Munakata, 2000). This framework gives the number of ways for interconnection of neurons, synaptic weight calculation, simulation of various cognitive processes, including characteristics of perception, attention, language, learning and memory.

The acquisition also takes place in case of psychoactive drug addiction. The rewarding property and repeated use of the drug increases the drug-associated stimuli (Wikler, 1973; Goldberg, 1976; Stewart, de Wit & Eikelboom, 1984; Childress, et. al., 1988; O'Brien et. al., 1992; Robinson & Berridge 1993; Di Chiara 1998). The psychoactive drugs normally attack the Mesolimbic dopamine pathway. The response reinforcement of drug use, activate the mesolimbic dopamine which is responsible for the operant conditioning for drug addiction (Montague, Dayan & Sejnowski, 1996; Schultz, Dayan & Montahue, 1997). A question rose by World Health Organization that why psychoactive drugs give a strong reinforcement that may lead to the development of drug dependence?

The whole project work is divided in four parts. First I read literatures for the basics about the brain processes, mainly the learning mechanism of the brain. Secondly I mimic some neurocomputational models made by researchers for showing acquisition and extinction phenomena. After that, the third phase is the designing of mathematical models for the dynamic behavior of dopamine and reward rate calculation in case of normal, dopamine depleted, pre-fed and over-weight animals. The fourth phase is the development of the neurocomputational models for the mechanism which are explained through mathematical models in third phase.

2.1. Classical Conditioning

In 1927, a Russian Physiologist Ivan Pavlov explained a learning mechanism called Pavlovian conditioning. It was the first type of learning studied hence it is also called Classical Conditioning. In this type of learning process, the environmental stimulus is combined with the naturally occurred stimulus to from a behavior. For example, a dog that is habituated to receiving food when its cage is opened begins to salivate even when the master opens its cage for cleaning. Due to these two simultaneously occurring events, the dog may learn that whenever the door will open, it will be for food. In this example, the master's opening the door of the cage is a natural stimulus. Pavlov called it Conditioned Stimulus (CS). Some behaviors which are unconditionally occur such as eye blink, is given a name by Pavlov as Unconditioned Stimulus (US). The behavior which is due to unconditional stimulus is called an Unconditioned Response (UR). If the CS and the US are repeatedly paired, eventually the two stimuli become associated and the organism begins to produce a behavioral response to the CS. Pavlov called this the Conditioned Response (CR).

Studies have shown that the conditional response may also occur after suppressing the conditional stimulus (Pavlov, 1927). This reappearance behavior of conditional response even after extinction session is called 'Spontaneous Recovery' (Rescorla 2004). This implies that the initial learning is stored somewhere which does not disappear during the learning of a new task.

2.2. Classical Conditioning in real world scenario

In the learning process of human beings, the classical conditioning theory doesn't fit completely. In Pavlovian conditioning, there are two stimuli involved whereas the humans learn to get some kind of rewards. The rewards may be the betterment of future comfort (Parents try to give good education to their children for better future), grow food grains for their survival; manufacture the products for their better

living etc. The Pavlovian conditioning can help to treat the anxiety and phobias. For example a teacher can arrange a positive environment for presentation by the students so that the student can stay calm and relaxed and can make new interconnection of neurons which can suppress the fears of standing in front of large audience.

2.3. Operant Conditioning:

In Operant conditioning (also instrumental conditioning), the association is between behavior and consequence, usually this is referred to as "goal directed behavior". Operant conditioning can be explained in terms of two technical terms – reinforcer and punisher. Any event or process, which strengthens the behavior i.e. more likely to occur, more frequently to occur, is termed as *reinforcer*. Any event which decreases the behavior i.e. makes it less likely to occur, decreases the frequency to occur, is termed as *punisher*. In operant conditioning the behavior produces the stimulus. The behavior produces because of some good or desirable consequences. The reinforcer and punisher can also work to decrease or increase the behavior (less likely or more likely to occur). So there are total four kind of operant conditioning:

2.3.1. Positive Reinforcement

In this form of conditioning the behavior which results in any kind of favorable consequence is strengthened. For example, a child crying for a chocolate will repeat this behavior if it leads to a reward (i.e. chocolate).

2.3.2. Negative Reinforcement

The subject increases the inclination for NOT repeating some behavior to get some good consequences. For example if the parents only buy a chocolate for the child when the child DOESNT cry, then the child will be less inclined to cry in the future.

2.3.3. Positive Punishment

The subject learns about some bad consequence that results from a particular behavior and learns not to repeat the behavior in future. Here some kind of learning

takes place to extinguish the behavior which results is unfavorable consequences. For example a thief gets punishment after getting caught red handed by the police.

2.3.4. Negative Punishment

A behavior is weakened by resulting in avoidance of a favorable consequence. For example, denying chocolate to a crying child results in weakening of the crying behavior.

2.4. Mesolimbic dopaminergic system

In recent years, studies from the field of drug addiction discovered that the nucleus accumbens (NAc) or ventral striatum and its inputs from the ventral tegmental area (VTA) of the midbrain are highly influenced by drug reward or natural rewards (Kobb et al, 2001; Wise 1998). Electrophysiological and neurochemical techniques prove that neurotransmitter *'dopamine'* is highly involve in reward pathways (Fibiger & Phillips, 1979; Wise, 1978). These brain elements influenced by dopamine form the *'mesolimbic dopaminergic system'*. The nucleus accumbens and ventral striatum are located in basal ganglia, which controls the various psychomotor behaviors.

Degeneration of dopaminergic cells in substantia nigra pars compacta and due to this, the loss of dopamine in striatum leads to the dieses like Parkinson's, schizophrenia and drug addiction (Nemeroff and Bissette 1988; Erenberg 1992, Koob and Nestler 1997, Wise 1998). The NAc receives excitatory connections from prefrontal cortex, hippocampus and amygdale (Swanson and Cowan 1975, Chronister et al 1981, Christie et al 1987, Parent 1990, Sesack and Pickel 1990, Pennartz et al 1994, Shinonaga et al 1994, O'Donnell & Grace 1995). The striatum receives the excitation impulses from other areas of cortex and the thalamus (Groves 1983, McGeorge and Faull 1989, Parent 1990).

2.4.1. Reward mechanism in VTA

The responsible reward and motivational nerve cells starts from the base of the brain VTA (ventral tegmental area), and sends pulses in the frontal brain, in more specific terms to the inner portion of the frontal cortex i.e. nucleus accumbens

(Nestler EJ, 2001). In humans, the reward pathway is more complex and constitutes other parts of the brain. Like amygdale which is a special part of the reward system, helps to make decisions for pleasurable or unpleasant feelings and make connections for these experiences (Nestler EJ, 2001). Hippcampus region is responsible for recording memories of an experience. It also records the cause and with whom it occurred (Nestler EJ, 2004). The frontal cortex coordinates and processes all these information and determines the ultimate behavior. At the last, the VTA accumbens pathway works to calibrate the weights of reward. More rewarding weight gives a strong impression and more likely to repeat the same behavior to get more reward (Nestler EJ, 2004). Psychoactive substances and many drugs give the same rewarding effect. This causes the dopamine-mimicking signals to nucleus accumbens and the reward pathway gets activated (Bozarth MA, 1994; Robinson TE and Berridge KC, 2001).

2.4.2. Dopamine Receptor Subtypes

Based upon pharmacological and biological criteria, all the five dopamine receptors are divided into two globally accepted families (Sokoloff and Schwartz, 1995; Missale et al 1998). D1 and D5 receptors lie under D1–like family. These receptors excite the enzyme adenylate cyclase by activating the $G_{s/olf}$ proteins. The D1 and D5 receptors can also activate other proteins like G_o and G_z (Sidhu, 1998).

Dopamine receptors D_2, D_3 and D_4 lie in D_2–like family. These receptors inhibit the enzyme adenylate cyclase by activating the $G_{i/o}$ proteins. These receptors also inhibit phosphatidylinositol turnover, increase in K^+ channel activity and inhibit mobilization of Ca^{2+} ions (Vallar L and Meldolesi J, 1989).

2.4.3. Learning in mesolimbic dopaminergic pathway

Organisms perform actions to survive in the environment. The actions may lead to some good experiences (gives higher reward) and also to the bad experiences (gives lower reward). Organisms try to retain those good experiences for future use. The retention of experiences which can give higher rewards from the environment is called *Reinforcement Learning.*

Reinforcement learning consists of four main subparts, namely a policy, a reward function, a value function and a model (Sutton R.S. et al 1998). Policy describes the organism's behavior during the learning phase. In other words, it is a searching process for the best state from the set of states to move from previous state to next state. It may involve simple data search from the table or complex calculation to find out the best state from the raw facts. The reward function describes the rewards achieved during the learning process. It tells about the good steps and bad steps by which the agent will move closure to or far from the final goal. The reward collected is based on the policy chosen for state selection process. If a new state gives low reward, the policy may be changed to select some other state for better reward collection. The value function describes the total reward which can expect to collect from current state to final state. It is also related with the policy making for selection of new state. The learning agent do not change the selection policy even if a selected new state gives low reward but the policy describe the expectance of high reward value in long run. The model of the environment gives the prediction of the result before taking the decision steps. For a relatively new environment, the trial and error method may be used to search the best states during the learning process.

2.4.4. Role of Dopamine in Reinforcement Learning

Dopamine is a neurotransmitter used by the reward pathway also called mesolimbic pathway. Dopamine has a great deal of influence over mesolimbic pathway (Berridge K.C. 1998). Improper regulation of dopamine may cause serious problem in complex brain processes like

- Memory
- Motivational and emotional responses
- Reward and desire
- Addiction
- Mental illness (hallucinations and schizophrenia)

Dopamine releases via two functionally independent components. One is phasic component of dopamine release in which the dopamine releases in short pulsating manner due to action potential and is rapidly removed from the synaptic cleft via reuptake (Grace AA, 1991). Another one is tonic component of dopamine which

exists at extrasynaptic receptor sites. Even the concentration level is too low but still it is sufficient to activate extrasynaptic receptors, including dopamine terminal autoreceptors (Berridge KC, 2004, Grace AA, 2002). Neurophysiological studies in behaving monkeys have shown that dopamine neurons respond to primary food and fluid rewards and to conditioned incentive stimuli predicting reward. These studies suggested that 'prediction error' signals are reported by phasic spiking activity of dopamine cells (Schultz W. et al, 1993).

2.4.5. Neurocomputational Models for Tonic and Phasic Activities of Dopamine

Computational models described the phasic activities of dopamine cells for learning processes and also showed that how this signal is used for reinforcement learning to get maximum reward (Sutton RS, Barto AG, 1990). The change in concentration level of dopamine, affect a wide variety of behavior. Some of which do not have any relation with phasic spiking activity of dopamine cell. Simple neurocomputational models to show reinforcement learning would not have capability to show free-operant behavior. Tonic level of dopamine also plays a significant role in energizing behavior. For studying the dopaminergic manipulation might force such influence over response vigor, a model should constitute the tonic level of dopamine also.

A neurocomputational model presented by Daw ND et al (2007) shows the free-operant behavior. The model calculates the reward based on response time and response rate in accordance to different reinforcement learning scenarios. Average rate of reward ($\bar{R}$) gives a significant purpose and act as an opportunity cost. If the average reward is high, every second is costly and performing the action quickly at the moments of high reward may lead to high reward collection even if the energetic cost of doing so is high.

The computational model has the simulated rat which is placed in operant chamber containing one lever and a food magazine. Now the subject can take several possible actions such as press lever, nose poking, grooming, sniffing etc. Here, a pair of action step the latency for doing that action (a, τ) actually calculates the reward. The goal of the subject is to select the best pair of action and latency to maximize the reward

in long run and also minimize the incurred cost per unit time. The result as shown in fig 2.2 is somewhat similar to the result as in fig 2.1 which was seen experimentally (Niv et al, 2005).

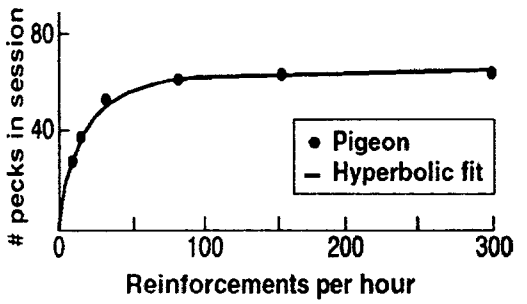


Figure 2.1: The relation between rate of responding and rate of reinforcement on a random interval schedule (Herrnstein, 1970)

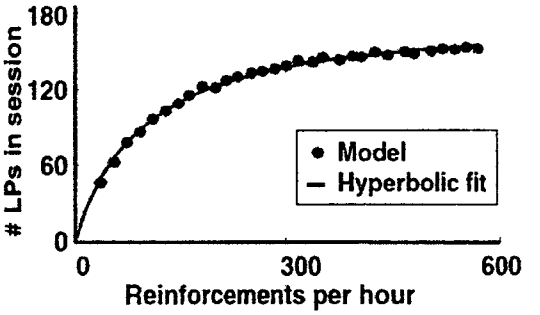


Figure 2.2: The Model simulation data showing the similarity with behavioral data.

Computational model also shows that selected latency for each step is inversely proportional to the average reward rate stating the higher response rate in case of high average reward rate. Further, it has been noted that in case of two levers, the response rate is similar to the reward rate of each lever. The numerical value of average reward rate ($\bar{R}$) plays a critical role in action-latency pair selection and it works as an opportunity cost. There is not any mechanism in this model by which one can correlate the tonic level of dopamine for the expected average reward rate. However, computationally we should expect the tonic average reward signal to be used predicatively and not only reactively, which would require it to be somewhat decoupled from the actual obtained phasic reward signal.

2.5. Neuron and its Computational Representation

Neuron is a biological cell just like any other cell in the body of species. The basic function of a neuron is to develop the behavior for short term as well as long term consequences so that the organism can survive in the environment. It has some special properties like dendrites, membrane potential, action potential etc. Dendrites work as input channel for chemicals called neurotransmitters to travel from sensory organs to neuron cell. The axon works as an output channel for the neuron so that the processed signal can pass-on to the next level of neurons for

further processing or to the motor mechanism for any kind of mechanical movement for the body.

The signals are in the form of ions called neurotransmitters, which travel throughout these dendrites, axons and also in the cell body. The fundamental principles behind the movement of neurotransmitters are electricity and diffusion. Because branches of dendrites are connected with various axons of other neurons hence one neuron cell can continuously receive and transmit the ions. Due to density of ions in the cell, the cell potential keeps changing. Every neuron has their specific property to accept the specific kind of neurotransmitter and also after some specific potential it only sends the neurotransmitter to the axon. This particular potential, only when a neuron sends the neurotransmitter is called action potential. The junction of dendrite and axon is called synapse. Dendrite releases the neurotransmitters over spikes in between narrow gap of dendrite and axon. The axon receives the neurotransmitters in the cell body through microtubules. The rest of the neurotransmitters which are not received by axon are taken back for reuse by the dendrite (*reuptake*).

In Computational models, the strength of synapse i.e. the capability of transmitting neurotransmitters between two neurons is considered as weight between the two neurons. There are two methods by which the ions move from one neuron to other. The first one is by the principle of electricity and second one is by the principal of diffusion. The movement on the principle of electricity, ions follows the well known Ohm's law:

$$I = \frac{V}{R} \tag{2.1}$$

Where, I is the amount of ions motion through the membrane of the neuron cell, V is the membrane potential and R is the resistance (blockage or barrier in the path of ion flow) made by the membrane and cell fluids. A conductance (G) can be defined as the inverse of the resistance ($G = \frac{1}{R}$) and represents how easily an ion can pass through the membrane. Now Ohm's formula in terms of conductance can be written as:

$$I = VG \tag{2.2}$$

By changing this resistance to conductance, we can easily understand that how a neuron defines the input quantity of ions in the membrane. It allows or disallows the ions by opening and closing the membrane channels by determining the conductance for each type of ion as a function of the input it receives.

One more cause for the ion movement is diffusion. When the concentration of one type of ions at one place increases then they start moving to low dense area to distribute the ion concentration evenly at all place. Unlike electricity, the same charged ions cannot replace each other. That's why the diffusion makes the concentration of every neurotransmitter even through cell membrane and channels. An electrical potential is also applied on the ion, so during the ion diffusion, the electrical potential counter acts on the diffusion force and that stops the diffusion. This particular potential at which the equilibrium point obtained is called equilibrium potential or the driving potential because now the flow of ion will drive the membrane potential towards this value. If the equilibrium potential (E) is known, the net membrane potential can be obtained by subtracting this value from the actual potential V:

$$I = G(V - E) \tag{2.3}$$

Equation 2.3 is called diffusion corrected version of Ohm's law. This equation can be applied on ion-by-ion basis to calculate the current contribution by each type of ion.

The four major ions which majorly play the role in of neurotransmitters are: sodium (Na^+), chloride (Cl^-), potassium (K^+) and calcium (Ca^{++}). Two mechanisms regulate the concentration of ions: sodium-potassium pump and selective permeability. Sodium-potassium pump pushes the Na^+ ions outside the membrane and collects some K^+ ions from outside the membrane. Selective permeability refers to channels allowing specific types of ions to pass from the channel. The opening and closing of the channels is governed by the inputs received by the neuron.

To calculate the total current for the neuron, the mathematical terms for every type of channels of a neuron must be derived. After that, the membrane potential may be

calculated for a very small duration of time because the membrane potential changes continuously (O'Reilly, 2001).

$$V_m(t+1) = V_m(t) + dt_{vm}[g_e(t)\bar{g}_e(E_e - V_m(t)) + g_i(t)\bar{g}_i(E_i - V_m(t)) + g_l(t)\bar{g}_l(E_l - V_m(t))] \quad (2.4)$$

Equation (2.4) gives the different values of conductance for different types of channels. The three different kind of ion channels are: excitatory synaptic input channel activated by the glutamate and passing the Na^+ ions (subscript e), second one is the inhibitory synaptic input channel activated by GABA passing the Cl^- ion (subscript i) and the third one is the leak channel which is always open and passing K^+ ions (subscript l). V_m represents the membrane potential.

The Sigmoidal Activation Function is the popular activation function for designing the neural network models (McCullagh, P. and Nelder, J.A., 1989).

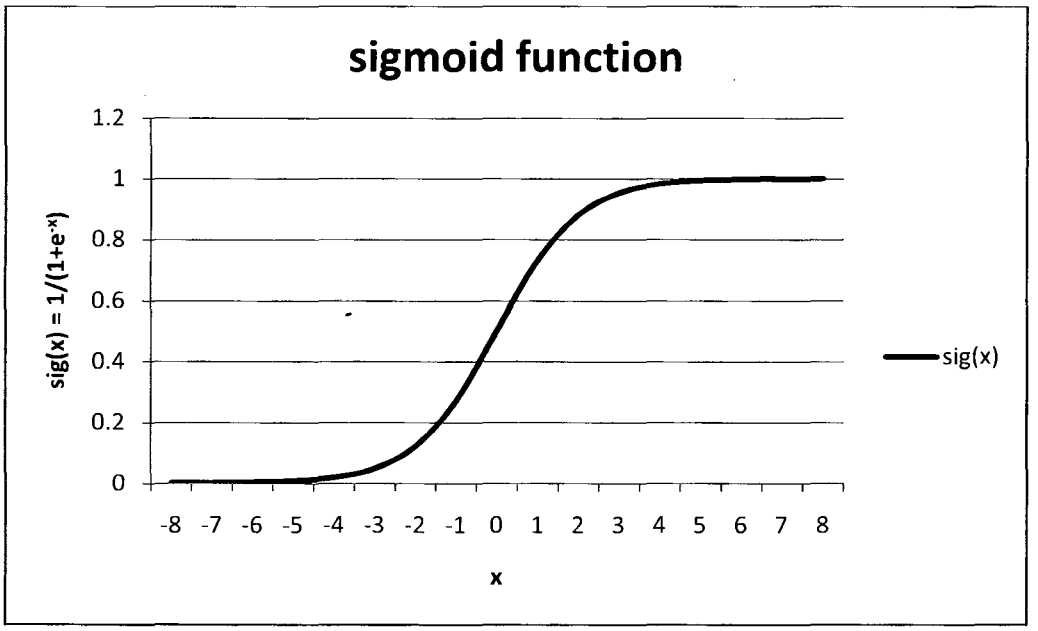


Figure 2.3. The Sigmoid Activation Function

The equation to define the sigmoidal activation function includes a net input to the unit function which is the sum of the individual weighted inputs from other units:

$$\eta_j = \sum_i x_i w_{ij} \quad (2.5)$$

Where η_j is the net input received by unit j, x_i is the activation value for sending unit i and w_{ij} is the weight of interconnection. The sigmoidal function transforms this net input value into an activation value (y_j), which is then sent to other units:

$$y_j = 1/1\text{-}e^{-\eta_j} \quad (2.6)$$

Cognition takes place in cerebral or neocortex. There are various other areas of the brain which also important for proper brain functioning. There are six primary layers of neurons in the brain (Sejnowski and Churchland, 1989). For neurocomputational modeling purpose, one can consider these six layers in to three layer architecture namely the first one is the input layer, second layer is the hidden layer which actually contains the information (weights) and the third layer is the output layer which is directly connected by the motor mechanism. Neurons work in groups. The information (signals from sensory organs) received by the input layer is transferred to hidden layer where it gets processed and the final output is given to the output layer for physical body movement or stored as cognitive phenomena.

In a group of neurons, the neurons go through the inhibition competition and the winner leads the group i.e. provide the output to the next layer. It might be possible that in one group one neuron leads the group for a task and for some another task some other neuron from the same group leads the group. There is bidirectional connectivity between groups of neurons so that the receiving units can send back the error signal to the sender. The sender then corrects the value and sends it back to the receiver side.

In an artificial neural network, there are two predominant weight calculation mechanisms, Hebbian learning and Error-driven task learning. Derived from the theory given by Canadian neuropsychologist Donald O. Hebb, Hebbian learning involves neurons adjusting their weight based on the neighboring neurons. When one cell is close enough to excite the firing of a neighbor cell and that firing can cause it to update its own weight then after some cycles the efficiency of one or both cells get increased significantly. It is a time taking process and result may lead to false learning. In case of error-driven task learning, the neurons update their

weights by receiving the error signals from remote neurons. The result in this case leads to true learning.

If the two learning mechanisms are combined, then one can achieve the actual biological learning processes of the brain (O'Reilly and Munakata, 2000).

$$\Delta w_{ij} = k_{hebb}(\Delta_{hebb}) + (1 - k_{hebb})(\Delta_{err}) \quad (2.7)$$

W_{ij} is the updated weight of a cell. Where k_{hebb} is a parameter defines the contribution of Hebbian learning in the total weight calculation. Δ_{hbb} and Δ_{err} are the updates in weights calculated using Hebbian principle and Error-driven learning principle respectively.

The artificial neuron network models designed based on the above equation gives some good learning results in comparison to the models using any one of the learning rules. However the percentage share for one weight calculation is critical for learning for long term reward learning and short term reward learning.

2.6. Psychological Experiments for Acquisition and Extinction Exhibition

Rescorla (2001) has shown some interesting results based on Pavlovian conditioning for acquisition and extinction behavior. Experiments were conducted to show whether retraining removes the decremental process developed in extinction or not. By repeating the retraining and extinction they noticed a spontaneous increase in the subjects' ability to evoke a response. The experiments indicate the superposition of a decremental process on the original conditioning that prevents the Stimulus-Outcome response while leaving the original S-O associations intact. Each experiment was based on Pavlovian conditioning using rats as subjects.

In his first experiment, Rescorla showed that insertion of retraining between extinction and the recovery test doesn't affect the recovery from the effects of extinction after passage of time. The second experiment is the replica of first one with a difference that this experiment is more close to spontaneous recovery. In first experiment, the testing was done immediately after training and extinction. Here in

experiment two the two sets of stimulus were tested and one was tested immediately and other one was tested after a delay. The result was very close to the result of first experiment and strengthens the belief that spontaneous recovery is not removed by the insertion of retraining. In third experiment Rescorla showed the amount of recovery with and without retraining. The results showed the equivalent level of responding for the two stimuli. It also suggested that any changes that took place over time following extinction were similar and irrelevant to trained stimulus. That also sowed the possibility that retraining has no impact on the magnitude of spontaneous recovery.

The goal of experiment four was to show the possibility of retraining conducted immediately after extinction would show a greater increase in the stimulus-outcome association that would retraining conducted after spontaneous recovery had occurred. The tests also suggest that retraining at a time when extinction was still recent led to a greater increase in the stimulus-outcome association that did retraining at a time when extinction was more distant. The aim of experiment five was to strengthen the results of experiment four by transferring the control to a response that has earned that same outcome. It is well documented that a stimulus will successfully transfer to an operant conditioning to the degree that the stimulus and the response share as association with the same outcome. The results for extinguished and retrained and stimulus suggest that extinction enhanced this difference and there was some elevation of responding during the same-outcome stimulus relative to responding in the baseline.

Extinction is also a one kind of learning and it is more context-specific than the original conditioning (Bouton M.E., 2004). According to Bouton (2004), there are four critical causes for extinction: the discrimination of a new reinforcement rate, generalization decrement, response inhibition, and violation of a reinforcer expectation.

It has been observed (Wagner and Brandon, 1989, 2001) that the lower rate of a new reinforcement affects the extinction process positively. This has also been tested by Haselgrove and Pearce (2003) and stated that when a conditional stimulus

is applied in extinction for long duration, the animal stopped responding in less number of trials. One behavior as given by Capaldi (1967, 1994) is that when the animal stopped responding when it generalizes the stimulus responsible for acquisition and the stimulus responsible for extinction.

2.7. Computational Models for Exhibition of Acquisition and Extinction Processes

A review by Noelle D.C. and Gupta A. in 2006 demonstrated neurocomputational models which clearly show the savings in memory during acquisition and extinction behavior. They used PDP++ (A software developed by O'Reilly and Munakata in 2000), which is well known software to design cognitive neural networks. They devised three models to show the savings in acquisition and extinction process.

In their first model, they learned the network for some defined output. After that to release the learned memory, they learned the network for another defined output. The process was repeated five times and it was found that, the number of cycles required to relearn the first-one process is less than the required cycles in initial epoch and also no. of cycles required decreases as they repeat the process. Same was true for the second-one task.

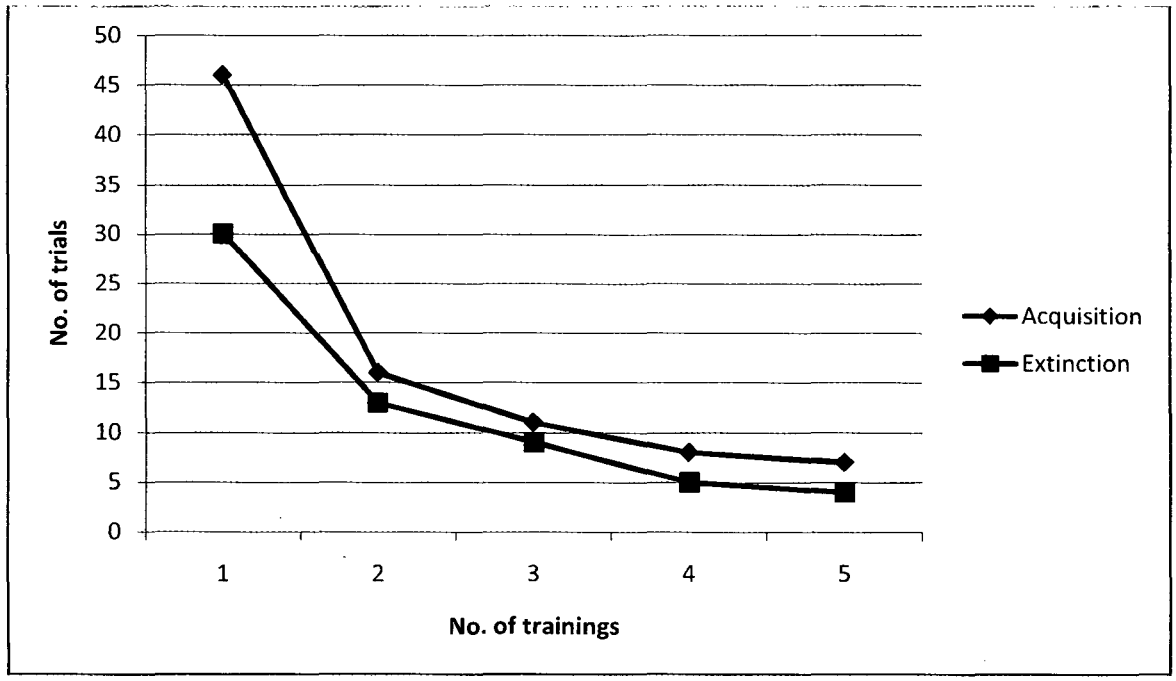


Figure 2.4. Speed of Acquisition and Extinction

They explained the result that there must be some savings of memory cycles in relearning process and the same network is trained for another output, some other neurons work together to learn the second process. These neurons over shadow the neurons which are engaged in previously learning process and lower down the weights so that they cannot trigger the previous output. But in reacquisition process (second time learning for the same process), the neurons has some previous weights and hence they learn the task in less number of cycles to attain the threshold level. This is also true for the extinction neurons. This model proves that there must be two separate sets of neurons for acquisition and extinction processes.

In their second model they showed that there must be some helping mechanism in between acquisition and extinction neurons. To explain this mechanism they designed a model based on Leabra modeling framework as earlier but used four different stimuli. In acquisition process, they trained the network with first and third stimuli whereas the second and fourth was applied in non-reinforced manner. In extinction process, the first and third were extinguishes and also the second and fourth was applied in non-reinforced manner. In second epoch, the network is trained for first and second stimuli. They repeated this re-acquisition and extinction process for 20 trials. During trainings, one can see that the network is retrained and extinguished for the first stimulus where as the network does not extinguish for the second stimulus.

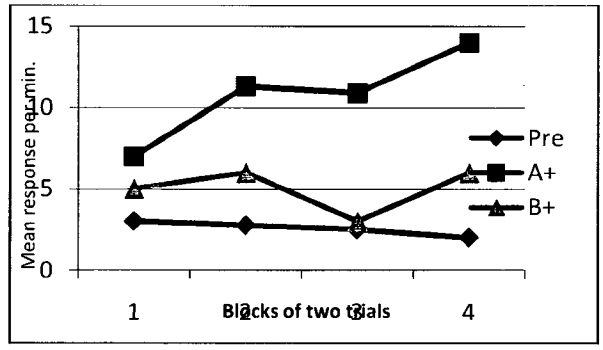


Figure 2.5. Rescorla Experiment

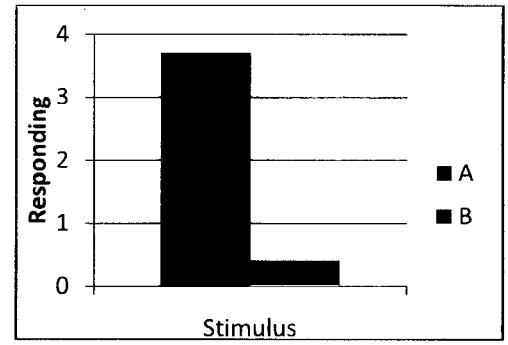


Figure 2.6. Result from simulation

The result from Rescorla's experiments based on classical conditioning and the results from this Neurocomputational model show some similarity. When the

network is tested for first stimulus and compared with the test result for response of second stimulus, the network gives the strong response for first stimulus than the second because the network gets acquaint and extinguish in each epoch whereas for the second stimulus the network only extinguish in each epoch.

In simulation they tested for the combined response of networks for the stimuli first and fourth with second and third. They found the same result as described earlier (Rescorla, 2002). This is because of the neurons involved for responding second stimulus get acquaint ed and third get extinguished in each epoch and have the

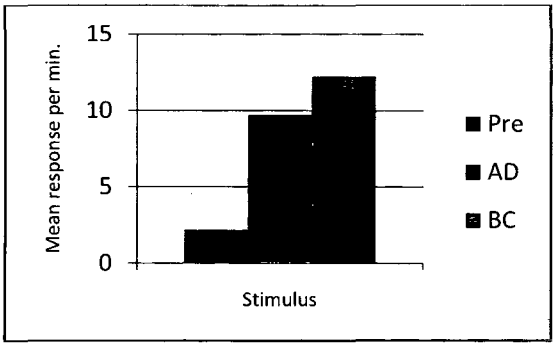


Fig 2.7. Result from Rescorla's experiment

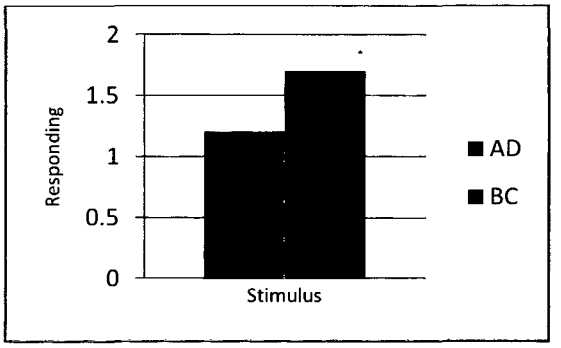


Fig 2.8. Result from simulation

strong response if they combined together than the first and fourth where for first stimulus the network has the strong response but fourth one does not help during the process of acquisition and extinction hence the combined response for first and forth has the less effect than the second and fourth.

In third model, they explained the behavior of the network when two stimuli are considered together in reacquisition process.

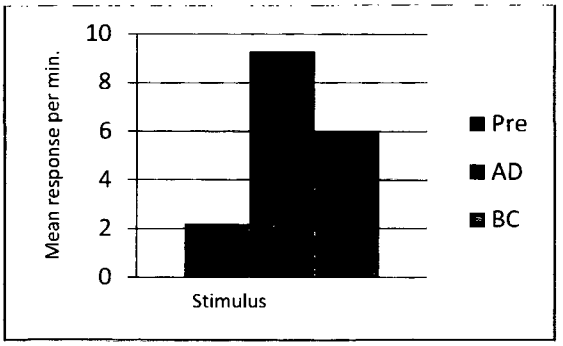


Figure 2.9. Result from Rescorla experiment

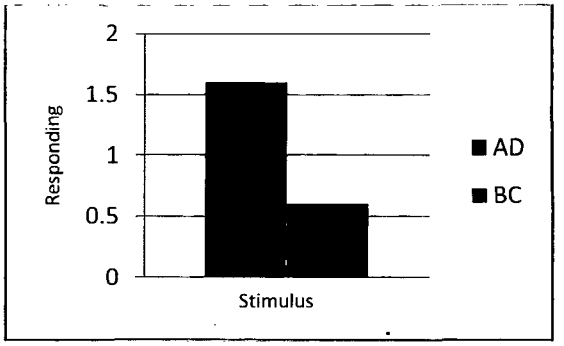


Figure 2.10. Result from simulation

They trained the network by four different stimuli. In first step of acquisition, they train the network with first and third stimuli followed by extinguishing the network with compound stimuli of first and third stimuli. In the reacquisition process, they reacquaint the network with the compound stimuli first and second. Here one can see that the network is trained by first stimulus in each epoch but the neurons which are involved to respond for the second stimulus never get extinguished and that was the reason that when they tested the combined effect of stimuli first and fourth with second and third, the response was higher for the first and fourth combined stimulus than the second and third. So there must be some over-shadowing mechanism involved which act and suppress the learned weights in the neurons which are not involved in the current learning process.

2.8. Neurocomputational Model without Catastrophic Interference

One of the challenging problems of artificial neural network is to implement the learning process as in human brain into the artificial brain (French R.M., 2003). Machines normally forget one learned task completely if machine try to learn another task. This phenomenon is called 'catastrophic interference'. This type of memory loss is significantly different than the human forgetting (Barnes and Underwood, 1959). Human brain learns the task, after dividing the task into some sub-task and learns sequentially. The brain is a distributed neural network and the artificial neuron connectionist network share single multiplicative weights and (e.g. feedforward and backpropagation networks) and do not mimic the human brain cognition process.

Gupta and Noelle (2005b) modeled a neurocomputational model using Leabra modeling framework to exhibit the machine learning free from catastrophic interference. They use the parameter kWTA (k – winner takes all) of the Leabra framework efficiently in their model to show sequential learning. After applying this parameter in the model, not all the units of hidden layer work together to learn for the desirable result but only some units (as per the value of kWTA) gives the sparse representation with few units active at a time.

They gave the assignment to learn the motion trajectory of a three joint planner arm. The learning required the sequence of task in which some of them were common in function and required same type of task learning. They found that in the process of task learning, the network showed savings of learning cycles during task learning process. Lower value of kWTA gave the better value since then the number of overlapped units for two tasks was less. One subtask used the units from previously learned units and showed the savings of learning cycles in case of similar task learning. This model also represents the sequential learning as it happens in the biological brain cognitive process.

2.9. Neurocomputational Model to Exhibit the Dual-Pathway Neural Network

Skill learning in human brain is a two step process, in the initial phase the task is in declarative sense. After that the task slowly shifted to more procedural steps (Anderson, 1981). Both phases acquire the motor skill separately. The first phase is administrated by controlled pathway and the second one is by automatic pathway. There is some association between these two pathways.

Gupta A. and Noelle D. (2007) designed a neurocomputational model to exhibit these two pathways. They trained a network for the key pressing motor arm which presses the numbers on a number-key pad sequentially. They designed a simple task learning network and added one more network in parallel to the previous one. The purpose of this second network is to transfer the copy of weights of automatic layer in previous step and provide these weights as an input to the next time step to the automatic layer. By applying the previous weights again and again gave the strong weights for the automatic layer. After some epoch, the automatic layer learned the task and gave better results in testing cycles.

Chapter 3

Cognitive modeling using PDP++

3.1. PDP++

Large number of neurons, their interconnection and lots of calculation are some of the obstacles for simulating the brain processes. Computer can help to design these kind of models but the algorithms which are used in designing the models only give the close results in comparison to biological processes. While the utility of computational models for implementing theories is now broadly recognized, the skills necessary to construct, analyze, and evaluate computational cognitive models have still created a hurdle for users and new researchers. To design the models, designer is expected to be proficient in computer programming, psychology and neuroscience so that they can design and analyze the output and can give some fruitful information (Noelle D., 2008).

Users, mainly from psychology and neuroscience are required to design the neurocomputational models for their hypothesis but it is rare to find users who are uniformly strong in all of the foundational skills that are important for the modeling work. This difficulty increases many fold when learning connectionist, parallel distributed processing, artificial neural network and computational cognitive neuroscience modeling all are required when a designer wants to construct a cognitive neural network model. These approaches to the modeling of cognition involve mathematical formalisms unfamiliar to many computer scientists and they depart radically from familiar characterizations of cognitive process in terms of stages or functional modules. Many users face significant challenges when attempting to learn these modeling framework, and even those who obtain a reasonable conceptual understanding often struggle to master the skills necessary to construct models of their own.

Designing cognitive model also needs lots of practice, patience and lots of hands-on exploration. Since computational model also requires the sound knowledge of

computer experience and those who has not strong computer skill, may face the problem during designing the models. Designing connectionist modeling insist that users face these obstacles and can reduce the burden by writing the functions in some standard high-level languages like C++, Java or MATLAB. Users also can reduce the computational burden through the use of cognitive modeling software packages. These software packages typically allow for the construction of cognitive models through helpful wizards, simple addition of objects or with the help of pre defined templates, already given in the software package. Now, user can select model type, specify a variety of model parameters, choose core connectionist algorithms and desired output form. Similar platform use gives the benefit of sharing the model among users, researchers and critics. Through the use of such simulation software, users without computer programming skills are often able to gain experience in building models, executing simulations, and collecting and analyzing data concerning model performance.

PDP++ is one such cognitive modeling simulation software package. Like many other simulators, PDP++ was primarily developed to support broad and ongoing research activities involving the construction, analysis, and evaluation of computational models of cognition. Thus, users learning cognitive modeling using PDP++ gain experience using a research-grade software tool can experience the actual biological cognitive phenomena in computer labs (O'Reilly and Munakata in 2000). PDP++ provides direct support for the most common connectionist architectures and learning algorithms, including:

- spreading activation, constraint satisfaction networks, including Hopfield networks (Hopfield 1982)
- Hebbian learning (Hebb 1949, Grossberg 1998), competitive learning (Rumelhart and Zipser 1986, Grossberg 1987), and self-organizing feature maps (Von der Malsburg 1973, Grossberg 1976a, Grossberg 1976b, Kohonen 2001)
- feed-forward backpropagation of error networks (Rumelhart et al., 1986)

- recurrent backpropagation of error networks, including simple recurrent networks (Elman 1990) and the "long short-term memory" architecture (Hochreiter and Schmidhuber 1997)

The package also provides support for computational cognitive neuroscience models intended to make more substantial contact with biological measures, offering an implementation of the LEABRA framework (O'Reilly and Munakata, 2000) as well as the Real-time Neural Simulator program. Models may be fabricated and manipulated in PDP++ using its elaborate graphical user interface, avoiding any need for users to write program code. The simulator is an open source software project, however, and utilities are provided to augment PDP++ with user-written C++ code if substantial departures from the provided algorithms are required. A detailed reference manual is available both as a printable document and as a collection of linked web pages. The entire PDP++ package, including example models, may be downloaded free of cost from the FTP site: ftp://grey.colorado.edu/pub/oreilly/pdp++/.

3.2. Leabra Framework:

LEABRA stands for "Local, Error-driven and Associative, Biologically Realistic Algorithm". It is a collection of computational formalisms for developing cognitive models that make contact with both observable behavior and detailed biological mechanisms. LEABRA models are constrained by our knowledge of processes at the level of membrane channels and individual neural functioning and also by our knowledge of gross brain anatomy and the role of various neurotransmitter systems. LEABRA is of particular interest because it incorporates many of the mechanisms that have appeared in the history of connectionist research. Its recurrent activation dynamics allow it to exhibit pattern completion and soft constraint satisfaction performance akin to that seen in Hopfield networks, other attractor networks, and spreading activation models.

Synaptic weight learning in LEABRA includes a Hebbian learning algorithm, allowing for self-organization learning, and an error-correction learning algorithm formally related to the backpropagation of error technique. Hebbian learning is performed using conditional principal components analysis (CPCA) algorithm (Olshausen &

Field, 1996) with correction factor for sparse expected activity levels. Error driven learning is performed using GeneRec, which is a generalization of the Recirculation algorithm, and approximates Almeida-Pineda recurrent backprop. The symmetric, midpoint version of GeneRec (O' Reilly, 1996a) is used, which is equivalent to the contrastive Hebbian learning algorithm (CHL). LEABRA networks can also make use of a reinforcement learning algorithm based on the role of the dopamine neurotransmitter system in learning. By bringing all of these mechanisms together, LEABRA provides a single focal framework through which a wide variety of connectionist concepts. LEABRA is fully supported in PDP++.

3.3. Neurocomputational Model Designing In Leabra Modeling Framework

There are two methods to design the models on this framework, first one is by manually creating the objects required in the simulated model and second one is with the help of in-built wizard. The second way is simple but gives the less flexibility to manipulate the model. The first one, which is the designing the model by adding the objects of the model in the workspace is somewhat tough but it provides the user to manipulate the model and create the desired form of output to verify and analyze the model.

The first step to create a neurocomputational model is to choose a new project option from the root window that will show the following window:

Figure 3.1. Root window

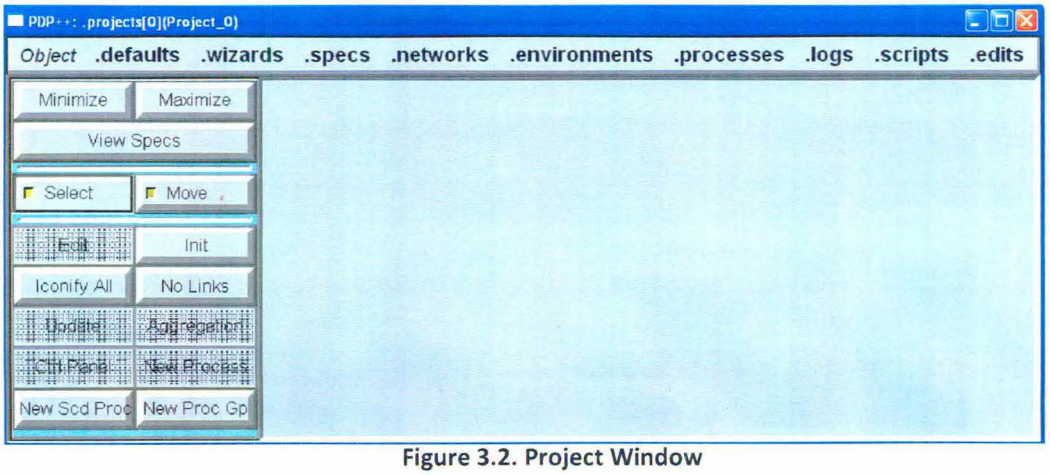


Figure 3.2. Project Window

We can consider this project window as our workspace because all the components of the model will be assembled here.

Now, the next step is to design network i.e. network of neurons. For that we have to choose the option new under network menu from the project window. The selection will open a prompt box asking the number of networks which we are interested in to create. After selecting the one network, one can see the window which is the design window for the network. Now user can choose the new layer button from the left side buttons in network window. Again a prompt box come-up asking the number of layer which has to be drawn. Here we will select a three tire network layer architecture which is most common type of neurocomputational model for cognitive processes.

The next step is to add the units in each layer and set up the connections in-between the layers. Units can be added by dragging the layer boundary and connection can be set by the selecting the two layers which has to be connect and defining the connection type (i.e. unidirectional or bidirectional) with the help of 'new projection' and 'connect all' buttons given in the left side on the network window.

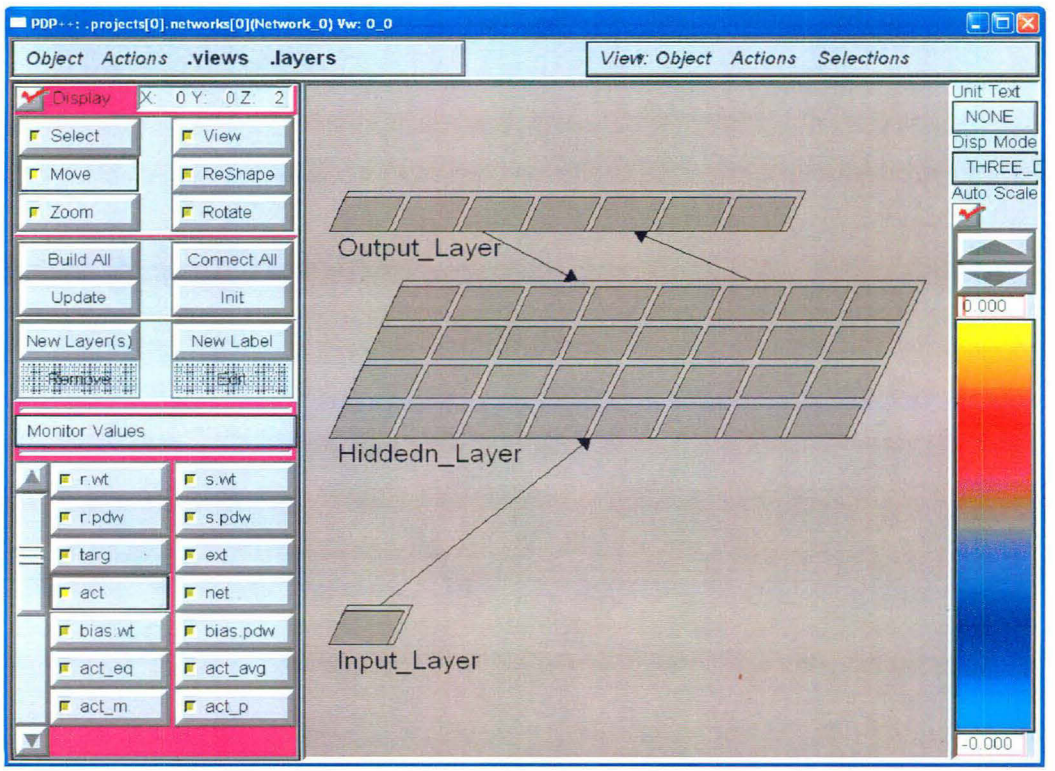


Figure 3.3. Network Window, after creating the network of neurons

A typical example with three layers, one input layer having one unit, one hidden unit containing 40 units and one output layer containing 7 units and unidirectional link between input and hidden layer and bidirectional connection between hidden and output layer can be setup and looks like the network window figure 3.3.

Next step is to define the task for which our network will learn and memories the output as defined in this step. For that purpose, we have to create the environment. To create the environment, user can select the 'new' option under 'environment' menu of the project window. An environment window will come up in which we have to create the events. Each event will represent the learning task, for example in case of AND gate simulation, one of the four events will represent the one input state and corresponding output state of the AND logic gate. Here, in our example, if we define the high input will activate the starting four units from the seven units and low input will activate the last here units from the output layer, then the environment window will look like the figure 3.4.

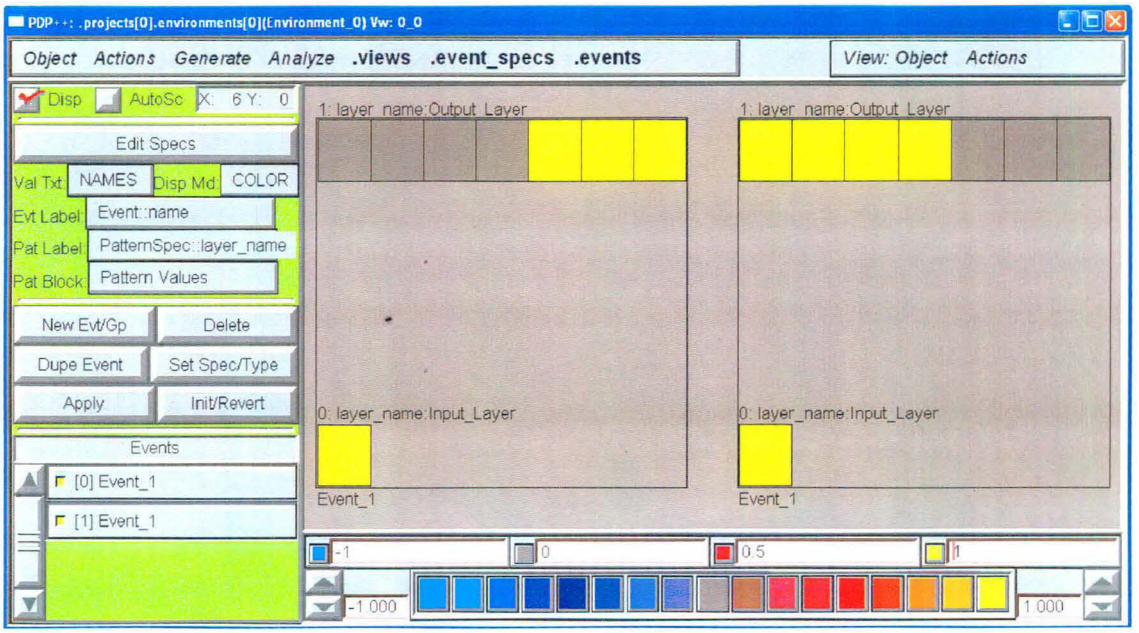


Figure 3.4. Environment window, after creating two events

The next step is to design the process (i.e. process of learning and testing) and adding reports (log) so that we can analyze the learning process of the network.

To create a process we have to add a new process from project window. The option 'new' is given under 'process' menu in the project window. Here we also have to specify the type of the process for example, if we are designing the model to show the cognitive behavior, then we have to choose the 'train process' because we want to train the network based on the desired output. There are five levels in a train process namely Train, Epoch, Trial, Settle and Cycle. Since the network learns the events in sequence, units adjust their weights based on the error signals received from the output layer. It requires some iterative cycles at each event to learn. Epoch level defines the repetition of all the events for training purpose. Settle and Cycle levels are required to adjust the unit weights according to the events.

To analyze the parameter and the error values during the learning process, we have to add log object to monitor the parameters, which can be added by choosing 'new' under the 'logs' menu in the network window. There are various kind of log reports available in this option in which the two namely 'Graph Log View' and 'Grid Log View' are highly useful. Graph log view can show the activation level of at each epoch and grid log view can show the values of each units displayed in grid like structure. After adding these objects the 'project window' will looks like the figure 3.5.

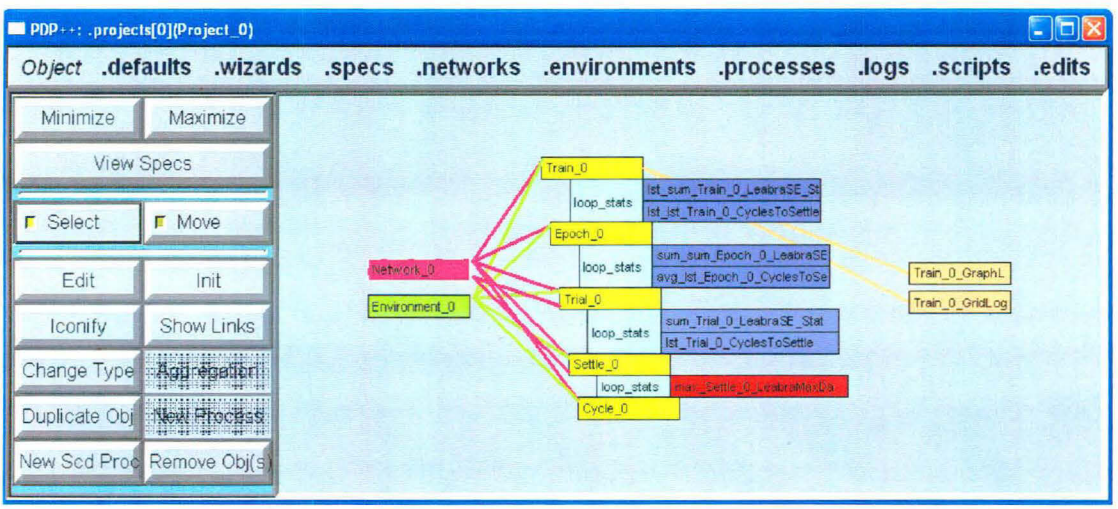


Figure 3.5. Project window, after adding all the essential objects for cognitive neurocomputational model

To train the network we have to open the control panel of the processes. For that we can select the 'control panel' option under 'process' menu on the project window. We can select control panel for any of the five levels. The control panel will show the buttons to initialize the network, run the learning process and we can stop the executing the run process in middle by pressing the stop button. We can also run the process in step manner. A train process stops when the trial process gives the no error (or error in the specified negligible region) for successive five epochs. The figure 3.6 of control panel is given below:

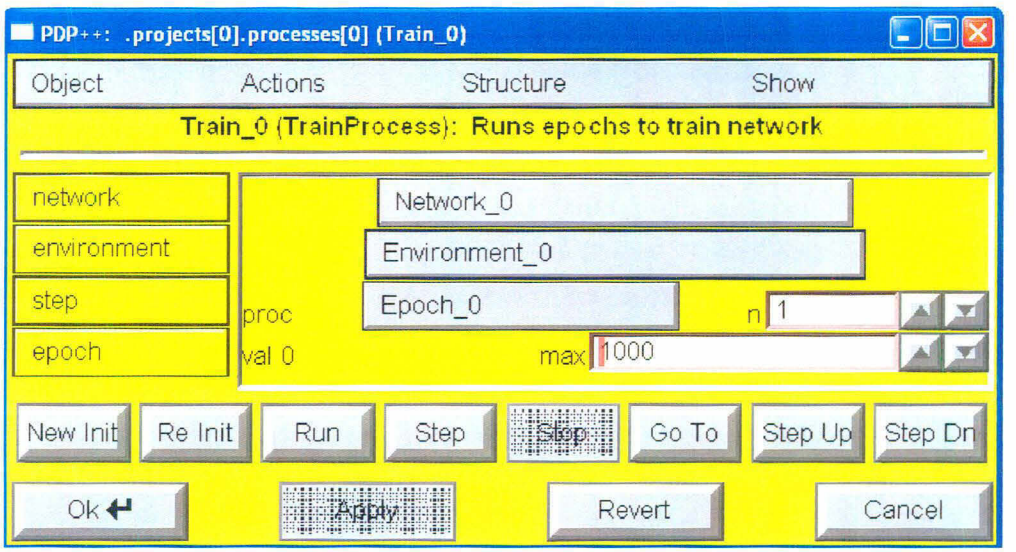


Figure 3.6. Control Panel for Trial Process

In this window, user can also specify the number of steps for the subsequent level (in this case the number of epochs). The network behavior can be analysed using the log windows.

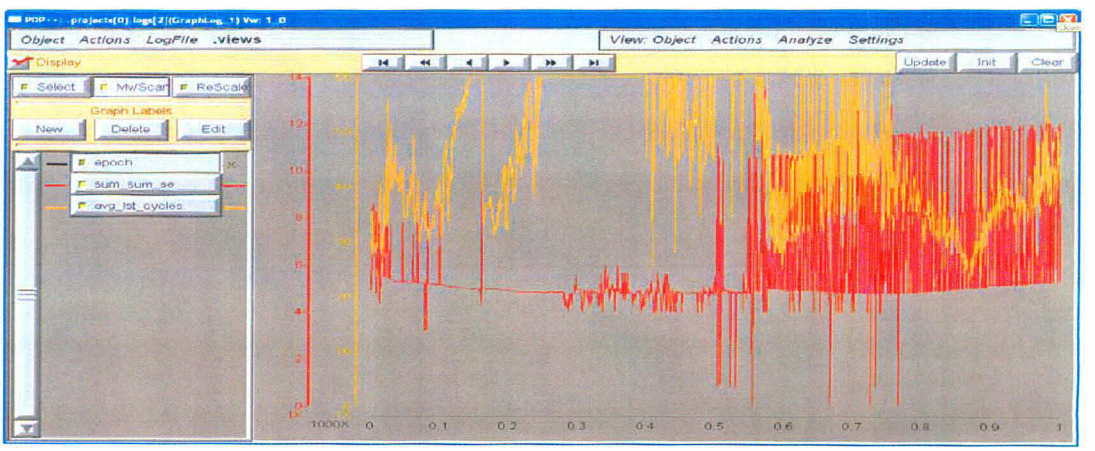


Figure 3.7. Graph Log View showing the 'sum square error' and 'average last cycle' at each epoch based on the above events and other default parameters

After designing the model, user can save the model. Each saved model contained various files as per the requirement of the model but saving the project from the root menu gives a compressed file which contained the entire required project file. This portable format feature helps the user to share the project among researchers.

The specifications of layers, units (neurons), connections, environments, events and logs can be edited at any time. For example to define the kWTA (k – winner takes all) parameter, we have to edit the layer specification by selecting the 'edit' option under 'specs' menu in project window. This option will open the edit window to edit the parameters of layer specification.

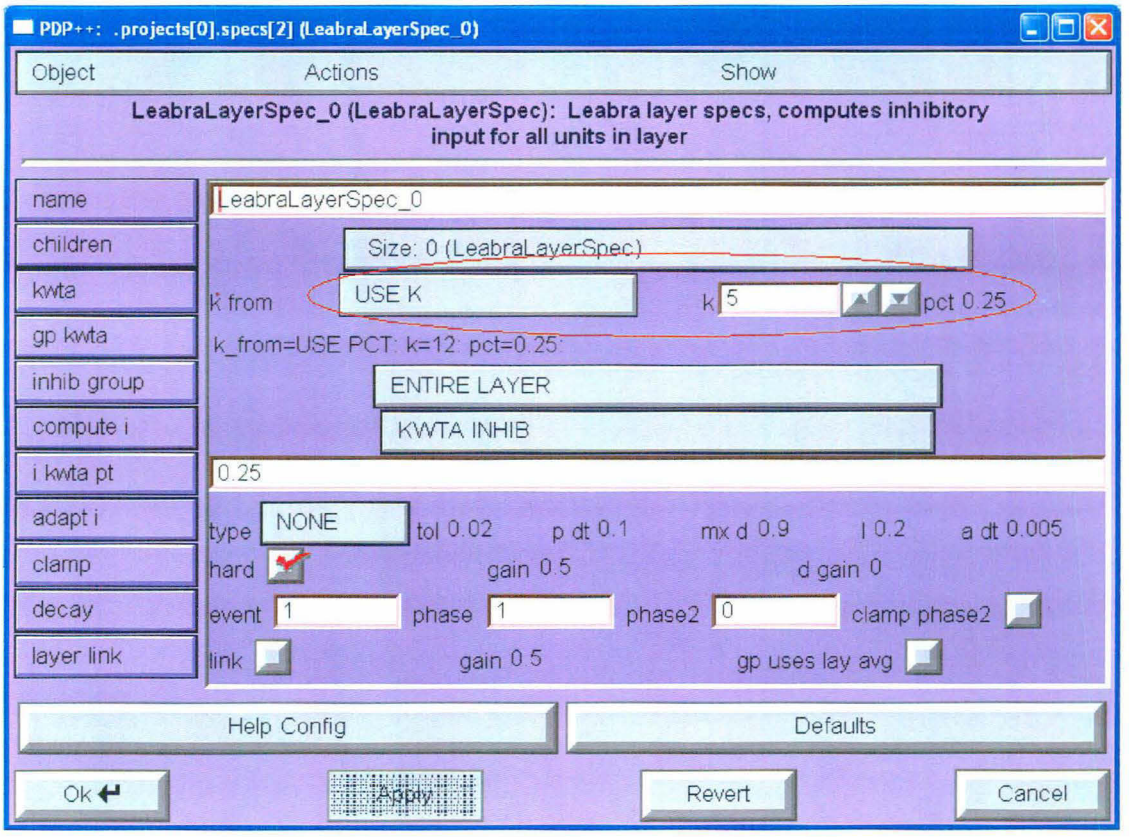


Figure 3.8. Leabra Layer Specification Window showing kWTA parameter

User can also change the weight calculation algorithm in this window and can be applied Hebbian or error-driven learning algorithm for weight calculation. Likewise user can edit any of the specification and can change the parameter according to the needs of the model. The software has the capability to store the continuous snaps during the learning cycles so that the user can show their work easily. The option

'Start Animation Capture' is given in network window under 'Action' menu. This option will automatically store the weight specific snaps of the network window during the learning process which the user can directly insert the snaps in their presentation software to show in seminar/ conferences.

Overall a Leabra modeling framework of PDP++ can give a flexibility to design the neurocomputational models which may depicts the real picture of biological processes of the brain. An advanced version of PDP++ has also come up called Emergent. Emergent has all the capability as PDP++ and additionally it gives a new 3-dimensional view of network, help on the specific object, more organized structure of the model's object hierarchy etc.

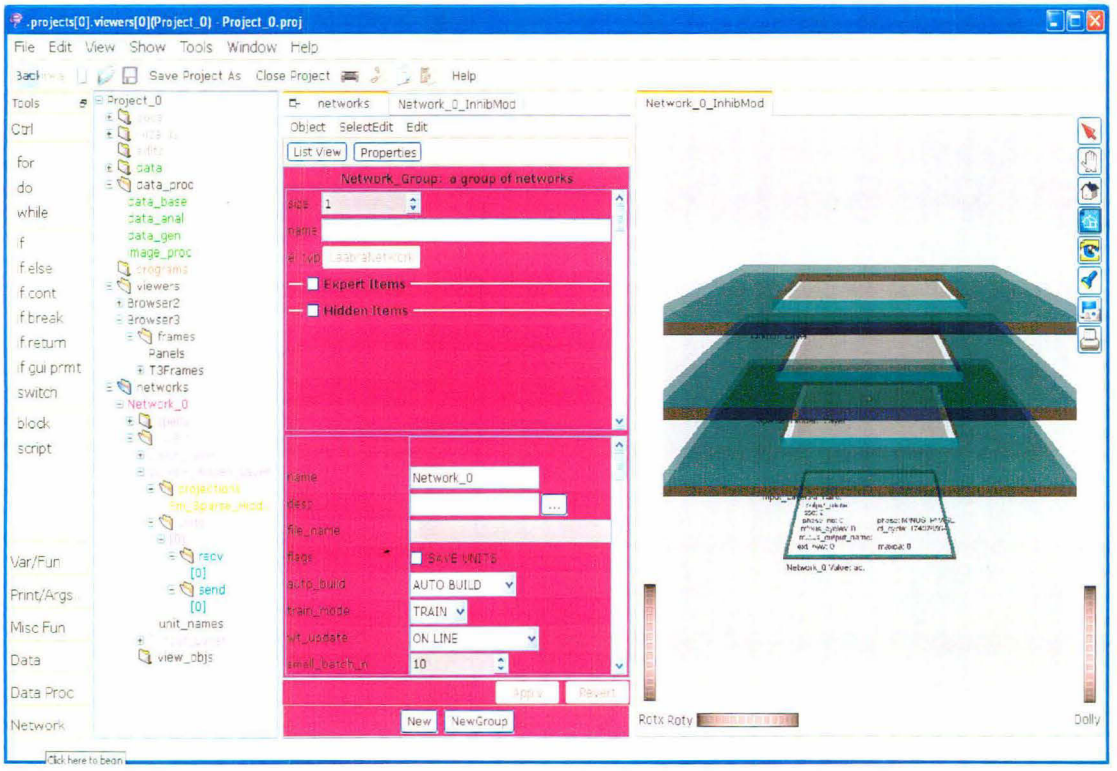


Figure 3.9. Emergent Package version 4.0.17

Chapter 4

Mathematical Model Explaining Relation between Hunger, Dopamine and Reward Rate

In various electrophysiological as well as behavioral experiments on animals, food is the reward for performing some task learning. One of the primary neurotransmitter system which is highly involve in reward pathway is mesolimbic dopamine pathway (Colle and Wise, 1988; Gratton et al., 1988; Old's 1990). While analyzing the behavioral data, most of the work fails to consider the level of hunger in animal conditioning. Hunger as well as food intake significantly change the level of dopamine in extracellular space. The dynamic change of dopamine level is responsible for vigor response for all dopamine – dependent behavior. We are giving a mathematical model, representing the dynamics of dopamine level in case of hunger and reward in subsequent sections.

4.1. Background

Hunger is seen to affect the action rate in two different ways. First, we explore the effects of chronic hunger. Animals are typically maintained below the base-line body weight for the duration of the conditioning studies. These studies can last from a few weeks to a few months period. What could be the effects of sustained starvation during this period, on the animal's behavior? It is observed that food deprived animals show lower level of extra-cellular dopamine as compared to the control animals (Bello et al., 2003; Carr et al.,2003). Behaviorally, these animals show a general sluggishness, but interestingly, they exhibit an above normal vigor for dopamine-dependent actions like eating or drinking (Pothos et al., 1995). Tonic release of dopamine is believed to be the primary source of extra-cellular dopamine, and hence, a drop in the extra-cellular dopamine level suggests some kind of deficiency in the tonic dopamine generation. The exact reason of this deficiency in case of chronic hunger is still unknown.

Studies in humans and animals have revealed that the dopamine system is controlled by a powerful homeostatic mechanism. This mechanism compensates for any changes in the level of extra-cellular dopamine by changing the rate at which tonic dopamine is synthesized and by changing the synaptic responsivity of the dopamine receptor neurons (Bello et al., 2003; Carr et al., 2003; Grace, 1991). Once again, the exact mechanism of this restoration process is not completely clear. However, it has been observed that a period of sustained dopamine decrease is followed by the development of new synapses in the dopamine receptor neurons. This increase has been reliably replicated in a number of studies where the dopamine levels were chemically suppressed (Grace, 1991). Recently, it has also been observed in studies where the dopamine levels were brought down by natural causes like food deprivation (Carr et al., 2003). It is believed that receptor responsivity could increase even without any visible development of new synapses (Carr et al., 2003; Grace, 1991; Pothos et al., 1995).

Hence, two opposing forces are at work in chronically hungry animals – The first mechanism decreases the dopamine level and the second mechanism compensates for the decreased dopamine levels. A side effect of the compensation mechanism is that the responsivity of the dopamine receptor neurons increases not only for the tonic dopamine, but also for the phasic dopamine release. Hence, over a period of sustained hunger, the animal's phasic dopamine system would become more and more responsive. This mechanism helps in explaining the selective increase in vigor for dopamine-dependent behaviors. Figure 4.1 summarizes these findings.

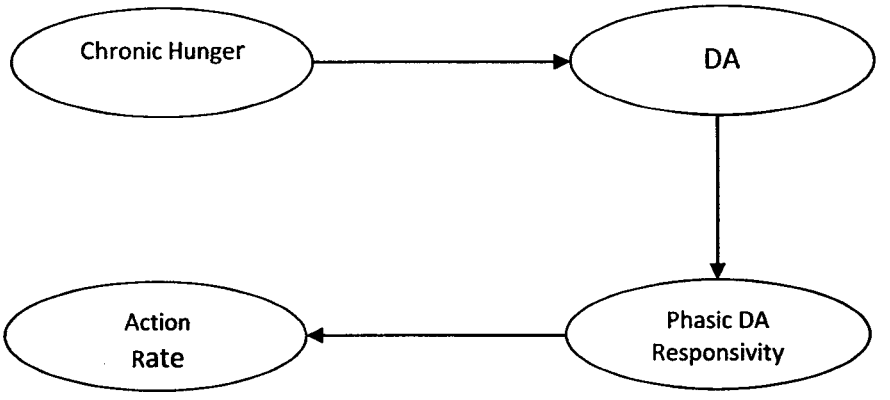


Figure 4.1: Chronic hunger results in a decrease in the amount of extra-cellular dopamine. This triggers the homeostatic mechanism, which increases the responsiveness of the phasic dopamine system. This results in a more vigorous dopamine-related response.

Next, we explore how hunger affects the action rate within an experimental session. It is observed that food consumption correlates with an increase in the dopamine level. This increase is seen only if the animal is hungry. Animals that have been fed ad-libitum do not show an increase in the dopamine level when they are fed (Bassero and Chiara, 1997).

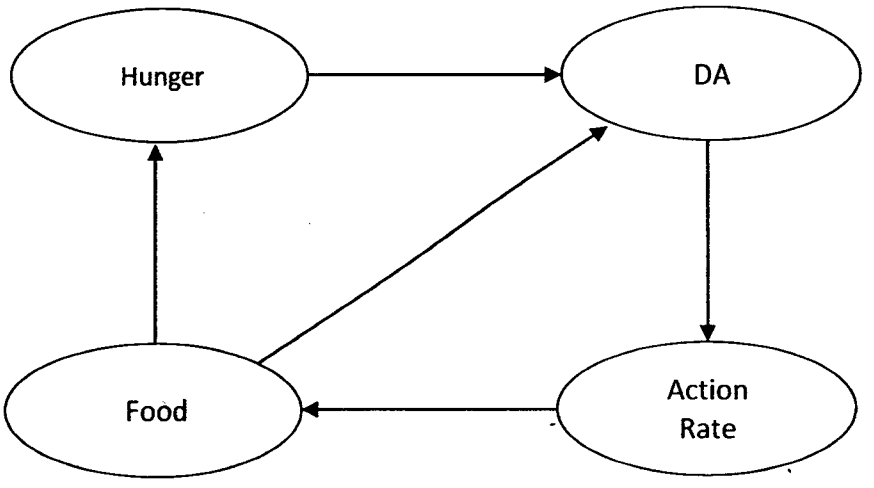


Figure 4.2: Action rate is directly proportional to the dopamine level. The rate at which food is received depends on the action rate. Food consumption decreases the level of hunger. Hunger and food consumption affect the dopamine level.

Second, changing the dopamine level artificially does not change the animal's appetite - the total amount of food consumed by the animal remains the same, despite the changes in the dopamine level (Salamone and Correa, 2002). These evidences, when pieced together, suggest that the process of food consumption, along with the current hunger level, together act to dynamically modulate the dopamine level (Young et al., 1992).

This dynamic regulation of the dopamine level seems to govern the action rate - an increase in the dopamine level causes an increase in the action rate and a decrease in the dopamine level causes a decrease in the action rate (Figure 4.2). First, there is evidence that the dopamine level directly correlates with the action rate (Young et al., 1992). Second, the animals that are administered with dopamine antagonists show slower rate of action (Salamone and Correa, 2002). Similarly, experiments in which the animals' dopamine level is enhanced show a more vigorous rate of action

(Cagniard et al., 2006). These evidences point towards the existence of a causal relation between the dopamine level and the action rate.

4.2. The Model

We present an abstract model of animal conditioning that is based on the behavioral and neuroscientific results discussed in the last section. The model consists of the following variables: Hunger represents the animal's hunger as a numerical score. EDA is a numerical representation of the extra-cellular dopamine level. ActionFreq signifies the action frequency. ChronicHunger signifies the chronic hunger, expressed as a percentage drop in the body weight. Reward is a binary value signifying whether the animal received a reward (a food pellet) or not. DAR is a numerical value representing the responsivity of the phasic dopamine system. In our model, time is divided into discrete steps, and the value of these variables is updated at each time step.

Action frequency in our model is directly proportional to the extra-cellular dopamine level as well as the responsivity of the phasic dopamine system. Hence, the action rate for the tth time step is computed as follows:

$$ActionFreq(t) = \alpha \times EDA(t) \times DAR \tag{4.1}$$

α is a model parameter. In the above expression, DAR does not depend on t. DAR changes over a time scale ranging from a few days to a few weeks. On the other hand, t, which represents the time period within a single experimental session, is too small to have any significant changes in the value of DAR.

In the above expression, the action rate is not shown to depend on phasic dopamine activity. Neuroscientific studies show that the phasic activation of dopamine neurons is responsible for triggering the reward-related action in the animals (Schultz, 2000). Phasic signals can be assumed to be incorporated implicitly since they are always required for action initiation. Also, the strength of these phasic signals, which varies during conditioning and extinction training can be modeled by varying the value of the parameter α in the above expression.

The level of hunger during a training session decreases as a function of the cumulative reward (food) value.

$$Hunger(t) = Hunger(t-1) - \beta \times Reward(t) \quad (4.2)$$

β is a model parameter. According to the above expression, the level of hunger can either decrease or remain the same during an experimental session. The training sessions in the behavioral experiments typically last between 30 to 60 minutes. This time period is too short for the animals to digest the consumed food and become hungry once again. Also note that the food consumed during previous training trials does not affect the hunger during the next training trial, as successive sessions of behavioral experiments are conducted after sufficient time gap to eliminate that possibility.

As described in the previous section, food consumption leads to the increase of dopamine and this release is contingent on the animal being hungry. This relationship between food consumption, hunger and the level of dopamine is captured as follows:

$$EDA(t) = \gamma \times EDA(t-1) + \lambda \times Sig\,(\phi \times Hunger(t) \times Reward(t)) \quad (4.3)$$

γ and λ are model parameters and Sig() represents the sigmoidal activation function. In the above expression, Hunger(t) x Reward(t) will be zero at time steps when no Reward is delivered. For time steps when reward is delivered, the value of the product will be equal to the level of Hunger. Hence, as per the above expression, the value of EDA will increase when a hungry animal gets a reward.

Dopamine responsivity in our model changes as follows:

$$DAR(T) = DAR(T-1) + \theta \times ChronicHunger(T) \quad (4.4)$$

θ is a model parameter. In the above expression, the use of T instead of t signifies the difference in time scales. t is used to represent time steps within an experimental session. In comparison, T denotes the number of days, as the change in dopamine responsivity does not happen over the period of a single experimental session. It requires a sustained food deprivation for a period of a few days to a few weeks.

We model the effects of dopamine depletion by scaling down the value of EDA. Hence, dopamine depletion to 80% of baseline in our model would mean that the value of EDA is 80% of the value that would be seen under normal circumstances.

Model parameters were determined through a combination of grid search as well as trial and error. The following values were used for the simulation results reported in subsequent sections: α = 0.3, β = 0.01, γ = 0.985, λ = 0.22, ϕ = 1.5 and = 0.4. The values of the variables were initialized as follows: Hunger(0) was set to 5.5 for hungry rats and 3 for pre-fed rats. EDA(0) was set to 0.5, and was restricted to range between 0.5 to 3.0. DAR was set to 1. As mentioned earlier, the value of Reward was either 1 (signifying the delivery of a food pellet) or 0 (no food) at each time step.

4.3. Model's Behavior in Different Scenarios

4.3.1. Chronic Starvation

Behavioral studies show that the degree and duration for which the animals are starved has a direct correlation with the responses vigor. For example, Cagniard et al (Cagniard et al., 2006) compared the responses of mice with varying degrees of starvation. Three sets of mice were used in their study, having 15%, 8% and 0% below baseline body weights. As expected, most vigorous response was seen in the mice maintained at 15% reduced body weight, followed by the mice maintained at 8% reduced weight followed by the normal mice. It is important to note that the mice learned the association between lever presses and food delivery before they were put on different food deprived regimes. Hence, the difference in the response rates cannot be attributed to the differences in the association strengths formed during conditioning. If these differences are not due to different association strengths, what could be the possible explanation?

Our model provides a simple explanation for these results. In our model, the effect of chronic hunger is captured by changing the dopamine responsivity (DAR) according to Equation 4.4. Since DAR is directly proportional to the action frequency, an increase in DAR results in a more vigorous response. Maintaining the animals at 10% below body weight for 5 days in our model resulted in an increase in the action rate from 34 per minute to 45 per minute.

Niv et al (Niv et al., 2005) have proposed an abstract reinforcement learning model of animal conditioning. Their model incorporates the effects of chronic starvation by changing the utility of the reward. For example, a reward that is worth 10 units to some animal would be worth 15 units to a hungrier animal. The action vigor in their model is directly proportional to the reward utility. Hence, changing the utility leads to a change in the action vigor. It is important to note that in their model, the utility of a particular type of reward would result in increased vigor only for the actions that are associated with that type of reward. In other words, for example, a starved animal should not drink water more vigorously. This is different from our model, where a change in DAR predicts that the animal will act more vigorously, not only for food, but for all dopamine-dependent actions.

Niv et al (Niv et al., 2005) do consider the evidence for a non-specific increase in action vigor in hungrier animals. They do so in the context of experiments where hungry animals are tested in a setup where they are free to perform some food-related action (like lever-press) as well as a drinking action. In such a setup, they explain away the increased vigor observed for the drinking action as the animal's desire to quickly revert to the food- related action. Our model predicts that the increased vigor of drinking is not contingent upon the availability of any food-related action. Instead, it is due to the increased dopamine responsivity.

4.3.2. Spontaneous Recovery

Spontaneous recovery is considered to be one of the most fascinating phenomenons related to the extinction of conditioning. The response vigor decreases during the extinction training. This decrement in the response vigor is seen to dissipate with the passage of time. This return of a portion of the originally learned behavior has been widely interpreted as evidence that extinction does not reverse the originally learned association (Gupta and Noelle, 2005; Rescorla, 2005). Many theories have been proposed as an explanation for the phenomenon of spontaneous recovery. One of the oldest explanations has been that the extinction-related associations get weakened with the passage of time due to the effects of non-specific interference. As another possibility, Bouton et. al. (Bouton, 2004) suggests that acquisition and extinction trainings are conducted in different temporal contexts and hence, these

effects tend to average out as the animal moves into a "new" context with the passage of time. In a similar theory, Devenport (Devenport, 1998) proposes that the response rate depends on the temporally weighted average reward value, with a higher weight given to the more recent events. Hence, just after extinction training, the weighted average is low, resulting in a less vigorous response. However, as the time passes, the temporally weighted average increases once again, resulting in increased responding.

In a typical spontaneous recovery experiment, the animals first undergo conditioning where an action is encouraged by associating it with a reward. This is followed by extinction training, where the action rate is seen to drop due to the withdrawal of the reward. Finally, after a period of rest, the animal is once again tested for its response rate. It is typically seen that the passage of time results in an increase in the response rate. Rescorla (Rescorla, 2001) incorporated an additional stage in the above described experiment. After the extinction training, he subjects the animals to a period of reacquisition. This was followed by the rest period and the test for response rate. He found that the animals responded with rates greater than those seen at the end of the reacquisition phase. He attributed this to the phenomenon to spontaneous recovery. He used this result as evidence in support of the dual pathway theory of animal conditioning. He suggested that extinction could involve the formation of a separate decremental association, and not a reversal of acquisition related association. Similarly, reacquisition might involve a slight strengthening of the acquisition related association and possibly a slight weakening of the extinction related association. Most of the extinction related association, however, must have survived through the reacquisition training. Over the rest period, this extinction related association must be going through a phase of decay or interference, causing the spontaneous recovery. He (Rescorla, 2005) wondered if there is a scenario where an opposite effect would be observed - i.e. the response rate decreases with the passage of time. He conducted a series of experiments, manipulating the order and duration of conditioning, extinction as well as rest periods. In all the cases, he found that that the rate of responding after the rest period was greater than the rate of responding before. In fact, in some of his

experiments, after the period of rest, the animals responded at rates that were higher than the rates ever reached during any of the training sessions. From this, Rescorla concluded that there must be something peculiar about the extinction related association that makes it weaker with the passage of time.

It should be noted that throughout his experiments, Rescorla maintained the animals at 80% of this normal body weight. This was true even for the rest period before the test of spontaneous recovery. The period of rest in his experiments was 5 days. Hence, it is likely that the phasic dopamine responsivity of these food deprived animals slowly increased during the rest period, causing a more vigorous responding during the spontaneous recovery test.

As explained before, in our model, the effects of extinction are captured by decreasing the value of the parameter α in equation 4.1. Decreasing α from the default value of 0.3 to 0.15 results, decrease in the lever press rate from 34 per minute to 7 per minute. Now, if we maintain the model 20% below the baseline body weight for the duration of 5 days, the lever press rate increases to a value of 16 per minute.

It should be noted that this explanation does not eliminate the need for the other theories of spontaneous recovery. In Rescorla's experiments, extinction training was conducted by the omission of rewards. Other conditioning studies, in which the extinction training was conducted by punishing the animals with a foot shock or some other undesirable event, report a different behavior – they report the spontaneous recovery of the fear response. Evidence suggests that fear responding is not governed by dopamine firing and we do not yet have a clear understanding of the mechanisms that might underlie the formation of fear related associations. In yet another variant, animals are trained to press a lever for food delivery. Extinction training involves the omission of reward for the lever pressing behavior in addition to reward delivery for a previously unrewarded behavior (like the pressing of some other lever). Over time, the animals stop pressing the first lever and start pressing the second lever. After a period of rest, it is seen that the behavior of pressing the first lever returns, such that animals now start choosing both the levers almost the

same number of times. While increase in phasic responsivity is still possible in such experiments, additional phenomena like memory consolidation (Wixted, 2004) may be playing some role as well.

Finally, a dopamine based explanation for the phenomenon of spontaneous recovery mitigates its applicability as an evidence for the dual-association hypothesis. Other evidence, however, still remain best explained via the dual association hypothesis (Gupta and Noelle, 2005) and hence, the dual-hypothesis cannot be completely ruled out.

4.4. Neurocomputational Model Exhibiting Relation between Reward and Hunger

As given in equation 4.2, the hunger is satiated by receiving the reward. Since reward is not given at each attempt, the hunger level goes high whenever the reward is not received. Hunger gets affected by the previous attempt reward value. It is something like Markovian process where the satiation of hunger is depends upon the reward received at previous time step.

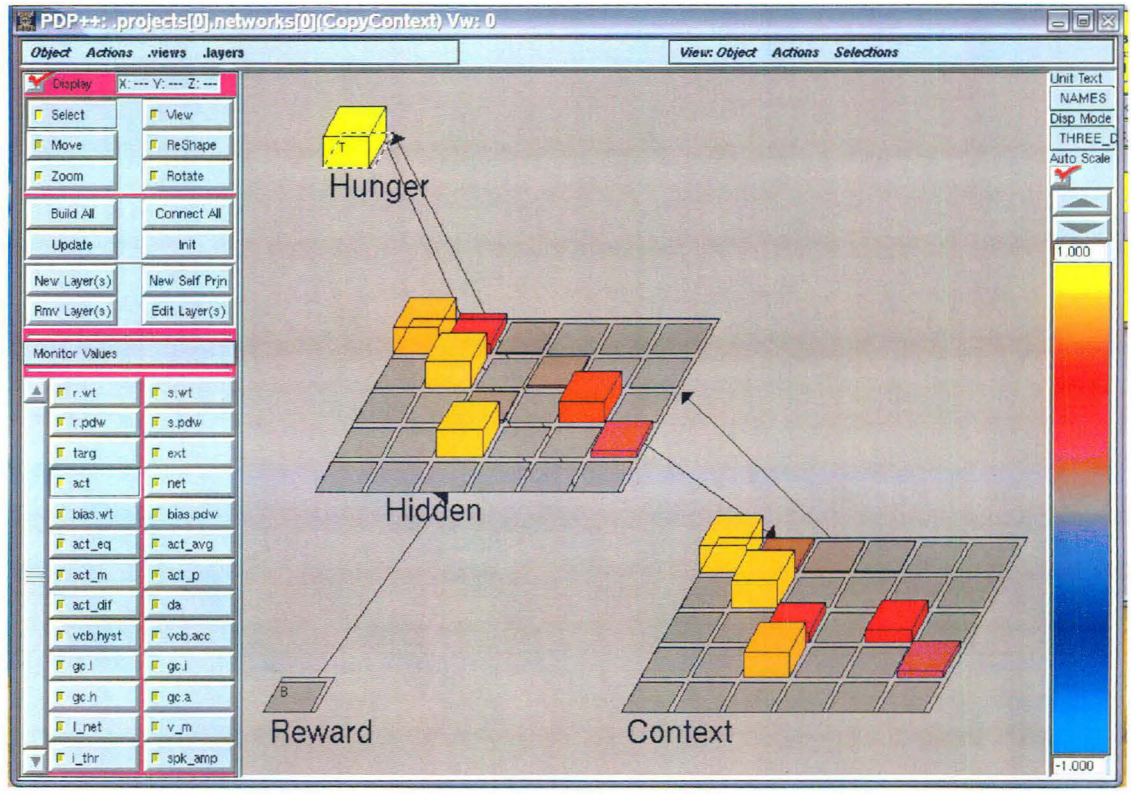


Figure 4.3: Neurocomputational model created on PDP++ showing the relation between reward and hunger and context layer to store and work as input layer for hidden layer for previous and next time step.

A small neurocomputational model is depicting the same behavior as of the equation 4.2. Here a context layer is introduced to store the values of hidden layer in previous step and works as an input layer in the current time step. Again the context layer stores the values of hidden layer at current time to use as input in the next time step. The size and unit specifications are the same for hidden layer and context layer so that the context layer can store the corresponding unit weights of hidden layer to supply in the next time step.

4.5. Conclusion

We have proposed a simple mathematical model based on the hypothesis that hunger and food consumption govern the dopamine levels and responsivity in the brain. These in turn govern the rate at which the animals perform the food-related action. Our model is able to provide parsimonious explanation for a number of behavioral phenomena. Numerous challenges still remain, however. First, our model is based solely on the evidence for correlations between hunger, dopamine levels and the response rates. The mechanism underlying these correlations remains to be explored. Second, we have only explored the effects of hunger on a previously learned behavior. The possible role of hunger and dopamine levels in the learning process also needs to be explored. Further, several factors other than hunger could possibly affect the dopamine levels. For example, there is evidence that stress results in the release of tonic dopamine. Also, action frequency could be affected by several factors other than just the dopamine level. Fatigue, other possible rewarding actions, other available rewards, time gap between action and reward and the amount of effort required for performing the action are some of the possible factors.

Finally, dopamine could be responsible for more than just governing the action rate. It could govern the action timing, action selection, reward selection, perception, and motor execution and so on. In summary, a number of challenges remain before we can have a clear understanding of the processes governing the conditioned behavior in animals.

5.1. Computational Model

Computer simulation is a process of representing a physical or theoretical system using computer. It is very convenient to model a system and to learn the dynamics of the system. Computer simulation gives various advantages like time-frame, cost, any type of danger involved in real situation etc. When a computer simulation added with computer graphics, virtual reality and animation, the system would become more realistic and interactive. With the help of interactive simulation software, a truck driver or aero plane pilot can be trained for every situation that could arrive in real situation. The person can learn without damaging the property, human life or hurting themselves (Fishwick PA, 1995).

Researchers use computer simulation to study the dynamic behavior of a system. Simulation frequently used by the researchers when:

- The model is very complex with many variables and interacting components
- The underlying variables relationships are nonlinear
- The model contains random variables
- The model output is to be visual in a 3D computer animation

Neurocomputational model is a type of neural network simulation, in which the neuronal level connectivity, connection weight calculation, and various other neuron level parameters can be manipulated, which helps to learn different brain processes. Researchers design the models by writing their computer programming codes or use any suitable pre-defined platform. There are number of neural network simulation software available in the market. Some are free and some are available after purchasing and accepting the license agreement. Research in artificial neural networks gives some hope for imitating the biological neural systems technically. All

neural networks are therefore designed to resemble in their biological counterpart. Simple biological neuron networks may consists of large number of neurons with dense interconnections. In the human brain there are about 10^{11} neurons where each one is connected to roughly 10^3 to 10^4 other neurons, i.e., there are more than 10^{14} interconnections called synapses (Hauser R et. al, 1991).

In this chapter, a neurocomputational model designed by us, showing the dynamics of dopamine in reward based learning. The basis of the model is a neurocomputational model designed by Frank MJ (Frank MJ., 2005) showing the dopamine modulation in the basal ganglia for medicated and non medicated Parkinsonism.

5.2. Dopaminergic Pathways in Basal Ganglia

In cognitive neuroscience, it is too difficult to model more than one brain regions if one region is affected by modulating functions in other brain regions. This issue can be easily seen in the effect of dopamine in the basal ganglia, which are critical for many aspects of cognition (Nieoullon, 2002). Many researchers assume that its function is to encode detailed aspects of SR mappings (e.g., Packard & Knolton, 2002). Others believe that different modulatory role of the basal ganglia to facilitate or suppress stimulus-response like associations that are represented in the cortex (Hikosaka, 1998; Mink, 1996). Frank proposed a neurocomputational model suggesting the dopamine activity in basal ganglia, as dopamine level changes in response to different behavioral events. Model also shows the variety of cognitive deficits due to dopamine dysfunction in basal ganglia, as in Parkinson's disease.

The cognitive deficits in Parkinson's disease can be divided into two general classes: patients impaired at task involving attentional processes (Rogers et al., 1998; Partiot Verin & Duois, 1996; Gotham, Brown & Marsden, 1988) or patients with implicit learning deficits (Knowlton, Mangels & Squire, 1996). Yet Parkinson's disease patients are impaired at implicit sequence learning and implicit categorization (Maddox & Filoteo, 2001; Jackson et al., 1995).

Studies show that Parkinson's disease involves damage to dopaminergic cells in basal ganglia (Kish et al., 1988) and the damaged basal ganglia is interconnected to a functional circuit with prefrontal cortex, producing frontal deficits. For more understanding, the role of dopamine in basal ganglia is required. There are two main types of cells in striatum responding differentially to phasic changes in dopamine occur during error feedback. This causes the two groups of striatum cells to independently learn positive and negative reinforcement values of responses, and ultimately acts to facilitate or suppress the execution of commands in the frontal cortex. Because these cortical commands may differ widely in content, damage to basal ganglia dopamine level gives rise to seemingly unrelated deficits.

Observational experiments on Parkinson's disease patients show that they are selectively impaired in cognitive procedural learning tasks that involve trial-by-trial error feedback. In case of observational implicit learning tasks (e.g., artificial grammar and prototype learning), patient performance is normative (Reber & Squire, 1999). Comparing both observation suggests that feedback-mediated learning occurs in basal ganglia and is therefore disrupted in Parkinson's disease. Feedback may modulate dopamine release in basal ganglia which, in addition to having a performance effect on response execution, is critical for cognitive reinforcement learning.

5.3. Phasic Burst and Dopamine Mediated Learning

Studies show that dopaminergic cells fire phasic bursts in response to unexpected reward (Schultz, 1998). Again the dopaminergic firing dips below baseline when a reward is expected but not received (Hollerman & Schultz, 1998; Schultz et al., 1993). These changes in extracellular levels of dopamine during feedback are critical for learning. Dopamine D1 receptor stimulation leads to long-term potentiation whereas D2 stimulation inhibits the long-term potentiation. Likewise, long-term potentiation is blocked by D1 antagonists and enhanced by D2 antagonists (Centonze, 2001). Due to modulation of dopamine in case of cellular excitability, associative learning may be enhanced in the presence of dopamine. Thus, the usefulness of recently active synapses may be reinforced by a burst of dopamine

acting as a "teaching signal," leading to a learning of rewarding behaviors (Wickens, 1997). Phasic bursts and dips of dopamine occur during positive and negative feedback which modifies the synaptic plasticity leads to effect on learning trial-and-error tasks.

5.4. Deficits Induced by Medication in Parkinson's Disease

The commonly used treatment for Parkinson's disease is L-Dopa therapy which is a dopamine precursor. Studies on medicated versus non-medicated patients show that it ameliorates task switching deficits in Parkinson's disease, but it impairs performance in "probabilistic reversal" (i.e. learning to reverse stimulus-reward probabilities after prepotent response are embedded) (Cools et al., 2001; Swainson et al., 2000; Gotham et al., 1998). Observations also show that dopaminergic damage in early stage Parkinson's disease is restricted to the dorsal striatum, leaving the ventral striatum with normal dopamine levels (Agid et al., 1993; Kish et al., 1988). Overdose medication refills the dopamine in ventral striatum and therefore detrimental to tasks that recruit it. Reversal learning depends on the ventral striatum and ventral prefrontal cortex in monkeys and recruits same area in healthy humans (Cools et al., 2002; Dias et al., 1996; Stern et al., 1995).

Frank concluded from the above explanations that during positive feedback, phasic bursts of dopamine may still be released. A higher level of tonic dopamine might functionally eliminate the effectiveness of phasic dips in dopamine during negative feedback. A dopamine agonist would continue to bind to receptors, as it is not modulated by feedback/ reward is endogenous dopamine. This by-product dopaminergic medication may eliminate an important aspect of the natural biological control system – the ability to quickly unlearn previously rewarding behaviors. To understand the dynamics of dopamine in basal ganglia, it is required a general description of basal ganglia circuits and functions.

5.5. Neuroanatomy and Biochemistry of Basal Ganglia

Basal ganglia can be thought to act as a brake on competing motor actions that are represented in the motor (or premotor) cortex. It modulates the motor execution by

signaling "Go" or "No-Go" (Hikosaka, 1989). At a time only the most appropriate motor command get executed.

Basal ganglia receive the input from striatum. The striatum is formed collectively by caudate, putamen and nucleus accumbens. The striatum receives input from various cortical areas and projects through the globus palidus and substantia nigra to the thalamus, and at last closing the circuit back to the area of the cortex from which it received i.e. premotor cortex (Alexander et al., 1986). Most of the projection cells in striatum which carry information to transmit to basal ganglia are GABAergic medium spiny neurons.

These neurons project to the globus and substantia nigra via two pathways which have opposite effect on the excitation/ inhibition of the thalamus (Alexander et al., 1990b). The direct pathway supports the execution of response whereas indirect pathway inhibits them. Cells in the direct pathway protect from striatum and inhibit the internal segment of globus pallidus. In the absence of striatal firing, the globus pallidus tonically inhibit the thalamus, so the excitation of direct medium spiny neurons and resulting globus pallidus inhibition serves to disinhibition the thalamus. The double-negative invoked by this disinhibition does not directly excite the thalamus, but instead simply enables the thalamus to get excited from other excitatory projections (Frank, Loughry et al., 2001; Chevalier & Deniau, 1990), thereby providing the gating function. Cells in the indirect pathway inhibit the external segment of the globus pallidus, which tonically inhibits it. The net effect of indirect medium spiny neurons excitation is then to further inhibit the thalamus.

When direct pathway striatal cells disinhibit the thalamus, excitatory thalamocortical projections increase the activity of the motor command that is currently represented in motor cortex so that its execution is facilitated. This implies that the direct pathway cells release a "Go" signal to select a given response. Indirect pathway activity, with its opposite effect on the thalamus releases a "No-Go" signal to suppress the response.

How these two pathways interact or are they work independently, is controversial. Both pathway converges to globus pallidus before either disinhibiting or further

inhibiting the thalamus, it would seen these two pathways act competitively to control basal ganglia output (Percheron & Filion, 1991).

As shown in figure 5.1, this competitive dynamic occurs in parallel for multiple responses. This may allow for selective control of different responses, so that one response may be enabled, whereas others are suppressed. Selection of a given resp-

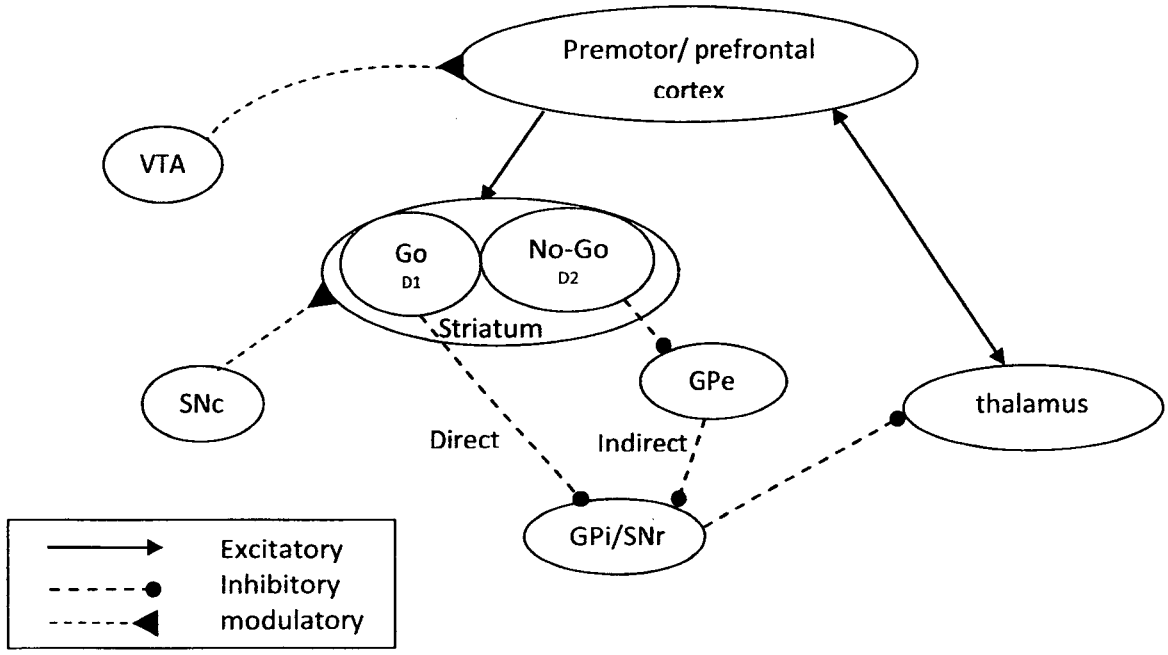


Figure 5.1: The corticostriato-thalamocortical loops, including the direct and indirect pathways of the Basal Ganglia. The Go cells express the D1 receptor, and the No-Go cells express the D2 receptor. DA from the VTA projects to the ventral striatum (not shown) and the frontal cortex. GPi = internal segment of globus pallidus; GPe = external segment of globus pallidus; SNc = substantia nigra pars compacta; SNr = substantia nigra pars reticulata; VTA = ventral tegmental area (Frank MJ, 2005).

onse may involve a Go signal to one area of thalamus, in conjunction with a "No-Go" signal to thalamus areas involved in competing responses. This competing dynamics is resolved by the modulatory signals given by SNc layer. These modulatory signals are received by the 'GO' as well as 'NO-GO' neurons in the striatum. From striatum, the two pathways are divided for disinhibiting or inhibiting the response signals of thalamus.

5.6. Cellular Mechanisms of Dopamine in Basal Ganglia

Basal ganglia circuitry is highly modulated by phasic changes in dopamine level. Dopamine is act as excitatory for D1 receptors and inhibitory to D2 receptors as

explained in previous section that striatal cells of the direct pathway contains D1 receptors where as indirect pathway striatal cells contains the D2 receptors and the relative levels of expression is asymmetrical. Thus, increased levels of dopamine activate the direct/ Go pathway and suppress the indirect/ No-Go pathway (Brown et al., 2004; Gurney et al., 2001).

5.6.1. D1 in Support of Go Cells

Evidences suggest that dopamine effectively enhances contrast or increases the signal-to-noise ratio by amplifying activity of the most active cells while inhibiting the least active one (Cohen et al., 2002; Foote et al., 1987; Rolls et al., 1984). Observations suggest that dopamine acting via D1 receptors can sharpen contrast by amplifying active cells that are in their up-states and inhibiting those in their down-state from firing spontaneously (Wilson et al., 1996). This may have the effect of increasing the signal-to-noise ratio, because cells encoding the relevant signal receive temporally coherent synaptic input from multiple cortical afferents and are therefore in their p-state, whereas those reflecting biological noise or other irrelevant background signals may be in their down-state and only firing spuriously. The increased signal-to-noise ratio in the direct pathway may help to determine which among several responses is most appropriate for selection.

5.6.2. D2 Inhibits No-Go Cells

The D2 inhibition effect on neuronal excitation was not found to be dependent on the membrane potential of the target cell hence D2 activation reduces the current which is enhanced by the L-type Ca channel, therefore reducing neuronal excitability. As explained earlier, D2 receptors predominate in the indirect/ No-Go pathway, which is thought to act as a break on a particular action or set of actions. Dopamine can then act in releasing the break, by inhibiting the No-Go activity via D2 receptors and allowing the Go pathway to exert more influence on basal ganglia output. This also gives to basal ganglia a clear state for selecting/ gating a response to be executed.

In case of Parkinson's disease patients, the lack of dopamine in basal ganglia results in difficulty initiating motor commands. Due to unreasonable dopamine release in basal ganglia, the system is in a tonic state of No-Go because as overactive indirect pathway leads to excessive cortical inhibition (Jellinger, 2002; Fillion & Tremblay 1991). This unreasonable release of dopamine, the balance is shifted to Go and the particular response that is executed may depend on level of activity of different subpopulations – representing different responses – in the direct/ Go cells. The D1 contrast enhancement mechanism described above would aid in selecting the most appropriate response by boosting its associated neural activity, while suppressing that of all other Go cells.

5.6.3. Dopamine in Basal Ganglia: Effects on Synaptic Plasticity

A consequence of dopamine Go/ No-Go activity level is that they drive activity-dependent learning to synaptic input. More active cells of basal ganglia undergo long term potentiation, whereas less active cells undergo long term depression (Bear & Malenka, 1994) and affect on plasticity in two basal ganglia pathways. If dopamine bursts during reinforcement are adaptive, they should have the complementary effects of increasing Go learning while decreasing No-Go learning so that reinforced responses are more likely to be facilitated in the future. This also enhances then long term potentiation in Go cells. The inhibitory effects of dopamine in the indirect pathway may induce long term depression in No-Go cells so that they learn to become less active.

Likewise, in case of dopamine dips, this, if they are adaptive, would enhance No-Go learning so that non-reinforcing responses are actively suppressed in the future. The increased No-Go activity should induce long term potentiation in No-Go cells because No-Go cells from indirect pathway acts as dopamine inhibitor. This hypothesis is indirectly supported by examining the effects of D2 receptor blockade, assuming that dopamine dips decrease D2 stimulation and should therefore have the same qualitative effects on the indirect pathway as D2 barrier (Finch, 1999; Robertson et al., 1992).

5.7. Neural Model of Basal Ganglia and Dopamine

To prove the hypothesis for two pathways in basal ganglia to work as Go/ No-Go associations can be substantially strengthened by testing its feasibility in a computational model. The model should incorporate all the key elements to generate novel prediction at the underlying source of cognitive dysfunctions in Parkinson's disease. If validated, it can also be used as a tool to understand complex involvement of dopamine in the basal ganglia in other neurological disorders.

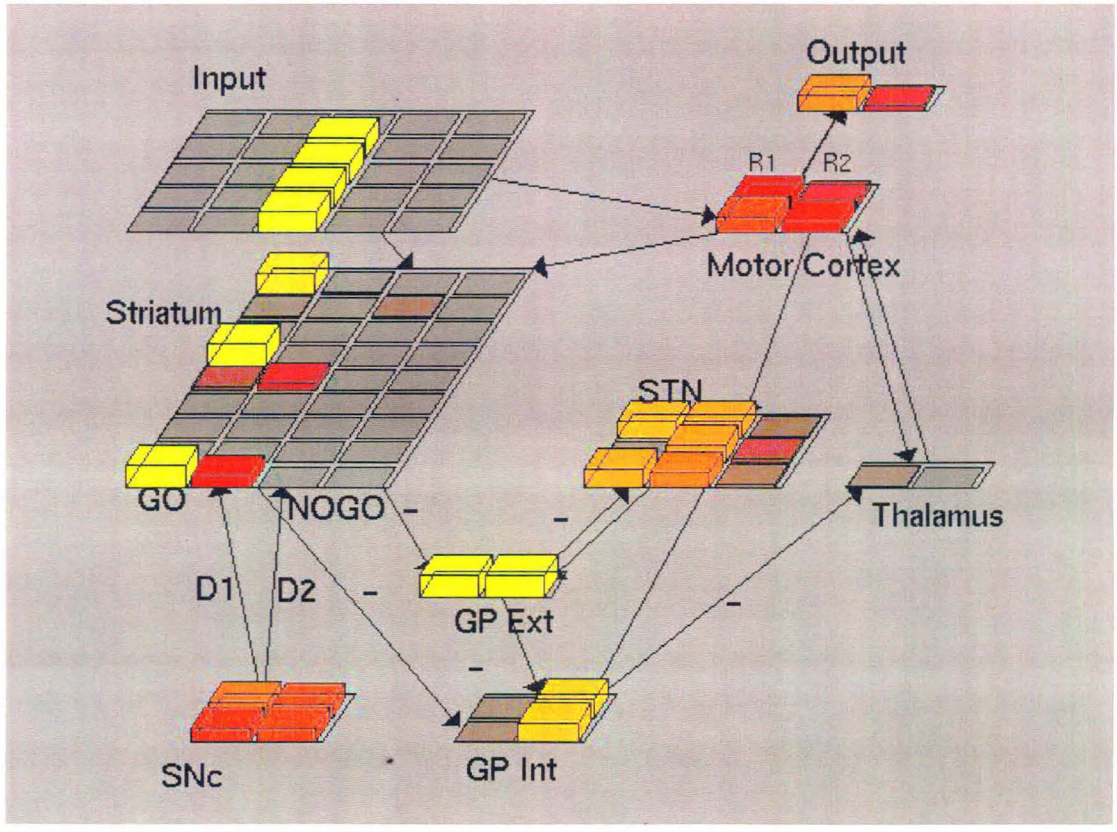


Figure 5.2: Neurocomputational model showing the basal ganglia and modulation of dopamine as suggested by Frank (Frank MJ, 2005).

A similar kind of model is proposed by Frank (Frank, 2005). The neural network model, as shown in figure 5.2, learns to select one of two responses to different input stimuli. As discussed in previous sections, the direct and indirect pathways in the model allow learning conditions that are appropriate for gating as well as those for suppressing. Parallel sub-loops independently modulate each response, allowing selective facilitation of one response with concurrent suppression of the other. The modulatory effect of dopamine is included in the projections from substantia nigra pars compacta to the striatum. Phasic bursts and dips in substantia nigra pars

compacta (shown as SNc in figure 5.2) firing arises from correct and incorrect responses respectively. These phasic changes drive learning by preferentially activating the direct pathway after a correct response and the indirect pathway after an incorrect response.

5.7.1. Mechanics of the Model

The model is based on Leabra framework (O'Reilly & Munakata, 2000) in which the units are "point-neuron" function using rate-coded output activation. There are separate excitatory and inhibitory synaptic input channels. Synaptic connection weights were trained using reinforcement learning version of Leabra. The learning algorithm includes two phases, allowing simulation of feedback effects and it is more close to behavioral phenomena than the standard error backpropagation. In the "minus phase," the network settles into activity states on the basis of input stimuli and its synaptic weights, ultimately resettles a response. In the "plus phase" the network resettles in the same manner but with the change in dopamine level. An increase dopamine level for correct response and a dip for incorrect response are applied by the substantia nigra pars compacta layer. The Connection weights are then adjusted to learn on the difference between activity states in the minus and plus phases.

5.7.2. Overall Network Division of Labor

The objective is to select either response 1 (R1 in figure 5.2) or response 2 (R2 in figure 5.2), depending on the task and sensory input. The input layer is directly connected to Premotor cortex but the connections are not so strong to trigger the response. The Premotor cortex also needs some bottom-up support from the thalamus. The job of the basal ganglia is to integrate stimulus input with the dominant response selected by the Premotor cortex and on the basis of what it has learned in previous experience, either facilitate (Go) or suppress (No-Go) that response.

In the model, there are two parallel sub-loops that are isolated from each other, separately modulating the two responses. This allows the selection of one response

while suppressing the other response. The four columns of striatum layer is divided into two parts, the left two columns are representing the "Go" units with separate columns for response 1 and response 2 and the two right columns are representing "No-Go" units with separate columns for response 1 and response 2. Thus, the four columns in the striatum represent, from left to right, "Go–Response1," "Go–Response2," "No-Go–Response1," and "No-Go–Response2."

The Go columns of striatum only projected to the layer "internal segment of globus pallidus" (represented as GPi in figure 5.2) form a direct pathway. The "No-Go" column to the layer "external segment of globus pallidus" (represented as GPe in figure 5.2) form indirect pathway. GPe columns inhibit the associated column in GPi, so that striatal Go and No-Go activity have opposing effects on the GPi. At last, each column in the GPi tonically inhibits the associated column of the thalamus, which is reciprocally connected to the Premotor cortex.

The network architecture simply supports the existence of connections, but how these ultimately influence behavior depends on their relative strengths. The network starts off with random weights and representations in both the basal ganglia and cortical layers are learned. Distributed activity within each striatal column enables different Go and No-Go representations to develop for various stimulus configurations during the course of training.

5.7.3. Simulated Effects of Dopamine

As discussed earlier, the dopamine acts differently on the units of striatum. To simulate this phenomenon, SNc layer project two different types of projections to striatum layer, one is excitatory D1 connection which only connects to "Go" column units and another other is inhibitory D2 connection which only connects to "No-Go" column units. One role of D1 activity is to enhance the strongly participating units. This can be done by increasing the activation threshold so that weakly active units do not exceed firing threshold and are suppressed. The effects of D2 activity are inhibitory, suppressing the No-Go cells. Thus, for a high amount of simulated DA, contrast enhancement in the direct pathway supports the enabling of a particular Go response, whereas the indirect pathway is suppressed.

5.7.3.1. Dopamine Modulates Learning

At the start of each trial in the minus phase, a tonic level of dopamine is maintained by setting the SNc units to be semi-active with activation value 0.5. At the initial stage of training, the network selects a random response, dictated by random initial weights in Premotor cortex. If the response is correct a phasic increase in SNc units set to have an activation value of 1.0 i.e. high firing rate. This dopamine phasic burst causes a more logical Go representation in the striatum to associate with the rewarding response that was just selected. In case of incorrect response, a phasic dip occurs with all SNc units set to zero activation results the network to learn No-Go for the selected incorrect response.

The model learns on the basis of the difference between activity states in the minus phase and plus phase, which only differ due to phasic changes in dopamine, hence it also ruled out any possibility of supervised learning. Weights from input layer and the Premotor cortex are adjusted so that over time, the striatum learns which response to support and which to inhibit in the context of incoming sensory input. Premotor cortex also learns to favor a given response for a particular input stimulus from the input layer. Thus the basal ganglia initially learns which response to gate via phasic changes in dopamine ensuing from random cortical response, and then this learning transfers to the cortex ones it starts to select the correct response more frequently. This also mimic the biological basis that basal ganglia is not a stimulus-response module rather it modulates the allowing or disallowing the response selected by the Premotor cortex.

5.8. Neurocomputational Model for Examining the Role of Dopamine Influenced by Reward Rate

Our model is based on Frank's model (Frank MJ, 2005), in which he showed the dynamic behavior of neurotransmitter "Dopamine" in basal ganglia. The most affected biological pathway in case of reward based learning is the dopaminergic pathway (see section 2.4). Frank's model is the proven one which shows the behavior of dopamine dependent actions in case of intact and Parkinson's disease

patients. This model is closely related with our work and could be used to show the rate of reward in case of normal and dopamine depleted animals.

5.8.1. The Model

In most of behavioral experiments conducted by researchers actually performed on animals. The animals used in such experiments are food deprived animals. Researches use food or juice as reward so that animal can learn some specific action required for the experiments. In neurocomputational model of basal ganglia proposed by Frank doesn't have such type of elements. The Frank's model is described in section 5.6.

We modified the model proposed by Frank, and added some elements to show the food reward mechanism. Studies revealed that Conditioned appetitive stimulus increases extracellular dopamine level in nucleus accumbens (Datla KP, 2002). The neurocomputational model should also depict the same dynamics behavior of dopamine as in the brain.

5.8.1.1. Satiation Layer

A satiation layer is added in the model to specify the level of satiation during the food intake. The unit weights of this layer are controlled by scripts. Whenever the model choose the correct response, the unit weights of satiation layer goes high showing the hunger goes down.

5.8.1.2. Connection between Satiation Layer to Striatum Layer

The connection between satiation layer to striatum layer is a tassel projection. Because the level of hunger will reinforce the network for correct response R1, hence we also have connected R2 column units among the 'GO' units and R1 column units among the 'NO-GO' units of striatum. Now for each correct response, when the network rewarded, the satiation layer value will be incremented (i.e. hunger decreases) by small amount (in our simulations 0.0005). The increase in satiation layer weights actually reinforces the R2 response selection from 'GO' pathway and R1 response selection from 'NO-GO' pathway.

5.8.1.3. Script Controlled Manipulation

PDP++ software also provides to write scripts and execute it at run time during the simulation. Since the behavior of satiation layer is somewhat dynamic i.e. we only have to increase units' weight when the model choose the correct response. To include this behavior, we added some scripts to manipulate the satiation layer units' weight. An object is defined in model which contains a current value of unit weights of satiation layer. During one trial, if the model chooses a correct response, this object weight is updated and assigned to the satiation layer units so that next time the unit weights will update from its previous value.

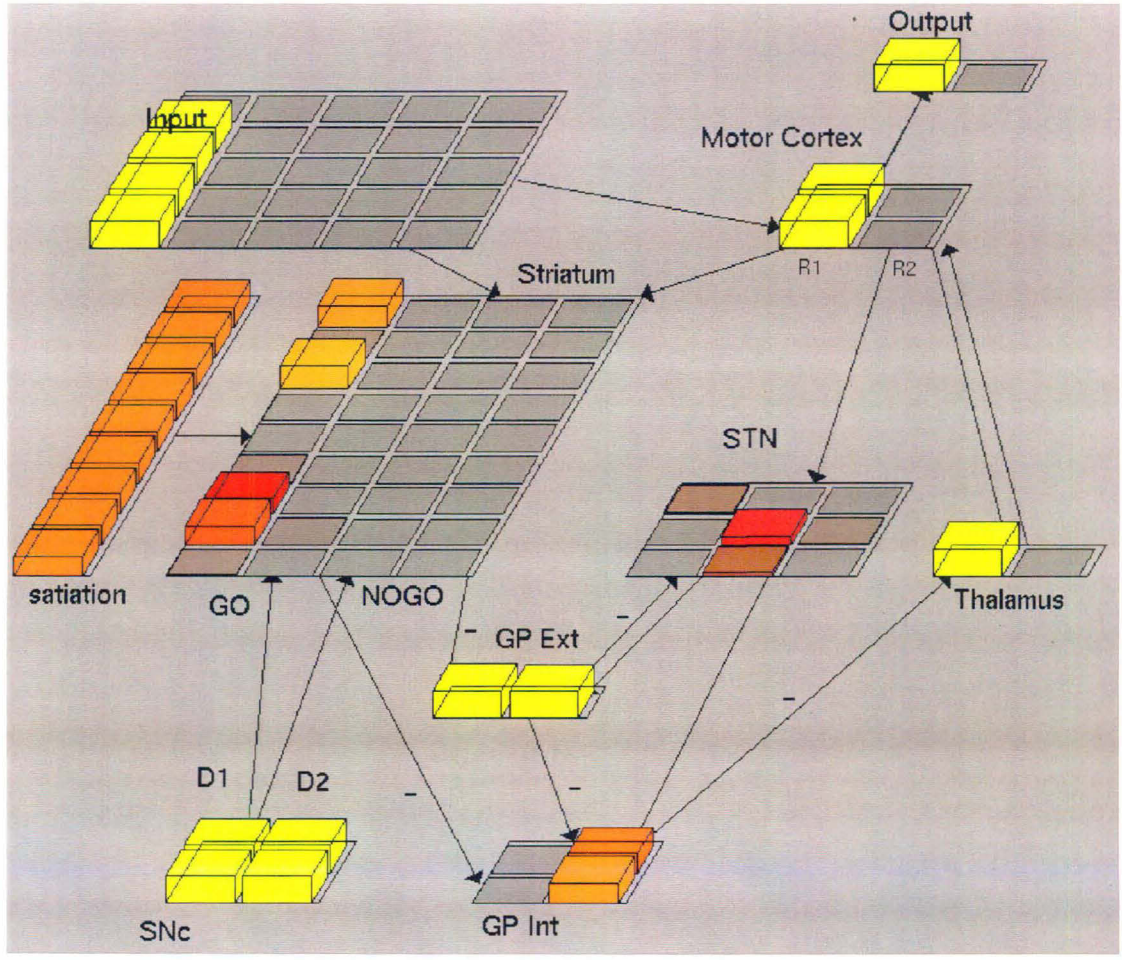


Figure 5.3: Neurocomputational model showing the satiation layer and connection between satiation layer and striatum.

5.9. Conclusion

The present model explores the dynamics of dopamine modulation in basal ganglia. This would help to understand the complex brain processes in experimental brain

research. The exploration is limited to the some basic mechanisms of the brain among the complex brain processes. Like it does not include the causes of phasic burst and dips in SNc layer. This also doesn't explore the fact that as learning progresses and rewards become expected, phasic burst of dopamine no longer occur during reward. The reward is instead transferred to an earlier stimulus that predicts reward. These all would be added in future work so that the model also proves the "temporal differences" as given in reinforcement learning algorithm (Sutton, 1988).

The behavioral experiments has performed on rats by researchers which describe the different scenarios for affected hunger level by manipulating the reward ratio or the by administrating the dopamine antagonist. In our work, the mathematical and neurocomputational models are been run for sufficient number of trials and the results are described in the following sections. The results are also compared with the behavioral data compiled by previous researchers.

6.1. Change in Action Frequency within a Session

6.1.1. Behavioral Experiment Results

Salamone et al (Salamone et al., 1995) conducted a number of conditioning experiments to explore the relation between dopamine level and response vigor. A typical experimental session in their study lasted for 45 minutes. In one set of experiments, using a continuous reinforcement schedule, they measured the number of lever presses over successive 15-minutes intervals. They found that the

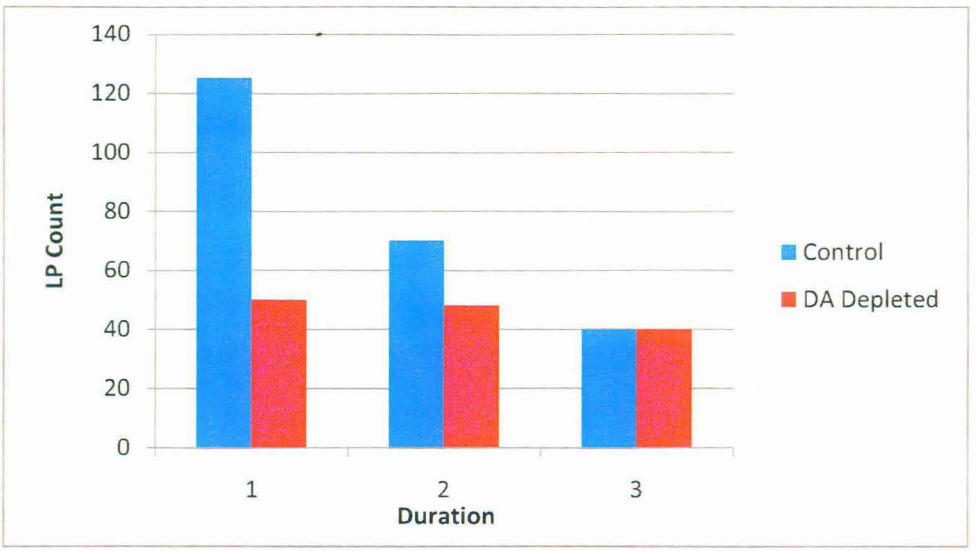


Figure 6.1: Behavioral results. In control rats, the lever press count is the highest in the first 15-minutes interval. The count falls in the second and the third 15-minutes intervals. The dopamine depleted rats show a significantly smaller lever-press count than the control rats during the 15 minutes subsequent intervals.

maximum number of lever presses occurred during the first 15-minutes. The number of lever presses decreased over the next 2 15-minutes intervals as shown in figure 6.1. They conducted the same experiment with rats that were administered with dopamine antagonists. The number of lever presses over the first 15-minutes period was significantly smaller in dopamine depleted rats as compared to the control rats. However, the number of lever presses over the next 2 15-minutes intervals, was comparable to that of the control animals.

The food pellets provided to the rats in their experiments weighted 45mg each. In case of control rats, the total number of lever presses, and as a result, the food pellets obtained were approximately 250. Hence, the rats consumed an average of 11.25 gm of food within the 45 minutes long experimental session. This quantity of food is roughly in the range of a food deprived rat's appetite. Another observation made by Salamone et al. (Salamone et al., 1995) in their experiment was that the effects of dopamine depletion start to fade off as the days pass by.

6.1.2. Mathematical Model Results

According to the mathematical model explained in section 4.2, the rate of responding for the control rats' drop during the second and third time intervals due to their hunger decreases with the consumption of food pellets. Dopamine depleted

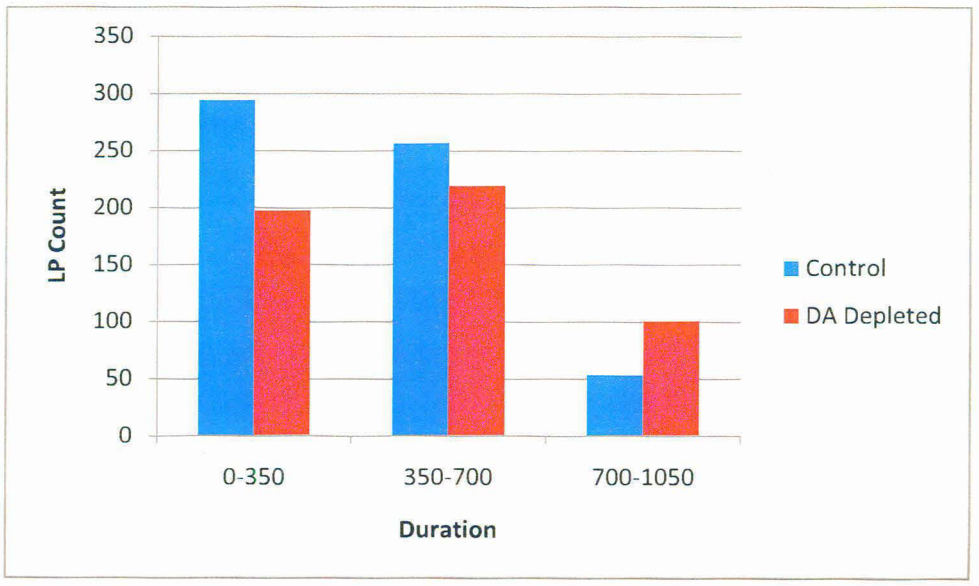


Figure 6.2: Simulation results. Number of lever presses was counted for three 350 time step intervals successively. Results match the behavioral data.

rats start acting at a lower rate. Hence, they get food at a slower rate. Delivery of food results in an increase in the dopamine levels, but this happens at a slower rate in the dopamine depleted rats. This is the reason for fewer responses in the first time interval. As the dopamine level builds up, the action rate increases and hence matches that of the control rats over the next 2 time intervals.

As shown in figure 6.2, our model explains these effects as well. Just as dopamine depletion triggered by food deprivation leads to heightened responsivity of phasic dopamine activity, a similar phenomenon is also seen in rats in which the dopamine level is artificially brought down through the administration of antagonists (Grace, 1991). As the time passes by after the injection of the antagonists, the homeostatic process slowly comes into action.

6.1.3. Neurocomputational Model Results

The neurocomputational model explained in section 5.8 is executed for 620 epochs. After 170 epochs, the network learns for the appropriate output and error goes to 0. After executing three consecutive 150 epochs and constant increase the unit weights of satiation layer, the number of correct response selection decreases. This is because the hunger decreases as the increase in weights of satiation layer units which directly affect the 'GO' units of striatum layer. In case of dopamine depletion,

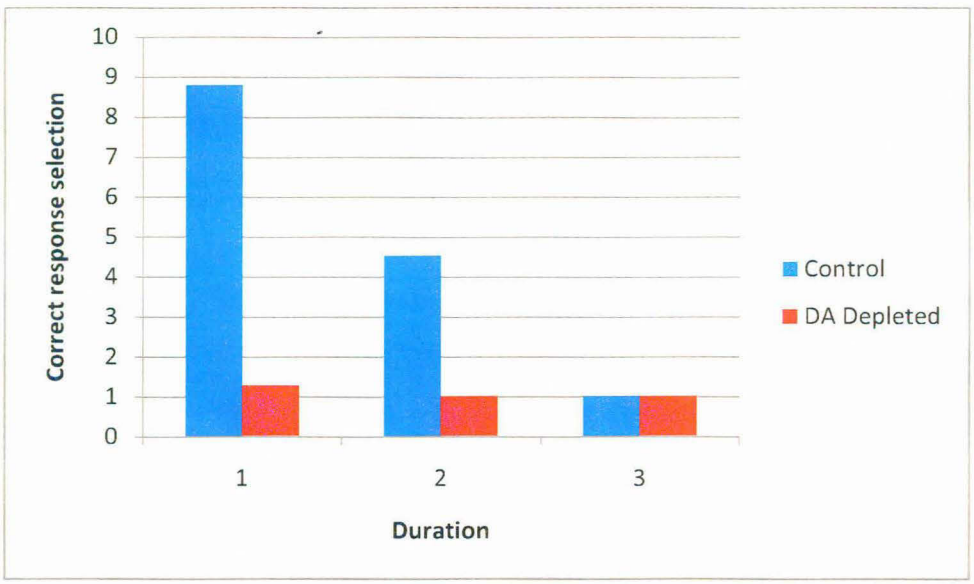


Figure 6.3: Neurocomputational model results. Number of correct response selection decreases as the hunger level goes down.

the number of correct response selection is much small in comparison to the control case for the initial phase of the simulation. After second and third phase, the number of correct response selection further decreases. The results for control and dopamine depleted case is compared in figure 6.3.

6.2. Effect of Dopamine Level Manipulation

6.2.1. Behavioral Experiment Results

Aberman et al (Aberman and Salamone, 1999) conducted experiments to study the relation between reward ratio and action rate. Reward ratios of 1, 4, 16 and 64 were used in their experiments, the number signifying the lever press count required for obtaining a food pellet. They found that the control rats showed an increase in the total number of lever presses with an increase in the reward ratio. They conducted the same experiments with dopamine depleted rats. In case of dopamine depleted rats, it was seen that for a reward ratio of 1, the number of lever presses was almost the same as in case of the control rats. The number of lever presses for a ratio of 4 was greater than that for 1, but it were significantly lesser than that for the control rats. Same was true for a ratio of 16. Interestingly, the number of lever presses for a ratio of 64 showed a significant drop - the number of presses was lesser than that for 16. Refer to figure 6.4.

Figure 6.4: Experimental results. In control rats, the number of lever presses increased as the reward ratio was increased. In dopamine depleted rats, the lever presses were significantly lower than the control rats.

6.2.2. Mathematical Model Results

Our mathematical model is able to explain the above results. First, for control rats, lower total lever presses are seen for lower ratios because the rats become slower once their hunger is satisfied. This happens quickly for low reward ratios due to more frequent food delivery. Dopamine depleted rat start the session by pressing the levers at a lower rate than the control rats.

For a reward ratio of 1, frequent reward delivery helps in a quick building up of the dopamine level, resulting in a gain in the lever press rate. The result is shown in figure 6.5. As a result, over the duration of the experimental session, they are able to

Figure 6.5: Simulation results. Three different reward ratios were used and the results show a qualitative match with the behavioral data.

press the lever enough number of times to be comparable to the control rats. For high reward ratios, however, the food delivery is less frequent. This, in addition with the fact that the rats are inherently slow due to the depletion makes it difficult for them to pick up enough speed. This becomes more and more prominent as the ratio becomes bigger.

The lever presses for higher reward ratio decreases the action rate because according to equation no. 4.3, the increase in EDA is less in comparison to the low reward ratio (most of the time, reward is 0), hence it also affect the equation no. 4.1 which gives less action frequency and the value action frequency decreases as shown in the figure no. 6.5.

6.3. Effect of Reward Rate Manipulation

6.3.1. Behavioral Experiment Results

Figure 6.6: Behavioral results show an increasing number of lever presses due to less reward ratio. Pre-fed rats require less number of lever presses because they are already satiated by some extend.

In the model proposed by Niv et al. (Niv et al., 2005), the response rate is proportional to the average reward rate. For the pre-fed rats (i.e. sated rats), the number of lever presses are less in comparison to control rats (refer to figure 6.7). For higher reward ratios, the curve is declining because for higher reward rate, the rats become less responsive. They might be satiated or their actions would not be correct due to less rewarded for pressing correct lever.

6.3.2. Mathematical Model Results

Our results through mathematical model depict the same behavior as given by the Niv et al. (Niv et al., 2005). The data is collected by executing the C++ programs based on the mathematical equations given in section 4.2. For control case, the initial hunger value is 5.5 and for pre-fed case, the hunger value is 3. The results, as shown in figure 6.8, are much similar to the behavioral data given in figure 6.7. For higher reward rate the lever press count is high. This can be explained as the hunger level is less frequent satiated due to less frequent rewards. To minimize the hunger, the model show more number of lever presses to get more reward.

Figure 6.7: Mathematical model results show an increasing number of lever presses for low reward rate but declining for higher reward rate.

In case of pre-fed i.e. less initial hunger value, the model receive the required reward in less number of level presses. But for the higher reward rate, the model requires more lever presses to get more cumulative reward.

6.3.3. Neurocomputational Model Results

The results obtained from running the simulated neurocomputational model explained in section 5.8 give the similar results as the behavior results done by

Figure 6.8: Neurocomputational Simulation results. Six different reward ratios were used and the results show a qualitative match with the behavioral data.

Aberman et al (Aberman and Salamone, 1999) and also in our mathematical model results. As results shown in the figure 6.6, in control case, the model show a vigor response for reward ratio 1 and less responsive for the higher reward ratios. The numbers used in reward ratios indicate the number of lever presses required to get a food palate. For higher reward ratio (in our case 6 or higher) gives a horizontal curve which can be explained as the model cannot learn for higher reward rate because the reward is less frequent in higher case. If we consider a behavior analysis for an animal, the animal can also deviate from what it has learnt in the training sessions due to less number of rewards for higher number of lever presses. One more reason could be that there is a physical limit to press the lever in one experimental session; hence at one point of time, the number of lever presses is maximum that could be pressed in one experimental session and that would be irrelevant with the satiation level of the animal.

For compilation of data, first the model is run for sufficient number of steps to train the network so that it can be responsive for the reward. After that the simulation is run for the number of trials until the network does not give error due to satiation layer value. Satiation layer projects its increasing value after each epoch to the striatum layer units so that their units responsible for response 2 selection from 'Go' units and response 1 selection from 'No-Go' units get more weights.

When the sufficient number of errors at each epoch is received, we stop the simulation for one reward ratio. Again for the reward ratio 2, first we train the network and then we change the script so that the network gets one reward for alternate phasic burst in SNc layer. Now again we run the simulation for sufficient number of trials to get the errors due to increasing satiation layer weights. Same procedure is followed for reward ratio 3, 4, 5 and 6. For pre-fed experiments, we choose the higher initial value of satiation layer and than start the simulation.

In case of pre-fed, the model starts the lever presses less number of times in comparison to the control case because now the model has some satiation layer value at the start of the experimental session. For higher reward ratio, the model requires more number of lever presses to satiate in comparison to control case.

6.4. Relation between Reward Rate and Action Rate

6.4.1. Behavioral Experiment Results

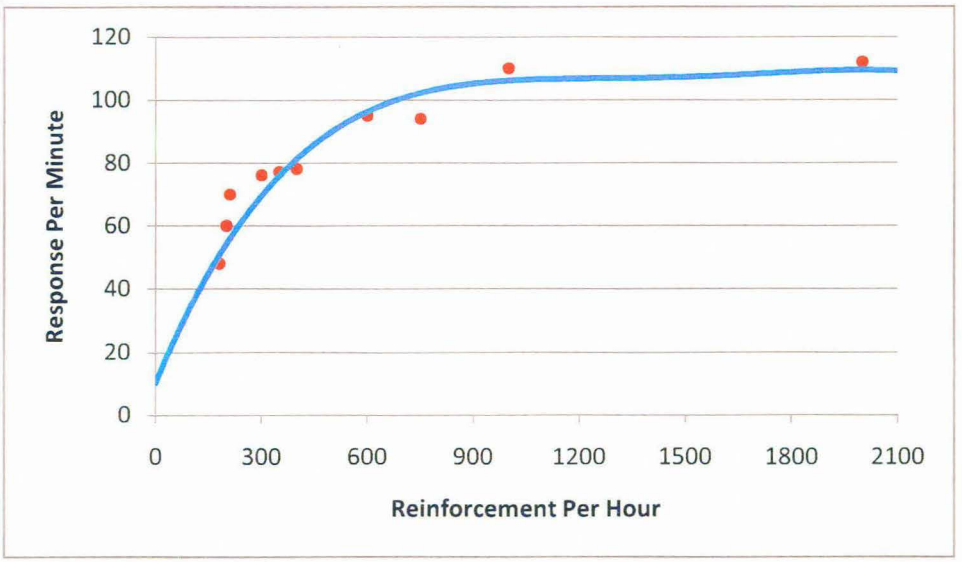


Figure 6.9: Behavioral experiment results showing an increasing trend of response rate for increasing the reinforcement rate.

The behavioral experiments done by the previous researchers on animals (Chung, 1966; Herrnstein, 1961; Niv et al. 2005), as in figure 6.7, show an increase in response rate by the increase in reinforcement rate. After some point of time, the increase in reinforcement rate doesn't affect the response rate as less frequent reinforcement doesn't make any learning and the animal presses the lever less frequent for reward.

6.4.2. Mathematical Model Results

According to our model, food consumption leads to an increase in the extra-cellular dopamine level, and the dopamine level, in turn, governs the action rate. A time step at which food is delivered, results in an increase in the dopamine level. In subsequent time steps, when food is not delivered, the dopamine level begins to fall back towards the baseline level. This continues until the next time step when the food is delivered once again. Hence, the running average of the dopamine level is determined by the time interval between successive rewards. Greater time interval leads to smaller average dopamine level. Since the action rate is dependent on the dopamine level, a greater rate is seen for a more frequent reward. For very high

reward rates, the dopamine level starts to asymptote near its maximum value, resulting in a ceiling effect on the action rate as well. Our model is very similar to the model proposed by Niv et al (Niv et al., 2005) in explaining this result. In their model, the action rate is governed by the average reward rate. The level of dopamine in our model is governed by the frequency of reward delivery, and hence, can be considered to be capturing the effects of average reward rate. The key point to note, however, is the reward rate at which the response rate asymptotes. From the data reported by Herrnstein (Herrnstein, 1970), it can be seen that in a typical rat, the response rate asymptotes for reward rates greater than or equal to 75 per hour. If we review the behavioral data reported in the previous section, the total number of food pellets received by the animals in the 45 minute long session were approximately 250, 250, 125 and 39 for the reward ratios of 1, 4, 16 and 64 respectively. This translates to a reward rate of 333, 333, 166 and 52 per hour. Hence, except for the reward ratio of 64, the response rate should reach its asymptotic value. However, the response rate cannot continue at this asymptotic value indefinitely. As described in the previous section, the response rate starts to drop as the animal's hunger is satisfied.

Figure 6.10: Mathematical model results showing an matching results as in the case of behavioral experimental results.

6.4.3. Neurocomputational Model Results

Figure 6.11: Neurocomputational model results showing a matching results as in the case of behavioral experimental and mathematical model results.

The neurocomputational model explained in section 5.8 is executed and data is collected, as shown in figure 6.10. The curve is depicting the similar results with the behavioral experiment results as well as mathematical results. The model uses 0.0 initial satiation layer value. Simulation is executed for reward ratio 1, 2, 3, 4, 5 and 6. As shown in the figure 6.10, for lower reward ratio, the number of correct response selection is less because the model get rewarded for each correct response selection and network show errors in initial trials and it learns quickly. For higher reward ratio, the number of correct response selection would not give the reward hence the network needs large number of correct response selection in which few of them leads to a reward and hence learning. Due to slow learning rate, network performs large number of correct response selection.

6.5. Effect of Reward Ratio Over Controlled and Over-Weight Animals

It is very often, before conducting the behavioral experiments on animals, the animals are kept on 80% of their body weight. The logic behind this is chronic hunger (section 4.3.1). If an animal is chronically hungry, it will show the vigorous response for dopamine-dependent actions. This also refers that if an animal is not chronically hunger i.e. its body weight is more than normal then the animal should not show the

vigorous response for any dopamine-dependent actions. We performed the experiment to prove this hypothesis using our set of equations and calculate the number of lever press counts in three different reward ratios. The result is shown in figure 6.11.

Figure 6.12: Simulation results. Three different reward ratios were used to show the relation between the reward ratio and lever presses in case of over-weight rat.

The simulation result is achieved by manipulating the ChronicHunger from 0.8 for controlled rats (depicting normal body weight) to 1.2 for over-weight rats (depicting the 20% extra body weight) in the equation 4.4. The simulation was performed for three successive time interval of 15 minutes. The result in figure 6.11 is somewhat proving that our hypothesis that the chronically over weighted animals have the same vigorous response as the case of controlled animals. In case of higher reward ratio, a normal animal presses more times lever for food than the over-weighted animals. This is a prediction made by our model and the proof remains by the means of behavioral experiments.

Chapter 7

Discussion

A good neurocomputational model can give a better understanding about brain processes. By manipulating various parameters, the complex brain processes can be explained more easily. Behavioral experiments use animals for collecting data for brain processes. Even after using sophisticated and costly instruments, sometimes it is too difficult to obtain the results from a live animal. Neurocomputational model helps a lot for collecting data and explaining the processes of the brain.

7.1. PDP++

I have chosen PDP++ software to create the model for dynamics of dopamine in reward pathway. Several software available for designing the neurocomputational model, but I found PDP++ is better than others with respect to reliability, easy to use, accepted by large research community, object oriented design approach and easy availability. The elements found in the brain are represented as object in this software. Like neurons can be expressed as units, a section of brain like thalamus can be constructed by a group of neuronal units. The connection between layers of neuron can be set by the available projection object.

Units' weight calculations can be set by the properties of the units and the properties of the projection. PDP++ provides the weight calculation by trial-and-error method or the Hebbian weight calculation. One can also set the combination of both types of weight calculation methods for a single unit or layer. Custom made weight calculation can be done using scripts. PDP++ provides the way to write scripts, which is similar to C++ language, by which the formula can be inserted so that they would execute during the simulation.

Overall, I think that PDP++ software is a complete software to create the connectionist models. One disadvantage regarding this software is the management of working display panels. For a large model, it is too difficult to manage the

windows for each and every object properties, processes and environments. The latest version of PDP++ i.e. Emergent is somewhat easy to manage the larger models because the objects and its properties are displayed hierarchically in separate window. So by selecting one object from one window, its properties are displayed in second window. PDP++ also has some feature for displaying the models in the form of presentations. It allows to create the snapshot of the models and to create animations of the cycles performed during the simulations.

7.2. Models and Simulations

Computational or mathematical models are helpful to understand the dynamic behavior of any system. In case of biological system, it is much more important because, it give a very convenient way to explore the biological processes. The model designing is very less time consuming, cheap, easy to collect data and accident free in comparison to the experiments in real biological environment. In our work we designed first a mathematical model and after that tried to develop the neurocomputational model to show the dynamic behavior of dopamine in reward pathway of the brain.

7.2.1. Mathematical Model

The action rate of an animal in a behavioral experiment is measured on the basis of lever presses for the food during a specified experimental session. Our mathematical equations calculate the amount of extracellular dopamine and update its arithmetic value when the model receives a reward. The received reward also satiate the hunger hence also lower the dopamine level. The chronic hunger is also considered in our mathematical model which is a long term processes. So the scale to calculate time to use in the equations for hunger and chronic hunger are different. The time scale for hunger equation is only measures a duration of an experimental session which is typically 30 – 45 minutes. In case of chronic hunger the time scale is of the order of days. Chronic hunger is a slow process and it would take from 5 – 10 days in case of rats. Our model considers all these factors for calculating the dopamine level for establishing the relation between hunger and action rate.

The results show that in normal animal, the action rate is higher than the animals having dopamine antagonist. Results also show that at initial stage, the action rate is higher in comparison to the subsequent time periods in an experimental session. Same behavior has also seen in case of dopamine depleted and pre-fed scenario. The results also depict the behavioral experiments done by researchers.

The initial values of the equation parameters are very crucial and they play very important role in simulation. The initial values of the equation parameters are calculated using some trial-and-error and hill climbing methods by considering some biological aspects and previous work done by researchers for all the equations.

7.2.2. Neurocomputational Model

As explained in mathematical model, the extracellular dopamine level is also depends upon the correct action performed and the receiving of reward. In the next move, the equation will calculate the dopamine level based on the previous correct or incorrect response. This particular task is modeled through context layer. The context layer stores the previous attempt result and gives the input at the next action attempt. The model is included in chapter number 4 along with the mathematical models.

To prove the mathematical models, we also have designed some neurocomputational models. These models are based on the neurocomputational model proposed by Frank (Frank MJ, 2005). This Model explains the two pathways for action signals received by Premotor cortex. The Frank's model is a proven one which exhibit the learning in case of response selection. Response selection is not only based on the input given by the sensory organs but it also includes the previous learned experiences. The model show how the pathways in basal ganglia works to select the response based on prior knowledge.

We added some elements in Frank's model and tried to explain the relation between the dopamine level and the action rate. The model depicts all the elements of basal ganglia and interconnection between them. Each connection has its own biological relevance. The addition of satiation layer made the model more close to the

proposed mathematical model. Since the reward will satiate the hunger, hence the dopamine level adjustment is done by the satiation layer.

7.3. Scope for Future Work

The work related to acquisition and extinction can be extended for the study of more dopamine dependent behavior such as effects of psychoactive substance use and drug withdrawal. The work related to dopamine dependent action rate can be extended to explain the dopamine dependent brain diseases like Parkinson's disease and Schizophrenia. The neurocomputational model can be extended to more complex models for explaining the dynamic effects dopamine on dopaminergic pathways.

The models can also be combined with the models on mesocortical pathways which is closely related with reward pathway i.e. mesolimbic pathway. The combined model can explore more details for reward related brain processes. These models can be easily transported to the latest version of PDP++ i.e. Emergent which can give more graphical presentations of layers and can be easily handled. It would happen that if one expand the model or combine the models the PDP++ would not be an easy platform to handle the view panels and understand the objects of the model. So converting the existing model to the new platform would be wise step before expanding the model.

References

Aberman JE and Salamone JD. (1999). Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. *Neuroscience*, 92(2):545-552.

Agid, Y., Ruberg, M., Hirsch, E., Raisman-Vozari, R., Vyas, S., Faucheux, B., Michel, P., Kastner, A., Blanchard, V., Damier, P., Villares, J., & Zhang, P. (1993). Are dopaminergic neurons selectively vulnerable to Parkinson's disease? *Advances in Neurology,* 60, 148–164.

Alexander, G. E., & Crutcher, M. D. (1990b). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neurosciences,* 13, 266–271.

Anderson, J. R. (1981). Interference: The relationship between response latency and response accuracy. *Journal of Experimental Psychology: Human Learning and Memory,* 7, 326-343.

Bassero, V. and Chiara, G. D. (1997). Differential influence of associative and nonassociative learning mechanisms on the responsiveness of prefrontal and accumbal dopamine transmission to food stimuli in rats fed ad libitum. *The Journal of Neuroscience*, 17(2):851-861.

Bear MF and Malenka RC. (1994). Synaptic plasticity: LTP and LTD. Current Opinion in *Neurobiology,* 4, 389–399.

Bello NT, Sweigart KL, lakoski JM, Norgren R and Hajnal A. (2003). Restricted feeding with scheduled sucrose access results in an upregulation of the rat dopamine transporter. *Americal Journal of Physiology - Regulatory, Integrative and Comparative Physiology,* 284:1260{1268.

Berridge, K.C. & Robinson, T.E. What is the role of dopamine in reward: Hedonics, learning, or incentive salience? *Brain Research Reviews,*28 (3), 308-367, 1998.

Berridge, KC. Motivation concepts in behavioral neuroscience.Physiology & Behavior 81 (2), 179 - 209, 2004.

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning and Memory,* 11:485-494.

Bozarth MA. (1994). Pleasure systems in the brain. In: Wartburton DM, editors. *Pleasure: The politics and the reality*. New York: Wiley & Sons.

Cagniard B, Balsam PD, Brunner D and Zhuang X. (2006). Mice with chronically elevated dopamine exhibit enhanced motivation, but not learning for food. *Neuropsychopharmacology*, 31:1362-1370.

Capaldi ED, Owens J, and Palmer KA. (1994). Effects of food deprivation on learning and expression of flavor preferences conditioned by saccharin or sucrose. *Animal Learning and Behavior,* 22 , 173-180.

Capaldi EJ. (1967). A sequential hypothesis of instrumental learning. In Psychology of learning and motivation, Vol. 1 (eds. K.W. Spence and J.T. Spence), pp. 67–156. Academic Press, New York.

Capaldi EJ. (1994). The sequential view: From rapidly fading stimulus traces to the organization of memory and the abstract concept of number. *Psychonomic Bull. Rev.* 1: 156–181.

Carr KD, Tsimberg Y, Berman Y, and Yamamoto N. (2003). Evidence of increased dopamine receptio signaling in food-restricted rats. *Neuroscience*,119:1157-1167.

Centonze D, Picconi B, Gubellini P, Bernardi G, and Calabresi P. (2001). Dopaminergic control of synaptic plasticity in the dorsal striatum. *European Journal of Neuroscience,* 13, 1071–1077.

Chevalier, G., & Deniau, J. M. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends in Neurosciences,* 13, 277–280.

Childress AR, Ehrman RN, McLellan AT, O'Brien CP (1988) Conditioned craving and arousal in cocaine addiction: a preliminary report. *NIDA Res Monogr* 81:74–80

Christie MJ, Summers RJ, Stephenson JA, Cook CJ, Beart PM. (1987). Excitatory amino acid projections to the nucleus accumbens septi in the rat: a retrograde transport study utilizing d[3H]aspartate and [3H]GABA. *Neuroscience* 22: 425–39.

Chronister RB, Sikes RW, Trow TW, De- France JF. (1981). The organization of the nucleus accumbens. In The Neurobiology of the Nucleus Accumbens, ed. RB Chronister, JF DeFrance, pp. 97–146. Brunswick, ME: Haer Institute

Chung SH. (1966). Some quantitative laws of operant behavior. Unpublished doctoral dissertation, Harvard University.

Cohen, J. D., Braver, T. S., & Brown, J. W. (2002). Computational perspectives on dopamine function in prefrontal cortex. *Current Opinion in Neurobiology,* 12, 223–229.

Colle, L.M. and Wise, R.A. (1988) Effects of nucleus accumbens amphetamine on lateral hypothalamic brain stimulation reward. *Brain Res.* 459, 361-368.

Cools, R., Barker, R., Sahakian, B., & Robbins, T. (2001). Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cerebral Cortex,* 11, 1136–1143.

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience,* 22, 4563–4567.

Datla K; Christidou MA; Widmer WW; Dexter DT. (2002). Citrus flavonoid tangeretin protects against dopaminergic neuronal loss in a rat model of Parkinson's Disease. AB*STR PAP AM CHEM S.* 224:U63-U63.

Daw ND. (2007). Dopamine: at the intersection of reward and action. *Nature Neuroscience*. 10: 1505 - 1507

Devenport, L. D. (1998). Spontaneous recovery without interference: Why remembering is adaptive. *Animal Learning and Behavior*, 26:172-181.

Di Chiara G. (1998). A motivational learning hypothesis of the role of dopamine in compulsive drug use. *Journal of Psychopharmacol* 12: 54-67.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature,* 380, 69.

Erenberg G. 1992. Treatment of Tourette syndrome with neuroleptic drugs. *Adv. Neurol.* 58:241–43.

Fibiger, H.C. & Phillips, A.G. (1979). Dopamine and the neural mechanisms of reinforcement. In A.S. Horn, B.H.C. Westerink, and J. Korf (eds.), The Neurobiology of Dopamine (pp. 597-615). New York: Academic Press.

Filion, M., & Tremblay, L. (1991). Abnormal spontaneous activity of globus pallidus neurons in monkey with MPTP-induced Parkinsonism. *Brain Research,* 547, 142–151.

Fishwick, Paul A. (1995):``Simulation Model Design and Execution: Building Digital Worlds," Prentice Hall

Foote, S., & Morrison, J. (1987). Extrathalamic modulation of cortex. *Annual Review of Neuroscience,* 10, 67–85.

Frank MJ., (2005). Dynamic Dopamine Modulation in the Basal Ganglia: A Neurocomputational Account of Cognitive Deficits in Medicated and Nonmedicated Parkinsonism; *Journal of Cognitive Neuroscience* 17:1, pp. 51–72

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between the frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience,* 1, 137–160.

French, R. M. (2003) Catastrophic Forgetting in Connectionist Networks. In Nadel, L. (Ed.) *Encyclopedia of Cognitive Science.* Vol. 1, pp. 431 - 435.

Goldberg SR, (1976). Stimuli associated with drug injections as events that control behavior. *Phamacol Rev.* 27: 325 – 340.

Gotham, A., Brown, R., & Marsden, C. (1988). 'Frontal' cognitive function in patients with Parkinson's disease "on" and "off " levodopa. *Brain,* 111, 299–321.

Grace AA, Amiel RJ (2002). Regulation of conditioned responses of basolateral amygdala neurons. *Physiology & behavior.* 77(4-5): 489-93.

Grace AA. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia. *Neuroscience,* 41(1):1-24.

Gratton, A., Hoffer, B.J. and Gerhardt, G.A. (1988) Effects of electrical stimulation of brain reward sites on release of dopamine in rat: an in vivo electrochemical study. *Brain Res. Bull.* 21: 319-324.

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

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

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, 11, 23-63.

Grossberg, S. (1998). Birth of a learning law. INNS/ENNS/JNNS Newsletter, 21:1–4. appearing with Neural Networks, 11(1).Hebb DO (1949). *The Organization of Behavior.* Wiley, New York.

Groves PM. (1983). A theory of the functional organization of the neostriatum and the neostriatal control of voluntary movement. *Brain Res. Rev.* 5:109–32.

Gupta, A. and Noelle, D.C. (2005a). Neurocomputational mechanism for generalization during the sequential earning of multiple tasks. In Proceedings of *the 2nd Indian International Conference on Artificial Intelligence.*

Gupta, A. and Noelle, D.C. (2005b). The Role of Neurocomputational Principles in Skill Savings, In *Proceedings of 27th Annual Conference of the Cognitive Science Society*, pp: 863-868.

Gupta, A. and Noelle, D.C. (2006). Lateral Inhibition Explains Savings in Conditioning and Extinction, IN *28th Annual Conference of the Cognitive Science Society.*

Gupta, A., Noelle, D.C. (2007). A Dual-Pathway Neural Network Model of Control Relinquishment in Motor Skill Learning. *IJCAI 2007*, pp: 405-410.

Haselgrove M, Pearce JM. (2003). Facilitation of extinction by an increase or a decrease in trial duration. *J Exp Psychol Anim Behav Process.* 29(2): 153-66.

Hauser R, Horner H, Männer R and Makhaniok M (1991). Architectural considerations for NERV — a general purpose neural network simulation system. *Lecture notes in Computer Science.* Heidelberg. 565: 183–195.

Hebb DO. (1949). The Organization of Behavior: A Neuropsychological Theory. Edition 99. Published by Wiley.

Herrnstein, R J (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior,* 4: 267-272.

Herrnstein, R. J. (1970). On the law of effect. *Journal of Experimental Analysis of Behavior*, 13(2):243-266.

Hikosaka, O. (1989). Role of basal ganglia in initiation of voluntary movements. In M. A. Arbib & S. Amari (Eds.), Dynamic interactions in neural networks: Models and data (pp. 153–167). Berlin: Springer-Verlag.

Hikosaka, O. (1998). Neural systems for control of voluntary action—A hypothesis. *Advances in Biophysics*, 35, 81–102.

Hochreiter S and Schmidhuber J (1997). Long short-term memory. *Neural Computation,* 9:1735–1780.

Hollerman, J., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience,* 1, 304–309.

Hopfield JJ (1982). Neural networks and physical systems with emergent collective computational abilities. In Proceedings of the *National Academy of Sciences*, 79: 2554–2558.

Igel, C. (2003) Neuroevolution for reinforcement learning using evolution strategies. In: *Sarker R, et al (eds) Congress on evolutionary computation, IEEE Press, New York*, Vol. 4, pp:2588 – 2595.

Jackson, G., Jackson, S., Harrison, J., Henderson, L., & Kennard, C. (1995). Serial reaction time learning and Parkinson's disease: Evidence for a procedural learning deficit. *Neuropsychologia,* 33, 577–593.

Jellinger, K. (2002). Recent developments in the pathology of Parkinson's disease. *Journal of Neural Transmission Supplement,* 62, 347–376.

Kish, S., Shannak, K., & Hornykiewicz, O. (1988). Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. *New England Journal of Medicine,* 318, 876–880.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science,* 273, 1399.

Kohonen T (2001). Self-Organzing Maps, volume 30 of *Springer Series in Information Sciences.* Springer, New York, third edition.

Koob GF, Le Moal M (2001): Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology* 24:97–129.

Koob GF, Nestler EJ. 1997. The neurobiology of drug addiction. *J. Neuropsychiatr. Clin. Neurosci.* 9:482–97.

Maddox, W., & Filoteo, J. (2001). Striatal contributions to category learning: Quantitative modeling of simple linear and complex non-linear rule learning in patients with Parkinson's Disease. *Journal of the International Neuropsychological Society,* 7, 710–727.

McClelland JL and Rumelhart DE (1985), Distributed Memory and the Representation of General and Specific Information, *Journal of Experimental Psychology: General*, Vol 114 No.2: 159 – 188.

McCloskey M. and Cohen N.J. (1989), Catastrophic interference in connectionist networks: The sequential learning problem, in G. H. Bower (Ed.) *The Psychology of Learning and Motivation*, Vol. 24, New York Academic Press, pp. 109 – 165.

McCullagh P and Nelder JA (1989). Generalized Linear Models. Second Edition. London: Chapman and Hall, pp. 511.

McGeorge AJ, Faull RLM. 1989. The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience* 29:503–37

Mink, J. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, 50, 381–425.

Missale C, Nash SR, Robinson SW, Jaber M, Caron MG. (1998). Dopamine receptors: from structure to function. *Physiol. Rev.* 78:189–225.

Montague PR, Dayan P, Sejnowski TJ (1996). A Framework for Mesencephalic Dopamine Systems Based on Predictive Hebbian Learning, *Journal of Neuroscience,* 16(5), 1936-1947.

Nemeroff CB, Bissette G. 1988. Neuropeptides, dopamine and schizophrenia. *Ann NY Acad. Sci.* 537:273–91.

Nestler EJ, Malenka RC. (2004). The addicted brain. *Sci Am*; 290: 78-85.

Nestler EJ. (2001). Molecular basis of long-term plasticity underlying addiction. *Nat Rev Neurosci*; 2:119–28.

Nieoullon, A. (2002). Dopamine and the regulation of cognition and attention. *Progress in Neurobiology,* 67, 53–83.

Niv, Y., Daw, N., and Dayan, P. (2005). How fast to work: Response vigor, motivation and tonic dopamine. In *Proceedings of Neural Information Processing Systems.*

Noelle D.C. (2008), Teaching Cognitive Modeling Using PDP++, *Brains, Mind & Media.*

O'Brien CP, Childress AR, McLellan AT, Ehrman R. (1992). Classical Conditioning in Drug-Dependent Humans. *The neurobiology of drug and alcohol addiction.* 654: 400-415.

O'Donnell P, Grace AA. (1995). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: hippocampal gating of prefrontal cortical input. *Journal of Neuroscience.* 15:3622–39.

O'Reilly RC, (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation,* 8(5): 895-938.

O'Reilly RC, (2001). Generalization in interactive networks: The benefits of inhibitory competition and Hebbian learning. *Neural Computing,* 13: 1199-1242.

O'Reilly, R.C. and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain.* MIT Press.

Olds, M.E. (1990) Enhanced dopamine receptor activation in accumbens and frontal cortex has opposite effects on medial forebrain bundle self-stimulation. *Neuroscience* 35, 313-325.

Olshausen B.A. and Field D.J. (1996). Emergence of simple cell receptive field properties by learning a sparse code for natural images. *Nature*, 381,607.

Packard, M., & Knowlton, B. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience,* 25, 563–593.

Parent A. (1990). Extrinsic connections of the basal ganglia. *Trends Neurosci.* 13:254–58.

Partiot, A., Verin, M., & Dubois, B. (1996). Delayed response tasks in basal ganglia lesions in man: Further evidence for a striato-frontal cooperation in behavioural adaptation. *Neuropsychologia,* 34, 709.

Pavlov IP. (1927). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex; Published by Oxford University press.

Pennartz CMA, Groenewegen HJ, Lopes da Silva FH. (1994). The nucleus accumbens as a complex of functionally distinct neuronal ensembles: an integration of behavioral, electrophysiological and anatomical data. *Prog. Neurobiol.* 42:719–61.

Percheron, G., & Filion, M. (1991). Parallel processing in the basal ganglia: Up to a point. *Trends in Neurosciences,* 14, 55–56.

Pothos, E. N., Creese, I., and Hoebel, B. G. (1995). Restricted eating with weight loss selectively decreases extracellular dopamine in nucleus accumbens and alters dopamine response to amphetamine, morphine and food intake. *The Journal of Neuroscience,* 15(10):6640-6650.

Reber, P. J., & Squire, L. R. (1999). Intact learning of artificial grammars and intact category learning by patients with Parkinson's disease. *Behavioral Neuroscience,* 113, 235.

Rescorla, R. A. (2005). Spontaneous recovery of excitation but not inhibition. *Journal of Experimental Psychology: Animal Behavior Processes,* 31(3):277-288.

Rescorla, R.A. (2001). Retaining of extinguished pavlovian stimuli. *Journal of Experimental Psychology: Animal Behavior Processes,* 27(2):115-124.

Rescorla, R.A. (2002). Saving tests: Separating differences in rate of learning from difference in initial levels. *Journal of Experimental Psychology: Animal Behavior Processes,* 28(4):369-377.

Rescorla, R.A. (2003). More rapid association change with retraining than with initial training. *Journal of Experimental Psychology: Animal Behavior Processes,* 29(4):251-260.

Rescorla, R.A. (2004). Spontaneous Recovery, *Learning Memory,* Vol. 11, pp: 501–509.

Robinson TE, Berridge KC (1993), The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Rev.* 18(3): 247-291.

Robinson TE, Berridge KC. (2001). Incentive-sensitization and addiction. *Addiction* 96:103-14.

Rogers, R., Sahakian, B., Hodges, J., Polkey, C., Kennard, C., & Robbins, T. (1998). Dissociating executive mechanisms of task control following frontal damage and Parkinson's disease. *Brain,* 121, 815–842.

Rolls, E., Thorpe, S., Boytim, M., Szabo, I., & Perrett, D. (1984). Responses of striatal neurons in the behaving monkey: 3. Effects of iontophoretically applied dopamine on normal responsiveness. *Neuroscience,* 12, 1201–1212.

Rumelhart DE and Zipser D (1986). Feature discovery by competitive learning. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 1. MIT Press, chapter 5, pages 151–193.

Salamone JD and Correa M. (2002). Motivational views of reinforcement: Implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioral Brain Research*, 137:3-25.

Schultz W, Dayan P and Montahue PR (1997). A neural substrate of prediction and reward. *Science*, 275, 1593.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80, 1.

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews, Neuroscience*, 1:199-207.

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, 13, 900–913.

Sejnowski TJ and Churchland PS, (1989). Brain and cognition. In M.I. Posner(Ed.), Foundation of cognitive science (pp 301 – 356). Cambridge, MA: MIT Press.

Sesack SR, Pickel VM. (1990). In the rat nucleus accumbens, hippocampal and catecholaminergic terminals converge on spiny neurons and are in opposition to each other. *Brain Res.* 527:266–79.

Shinonaga Y, Takada M, Mizuno N. (1994). Topographic organization of collateral projections from the basolateral nucleus to both the prefrontal cortex and amygdaloid nucleus accumbens in the rat. *Neuroscience* 58:389–97.

Sidhu A. (1998). Coupling of D1 and D5 dopamine receptors to multiple G proteins: implications for understanding the diversity in receptor-G protein coupling. *Mol. Neurobiol.* 16:125–34.

Sokoloff P, Schwartz J-C. (1995). Novel dopamine receptors half a decade later. *Trends Pharmacol. Sci.* 16:270–75.

Stacy Braslau-Schneck, (2003). Animal Trainer's Introduction To Operant and Classical Conditioning, URL: *http://www.wagntrain.com/OC/#Learning.*

Stern, C., & Passingham, R. (1995). The nucleus accumbens in monkeys (Macaca fascicularis). *Experimental Brain Research*, 106, 239–247.

Stewart J., de Wit H., Eikelboom R. (1984). Role of unconditioned and conditioned drug effects in the self-administration of opiates and stimulants. *Psychological Review*, 91(2): 251-268.

Sutton RS and Barto AG (1990). Time-derivative models of Pavlovian reinforcement. In J.W. Moore & M. Gabriel (Eds.), *Learning and computational neuroscience* (pp 497-537). Cambridge, MA: MIT Press.

Sutton RS and Barto AG (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press.

Sutton RS and Whitehead SD (1993). Online learning with random representations. In *Proceedings of the Eleventh International Conference on Machine Learning*, pages 314-321. Morgan Kaufman, San Francisco, CA.

Sutton RS. (1988). Learning to predict by the method of temporal differences. *Machine Learning,* 3: 9–44.

Swainson, R., Rogers, R. D., Sahakian, B., Summers, B., Polkey, C., & Robbins, T. W. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: Possible adverse effects of dopaminergic medication. *Neuropsychologia,* 38, 596.

Swanson LW, Cowan WM. 1975. A note on the connections and development of the nucleus accumbens. *Brain Res.* 92:324–30.

Vallar, L and Meldolesi, J. (1989). Mechanisms of signal trans- duction at the dopamine D, receptor,*Trends Pharmacol. Sci.,* 10:74-77.

Wagner, A. R., & Brandon, S. E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). IN S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 149-190).

Wagner, A. R., & Brandon, S. E. (2001). A componential theory of Pavlovian conditioning. In R.R. Mower & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 23–64). Mahwah, NJ: Lawrence Erlbaum Associates,Inc.

Wickens, J. (1997). Basal ganglia: Structure and computations. *Network: Computation in Neural Systems,* 8, R77–R109.

Wikler A. (1973); Dynamics of Drug Dependence: Implications of a Conditioning Theory for Research and Treatment, Archives of *General Psychiatry,* Vol 28, no.5: 611 – 616

Wilson, C., & Kawaguchi, Y. (1996). The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *Journal of Neuroscience,* 16, 2397–2410.

Wise RA. (1998). Drug-activation of brain reward pathways. *Drug Alcohol Dependence* 51:13-22.

Wise, R.A. (1978). Catecholamine theories of reward: A critical review. *Brain Research* 152: 215-247.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, 55:235-269.

Young, A. M. J., Joseph, M. H., and Gray, J. A. (1992). Increased dopamine release in vivo in nucleus accumbens and caudate nucleus of he ratt during drinking: A microdialysis study. *Neuroscience*, 48(4):871-876.