ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

LEARNING HOW TO SELECT AN ACTION: FROM BIFURCATION THEORY TO THE BRAIN INSPIRED COMPUTATIONAL MODEL

M.Sc. THESIS

Berat DENIZDURDURAN

Electronics and Communication Engineering Department

Electronics Engineering Programme

MAY 2012

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

LEARNING HOW TO SELECT AN ACTION: FROM BIFURCATION THEORY TO THE BRAIN INSPIRED COMPUTATIONAL MODEL

M.Sc. THESIS

Berat DENIZDURDURAN
(504101203)

Electronics and Communication Engineering Department

Electronics Engineering Programme

Thesis Advisor: Assoc. Prof. Dr. Neslihan Serap ŞENGÖR

MAY 2012

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

NASIL EYLEM SEÇİLECEĞİNİ ÖĞRENME: DALLANMA TEORİSİNDEN BEYİN ESİNLENMELİ HESAPLAMALI MODELE

YÜKSEK LİSANS TEZİ

Berat DENİZDURDURAN
(504101203)

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Elektronik Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Neslihan Serap ŞENGÖR

MAYIS 2012

Berat DENIZDURDURAN, a **M.Sc.** student of **ITU Graduate School of Science, Engineering and Technology** with student ID **504101203**, successfully defended the thesis entitled **"LEARNING HOW TO SELECT AN ACTION: FROM BIFURCATION THEORY TO THE BRAIN INSPIRED COMPUTATIONAL MODEL"**, which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assoc. Prof. Dr. Neslihan Serap SENGOR**..........................
Istanbul Technical University

Jury Members : **Prof. Dr. Serdar ÖZOĞUZ** ....
Istanbul Technical University

Prof. Dr. Ertan YURDAKOŞ ....
Istanbul University

Date of Submission : 04 May 2012
Date of Defense : 29 May 2012

To my parents,

FOREWORD

First and foremost, I would like to thank to my supervisor of this project, Assoc. Prof. Dr. Neslihan Serap Sengor for the valuable guidance and advice. She couraged me greatly to work on computational neuroscience. I also would like to thank to members of Neuroscience Modelling and Researh Group, especially Assist. Prof. Dr. Ozkan Karabacak for supporting me during my M.Sc. studies.

The mobile robot Khepera II used for simulations and implementations belong to I.T.U. Artificial Intelligence and Robotics Laboratory. I would like to thank the laboratory staff and especially the coordinator Assist. Prof. Dr. Sanem Sarel Talay for their guidance and for their sharing knowledge.

Also, I would like to take this opportunity to thank to the Thursday Group in Faculty of Medicine at Istanbul University for offering us to participate their journal club. It gave us an opportunity to participate and learn about the networks in the brain, the fundamental concepts and syndromes in behavioral neurology.

Finally I would like to thank my collegue, Ömer Deniz Akyıldız, for his unstoppable jokes during the time I studied.

This work is partially supported by I.T.U. BAP project no: 55522, and TUBITAK project no: 111E264.

May 2012

Berat DENIZDURDURAN
(Electronics Engineer)

TABLE OF CONTENTS

Page

ABBREVIATIONS

AS	: Action Selection
BG	: Basal Ganglia
GDB	: Goal-Directed Behaviour
RL	: Reinforcement Learning
PD	: Parkinson's Disease
HD	: Huntington's Disease
DA	: Dopamine
TDE	: Temporal Difference Learning
Str	: Striatum
STN	: Subthalamic Nucleus
SNr	: Substantia Nigra pars Reticulata
SNc	: Substantia Nigra pars Compacta
GPi	: Globus Pallidus interna
GPe	: Globus Pallidus externa
Thl	: Thalamus
FB	: Fold Bifurcation

LIST OF TABLES

LIST OF FIGURES

Page

LEARNING HOW TO SELECT AN ACTION:
FROM BIFURCATION THEORY TO THE BRAIN INSPIRED COMPUTATIONAL MODEL

SUMMARY

The computational models of cognitive processes affirm our understanding of the ongoing mechanisms and robot models are a further step in computational neuroscience. The main point of this thesis is to show the potential use of robot models for tasks requiring high order processes like action selection and reinforcement learning.

Neurophysiological experimental results suggest that basal ganglia take part in selecting an action amongst different choices based on the saliencies of each possibility. There are computational models based on these experimental results for action selection. This work focuses on modification of action selection by dopamine release and a computational model capable of adapting it behaviour with parameter change is proposed. In this work, the aim is to investigate the effectiveness of the cortico-striato-thalamic model in a scenario based on rat's behavior, so behavior of a rat is simulated on the mobile robot Khepera II. The proposed model has the ability of selecting the appropriate actions under changing environmental conditions, thus it is suitable to implement learning to become familiar with a new environment.

The differences between the sensory systems of the mobile robots and the rat is resolved in order to mimic the behavior of a rat. The mobile robot is trained to learn to recognize the food and the place of the nest and it is capable of completing the task even though the conditions in the environment changes. In all of 30 trials, mobile robot recognizes the food approximately in 6 or 7 steps and also approximately at its 4^{th} trial the robot learns the place of the nest and deposits the food there.

The ultimate goal of this thesis is to investigate the high-order process, goal-directed behaviour, and to utilize the reinforcement learning to determine the choices and using a simpler model of cortico-striato-thalamic circuit for action selection. The contribution of this thesis is to focus on bifurcation analysis of the dynamical system proposed for goal-directed behaviour. Based on this bifurcation analysis, we investigated the updating of action selections during reinforcement learning, and explain how this updating effects the dynamic systems behaviour. So an explanation of Basal Ganglia circuit for action selection is given, and these results are implemented on a mobile robot to solve a foraging task.

NASIL EYLEM SEÇİLECEĞİNİ ÖĞRENME:
DALLANMA TEORİSİNDEN BEYİN ESİNLENMELİ
HESAPLAMALI MODELE

ÖZET

En ilkel canlılardan en gelismis primat kabul edilen insana kadar doğru zamanda doğru kararlar verebilme yeteneğini üstlenen beyin bölgesi temelde aynı yapılardan oluşmaktadır. İnsan beynini üstün kılan neokorteksin bu devreyi referans alarak evrimleştiği iddia edilmektedir. En temel hayati kararlardan duyguların da devreye girdiği karmaşık kararlarda bilim insanlarının işaret ettiği beyin alt yapısı basal ganglia çekirdekleridir. Yapılan çalışmalar göstermektedir ki birden fazla döngü ile iç içe geçmiş bu bölge beynin birçok bölgesine aksonlarla bağlanmışken birçok bölgeden de uyarılar almaktadır. Basal ganglia devresine ilişkin herbir döngüyü bağlantı aldığı bölgelere bakarak fonksiyonel olarak sınıflandırabilmek mümkün olduğu gibi bu döngülerin birbirleriyle de sıkı bir ilişki içinde olduğunu söylemek mümkündür. Karar verme üst başlığı altında ilgilenilen yaklaşım eylem seçimidir. Eylem seçimi birden fazla seçeneğin olduğu durumlarda ortam şartlarına da bakarak en doğru eyleme yönelmemiz olarak tanımlanabilir. Bilişsel süreçlere ilişkin geliştirilmiş hesaplamalı (matematiksel) modellerin amacı canlıların davranışlarında rol alan bu mekanizmaları açıklamaya çalışmaktır. Bu modeller özellikle son yıllarda insansı robot çalışmalarında da kullanılmaktadır. Bu çalışmada vurgulanmak istenen ise karar verme ve pekiştirmeli öğrenmeye ilişkin bilişsel süreçlerin hesaplamalı modellerinin robotik uygulamalarını gerçekleştirmektir. Bu çalışmada özellikle motor hareketlerden sorumlu olduğu düşünülen dorsal korteks-basal ganglia-talamus devresi ile ilgilenilmiş ve bu bölgenin fonksiyonel yapısı sistem seviyesinde geliştirilen bir matematiksel model ile açıklanmak istenmiştir. Geliştirilen model lineer olmayan dinamik sistemler disiplininde ele alınıp ele alınan devreye sistem seviyesinde yaklaşan bir modeldir.

Hesaplamalı sinirbilim son yıllarda birçok disiplinden bilim insanının ilgilenmeye başladığı bir disiplin haline gelmiştir. İnsan beyninden esinlenerek geliştirilen makine öğrenmesi algoritmalarından daha hızlı ve daha verimli yazılım teknikleri için "beyin gibi programlama" tekniklerine kadar özellikle mühendislerin ilgilendiği birçok konunun temelini oluşturmaya, bilim insanlarının meraklarının bu yöne çekilmesine sebep olmuştur. Günümüzün çözülememiş en büyük sorularından birinin beynin nasıl işlediği oluşu, bu soruna mühendislik bakış açısıyla da bakılmasını zorunlu hale getirmiştir. Beynin işleyişine ilişkin matematiksel modeller ve bu modellerin test edilmesi güçlü yazılım tekniklerine ve de matematiksel bakış açısına sahip olunmasını gerektirmektedir.

Ele alınan bu tezde, beynin işleyişine ilişkin en popüler sorulardan biri olan eylem seçimini nasıl veririz, bu bilişsel süreci kontrol eden beyne ilişkin devreyi modelleyebilir miyiz ve modelimizi nasıl test ederiz sorularıyla başlandı. Amaca yönelik davranışlar ve pekiştirmeli öğrenmeye dair bilişsel süreçlerin hesaplamalı

modelleri merkezi sinir sisteminin fonksiyonel birimleri ile sinir taşıyıcılarına dair bulgulara dayanmaktadır. Yapılan çalışmalar göstermiştir ki, bulunulan ortamdan alınan çeşitli uyaranlara bağlı olarak farklı seçenekler içerisinden yapılan seçimlerde basal ganglia çekirdeklerinin önemli bir rolü vardır.

Nörofizyolojik bulgular basal ganglia çekirdeklerinin eylem seçiminde görev almasının yanında özellikle ödüle dayalı öğrenme ile de sıkı bir ilişkisinin olduğunu işaret etmektedir. Bu bulgulara dayanarak, geliştirilen matematiksel model ile zamansal fark öğrenme algoritması birlikte ele alınmış ve dinamik sistemler yaklaşımında geliştirilmiş olan modelin dallanma analizleri yapılarak basal ganglia çekirdeklerinde öğrenmenin nasıl gerçekleştiğine ilişkin bir metot geliştirilmiştir. Model ve geliştirilen öğrenme metotu tıpkı nörofizyolojik bulgularında işaret ettiği gibi basal ganglia çekirdeklerinin en önemlisi olan striatumun ve striatuma etki eden dopamin hormonunun öğrenmedeki etkisini modelleyebilmektedir. Ödül ve doğru karar arasında kurulan ilişki dopamine karşı düşen parametre ile kontrol edilebilmektedir.

Modelin geliştirilmesi ve dallanma analizlerinin yapılmasından sonra hesaplamalı sinirbilim literatüründe özellikle son yıllarda önem kazanan robotik bir uygulama ile test edilmesi amaçlanmıştır. Geliştirilen matematiksel modelin robotu kural tabanlı bir kodlama ile değil de tamamen modelin kararlarına bakılarak yönlendirmesi amaçlanmıştır. Test ortamında en temel üç hayati eylem ele alınmıştır. Bilinmedik bir ortamda bir farenin hayatta kalabilmesi için gerekli olan bu üç eylem, ortam içerisinde yem bulmak için hareket edebilme, önüne çıkan engellerden kaçınma ya da yem olduğunu düşündüğü cisimleri alma, ve eğer yem bulduysa güvenilir bulduğu yuvasına bu yemi taşıma olarak sınıflandırılabilir. Robot testimizde robot için de bu üç eylem yeniden organize edilmiş ve tekerleklere ilişkin motorları çalıştırıp ortam içerisinde rastgele hareket etme, engel, yem ve taşıyamayacağı kadar ağır olan yem arasındaki farkları ayırdedebilme ve potansiyel yemini kaldırma ve yine en son olarak ışıkla işaretlenmiş yuvasına dönme olarak dizayn edilmiştir. Robot üzerinde kullanılan uzaklık ve ışık sensörleri farenin dış dünyayı tanıyabilmesi, duyumsayabilmesi için kullandığı koku alma ve görme duyularını sembolize etmeye, anlamlandırmaya çalışmaktadır. Yapılan 30 simülasyon göstermiştir ki, bu çalışmada kullanılan hesaplamalı model ve pekiştirmeli öğrenme algoritması robota yem olarak tanıtmaya çalıştığımız silindirleri, engellerin olduğu bir ortamda, yaklaşık 6. ya da 7. denemesinde tanımış sürecin devamında yani yemi öğrendikten sonra yuvasını ise 3. ya da 4. denemesinde öğrenebilmiştir. Beklenildiği üzere robot bu 30 testin tamamını öğrenmeyle neticelendirmemiş, değişen ortam şartlarına bağlı olarak öğrenememe ile de karşılaşmıştır.

Geliştirilmiş olan bu robotik test ortamı matematiksel model, bu modele ilişkin dallanma analizleri ve pekiştirmeli öğrenme aynı anda ele alınarak çözülmüştür. Test ortamı kurulurken görevlerin çok daha zorlu olmasındansa modelin ve öğrenme metotunun işlerliğinin görselleştirilmesi amaçlandığı için robota ilişkin görev basit tutulmuştur. Modelin ve öğrenme metodunun bu test ortamından başarı ile test edilmiş olması çok daha karmaşık testlerin de çözülebileceğine ilişkin bir ilk adım olarak düşünülmektedir.

Geliştirilen matematiksel model beyne ilişkin alt yapılara ve alt yapıların birbirleriyle olan ilişkisi gözönüne alınarak modellenmiştir. Matematiksel modellerin bilişsel bilimdeki en önemli özelliklerinden biri de modellere bakarak tam tersine ele alınan beyne ilişkin alt yapılar hakkında yorumlar yapabilmenin mümkün olabilmesidir. Ele alınmış olan bu matematiksel model daha önce de vurgulandığı gibi birçok özel fonksiyonları karşılayabilme yeteneğine sahiptir, tıpkı dopamin ile öğrenme arasında bir ilişkiye işaret edebilmesi gibi. Matematiksel modelin bir diğer özelliği de ele alınan beyne ilişkin basal ganglia çekirdeklerinin fonksiyonel olarak bozulmasında görülen davranışsal bozukluklara da işaret edebilmesidir. Dopamin azlığının günümüzün en önemli rahatsızlıklarından biri olan Parkinson ile ilişkilendirilmesi ve Parkinson hastalarında belirgin olarak görülen eylem seçimine ilişkin görülen kararsızlık matematiksel model tarafından da karşılanabilmektedir. Benzer şekilde dopamin fazlalığının sebep olduğunun bilindiği genetik bir rahatsızlık olan Huntington hastalığının da modeldeki dopamin kontrol edilerek işaret edilebildiği gösterilmiştir. Eylem seçimine ilişkin elde edilen sonuçlar tıpkı Huntington hastalığında olduğu gibi aynı anda birden fazla eylemin çok kolay bir şekilde seçilebilmesini gösterebilmektedir.

Bu çalışma ile, nöral yapılara dayanarak geliştirilen hesaplamalı modellerin robotik uygulamalarının, bilişsel süreçlerin anlaşılmasının yanı sıra makine öğrenmesi için de gerçekleştirilecek çalışmalara yeni bakış açısı kazandırabileceğine dair bir örnek teşkil etmesi amaçlanmıştır.

1. INTRODUCTION

The computational models of cognitive processes based on neural substructures affirm our understanding of the ongoing mechanisms during these high order processes. Thus, computational models of neural systems can be regarded as tools for understanding the cognition. Obtaining these models and showing their effectiveness would stimulate studies in cognitive science and inspire developing further tools for machine learning. These models also inspire new approaches and techniques for implementing intelligent systems. The comprehensive studies in computational neuroscience provides an overview on the research conducted with brain-inspired intelligent systems and range from building Very-large-scale integrated circuit (VLSI), to the cognitive robotic's application further to the medical robotics. All became important topics in contemporary neuroscience. The aim of these researches can be stated as creating a new approach in understanding the human brain studies. Eventually, brain-inspired systems will become a tool to the next generation technology once its secret is revealed and its potential is understood.

One of the interesting questions of contemporary neuroscience is how primates make appropriate decisions at the right time, literally defined as Action Selection (AS). Informally speaking, AS is the challenge of the next step to create the best choices considering the list of available actions. As we are interested in AS from the perspective of computational neuroscience, we have to answer another question to understand AS. Which part of brain or what kind of network has a role in selection of the approriate action for motor control? Neurophysiological experimental results suggest that basal ganglia (BG) plays crucial role in AS while dopamine (DA) modifies this process [1]. There are computational models based on these experimental results for AS [1–7].

AS is a part of high order cognitive process which is called Goad-Directed Behaviour (GDB). Amongst wide spectrum of high order cognitive processes such as planning, selective attention, decision making; GDB has driven a specific attention and numerous computational models have been proposed for AS [2, 3, 8] and reinforcement learning (RL) [1, 4], which together constitute GDB. The neural substructures taking part in GDB are in general orbito and medial prefrontal cortex, dorsolateral prefrontal cortex, BG and related regions of thalamus, so the computational models consider the interaction of these neural substructures [1–9]. Not only behavioral disorders such as addiction and neurodegenerative diseases as Parkinson's Disease (PD), Huntington's Disease (HD) are modeled based on the models of BG [10, 11] but also different cognitive processes as feature detection are modeled [6].

Since the information processes in brain are dynamic, the proposed model is constituted with discrete-time dynamical systems [12]. The brief notions for dynamical systems and bifurcation theory will be given in Chapter 2. As the role of BG circuits in AS is well-known, motivated by these studies, a computational model is proposed for AS which includes both direct, indirect and hyperdirect pathway. The proposed model for AS will be discussed in Chapter 3. This work focuses on modification of AS by dopamine (DA) release and the proposed computational model is capable of adaptation with parameter change. The adaptation ability of the proposed model will be exploited with bifurcation analysis and the modification of the parameters will be carried out by RL.

The dynamical system considered for AS in the model is analyzed using with XPPAUT to discuss the effect of DA on AS. These analyses reveal that the DA effect on the selection processes. Based on the results of this bifurcation analysis, the role of RL on AS is discussed. In chapter 4, we will focus on the learning architecture for BG circuits by means of temporal difference learning (TDL) and bifurcation theory. After the whole model's architecture is set up and analysis is carried out, in Chapter 5, we will give our results where we tested our model with a robot application. The model is implemented on mobile robot and foraging task is realized where an exploration in an unfamiliar environment with training in the world is accomplished. Here, an implementation of GDB on Khepera II mobile robot will be presented. The main point of this work is to show the potential use of robot models for tasks requiring high order processes like GDB. Thus, this work fulfills its aim of showing the efficiency of brain-inspired computational models in controlling intelligent agents. The model and the approach of analyzing it could also be informative in understanding the effect of DA in physiological disorders as PD and HD.

2. DYNAMICAL SYSTEMS

Computational neuroscience studies, involving different point of views, have provided key insights into brain structures. As the information processing in primate brain points out the formation of brain function is dynamic [12], so we proposed a computational model based on nonlinear dynamical systems theory. In this chapter we will introduce brief notions for dynamical systems [12–15] which are required to follow the work carried out in this thesis.

2.1 Definition of Dynamical Systems

Dynamical system is defined with two sets, one is set of states, the other is set for time, and the evolution of its states are described to be a deterministic process [12–15]. In dynamical system, the next state variable depends on time and its previous value. In brief, there exists a law that defines the system's evolution from the past states to the future states which depends on time. There are a number of events that can be defined as dynamical systems, such as: economical, social, physical and even neurological processes [14]. To give an example for dynamical systems considered in brain studies, the well-known neuron model, Hodgkin-Huxley with a modified version for Medium Spinny neurons in Striatum will be considered in this thesis, see Appendix A.1. This system has four-dimensional ordinary differential equations where one of its states describes the membrane potential of neuron [16]. The other equations define the dynamics of K^+ persistent and Na^+ transient ion channels probability of being open and closed. All four state variable simulates the real neuron behaviour. The computational model of a neuron, simulating its behaviour, defines the dynamical processes which depend on initial states. The state variables of dynamical systems, in this case neurons, are evaluated with a rule of the continuation to converge to one of the possible state spaces. Even the analysis of the nonlinear dynamical system is difficult, dynamical system theory allows us to describe the complex structures in real life. As we mentioned above, since the real-life is a dynamical system, the comprehensive studies to explain the natural processes is typically constituted with nonlinear dynamical systems.

Although dynamical systems have key role in the development of pioneered technologies, chaos has became prominent in the beginning of the latter half of 20^{th} century [14]. In 1960s, with the progress of the computer technology, Edward Lorenz discovered the strange attractor when he studied the convection rolls in atmosphere. This interesting experiments allow us to understand the chaotic notions in mathematics [17]. The solution of the system he developed never converges to any equilibrium or limit cycles even the iteration step is increased too much, thereby even the system is deterministic long-term prediction in these systems is imposibble [18]. For example a simple dynamical system named Lorenz System has a *strange* solution and when this solution is plotted on, it is called Lorenz Attractor. Behaviour of this system depends on initial conditions. Solution of the dynamical system giving Lorenz Attractor is given in Figure 2.1. As it can be followed in Figure 2.1, solution of the system

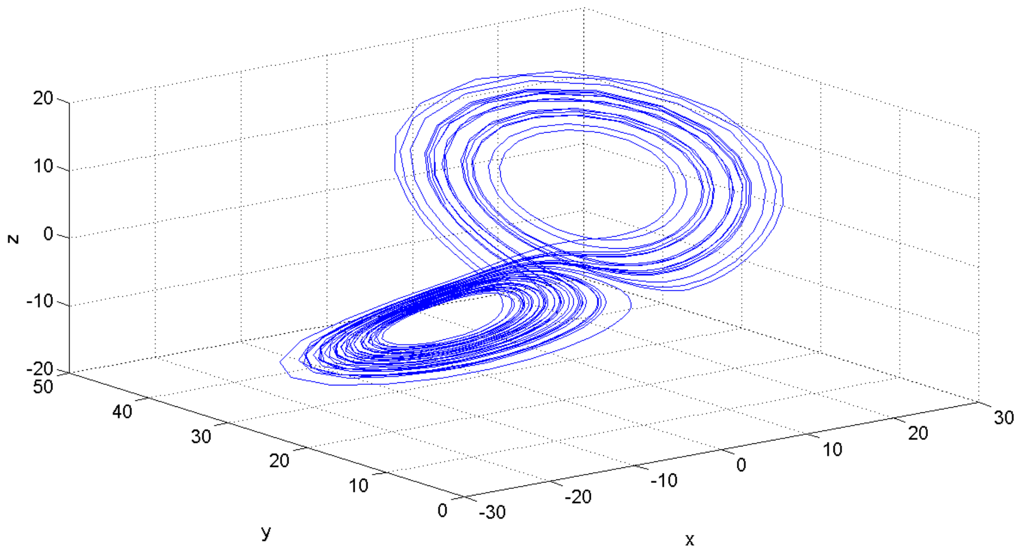

Figure 2.1: Lorenz attractor, $\rho = 28,\ \sigma = 10,\ \beta = 8/3$.

depends on initial conditions which make the system more sophisticated. Information processing in brain is another complicated example for dynamical system. As we deal with computational neuroscience, the further details of the dynamical systems will be given via modified Hodgkin-Huxley neuron models [16]. The equations for modified Hogdkin-Huxley can be found in Appendix A.1. Now we will shortly give definition of some quantities, following from [13]. The figures that we will give below are simulated in XPPAUT [19] and in house built in MATLAB codes.

Definition 2.1 State Space: The state space, X, is a set of possible solutions of the given system [13] ■

An example of a state in state space can be seen in Figure 2.1. These possible solutions describe the current position of a given system and with state space we can see the evolution of to solution of the system.

Definition 2.2 Time: The transition between the past states to the future states can be defined as an evolution of the given systems. These evolutions depend on time $t \in T$, where T is a number set [13]■

We can consider the dynamical systems where the time can be continuous (real) $T = \Re^1$ or discrete (integer) $T = \mathbb{Z}$ and we can define these systems such as continuous-time and dicrete-time dynamical systems, respectively. The model that we deal with (see Section 3) will be defined as a discrete-time dynamical system on the other hand Hodgkin-Huxley neuron model is a continuous-time dynamical system as you can see the phase portrait in Figure 2.2. Discrete-time dynamical systems are much more easy to solve rather than the continuous-time dynamical systems.

Definition 2.3 Evolution Operator: In a dynamical system, to determine the next state of the system at time t, x_t, system evaluates it using the initial state, x_o, and state transition law, φ which is called evolution operator (in continuous-time case it can be called *flow*, $\varphi_{t \in T}$) ■

$$\varphi^t : X \to X \qquad \textbf{(2.1)}$$

where system starts an initial state, x_o, and transforms to the next state, x_t, at time t:

$$x_t = \varphi^t x_o \qquad \textbf{(2.2)}$$

As the dynamical system is a deterministic process, there are two natural properties for evolution operator:

$$\varphi^0 = id \qquad \textbf{(2.3)}$$

where id can be defined as identity map on X, id $x = x$ for all $x \in X$. This property of the evolution operator means that the system's state does not change "spontaneously" and second property is:

$$\varphi^{t+s} = \varphi^t \circ \varphi^s \qquad \textbf{(2.4)}$$

It can be also written as:

$$\varphi^{t+s}x = \varphi^t(\varphi^s x) \qquad \textbf{(2.5)}$$

for all $x \in X$, and $t, s \in T$. According to this property, we can define the system as an "autonomous" which means that behaviour of the system does not change in time.

In discrete-time case, system can be defined only one map , $f = \varphi^1$, called "time-one map". We can generate the next state by using the "time-one map", such as:

$$\varphi^2 = \varphi^1 \circ \varphi^1 = f \circ f = f^2 \qquad \textbf{(2.6)}$$

In this case, the second iterated map, f^2 is computed by using "time-one map". Similarly:

$$\varphi^k = f^k \qquad \textbf{(2.7)}$$

for all $k > 0$.

After explaining the evolution operator, now we can define the dynamical systems, formally as:

Definition 2.4 A dynamical system is a set of time T, a set of state X, and a set of evolution operators $\varphi^t : X \to X$ with satisfying the Equation 2.3 and 2.4 [13] ■

2.1.1 Phase portrait

A formal definition of orbits (trajectories) can be given as follows:

Definition 2.5 $Or(x_0) = x \in X : x = \varphi^t x_0 \; for \; all \; t \in T \; such \; that \; \varphi^t x_0 \; is \; defined$ ■

Orbits of a continuous-time system are curves while in discrete-time systems are sequences of points. The composition of orbits constitutes the phase portrait. Let us consider the behaviour of the neuron to explain the phase portrait in Figure 2.2. The typical behaviour of the neuron is quescient or spiking, in special cases bursting. Here, we can see the bursting behaviour of the neuron while the given system's phase portrait seems like a torus [20]. This phase portrait is obtained with in a special case, see Appendix A.1. From the computational neuroscience point of view, the state of this

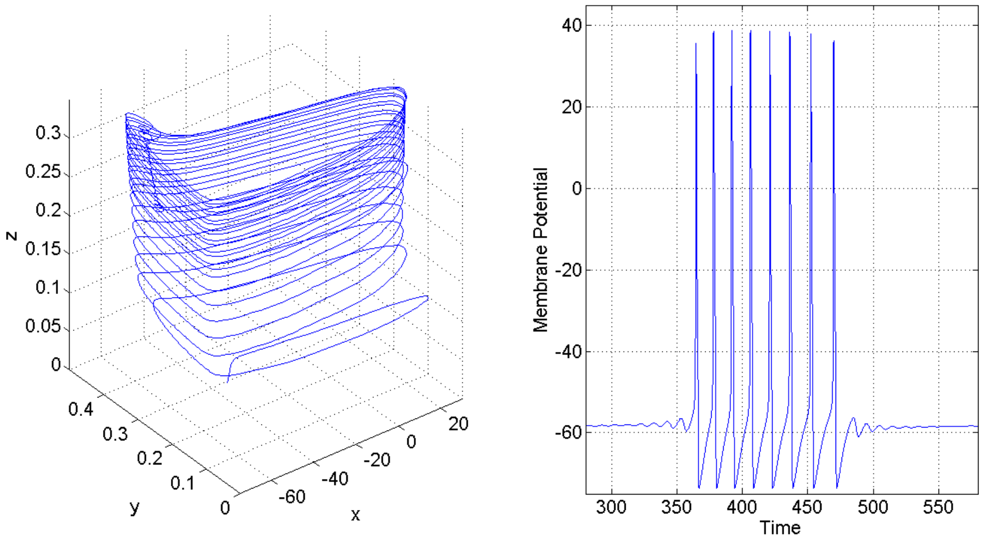

Figure 2.2: a. An example of phase portrait with Hodgkin-Huxley neuron model. b. Evolution of the solution by time.

behaviour is constituted with a resting and spiking modes between a transition period which can be called bursting [12].

2.1.2 Equilibrium (fixed point) and periodic orbit (limit cycle)

The classification of all possible orbits can be gives as: equilibrium (in discrete-time fixed points), periodic orbits (limit cycles) and strange attractors [14]. We can describe the equilibrium (fixed point) that phase space maps into itself either as time approaches infinity or as time approaches negative infinity and the formal definition as follows for continuous time:

Definition 2.6 $x' = f(x)$: x is an equilibrium if $f(x) = 0$ ■

while in discrete-time it is defined as follows:

Definition 2.7 $x_{n+1} = f(x_n)$: x is a fixed point if $x = f(x)$ [13] ■

An example of equilibrium point can be seen in Figure 2.3a. In two-dimensional manifold, a periodic orbits (in discrete-time limit cycles) can be defined as a closed trajectory when the other trajectories spirals into it in all circumstances (see Figure 2.3b). To give a formal definition for continuous-time case:

Definition 2.8 L_0 : for all $x_0 \in L_0$ is a periodic orbit if $\varphi^{t+T_0} x_0 = \varphi^t x_0$ for all $t \in T$ ■

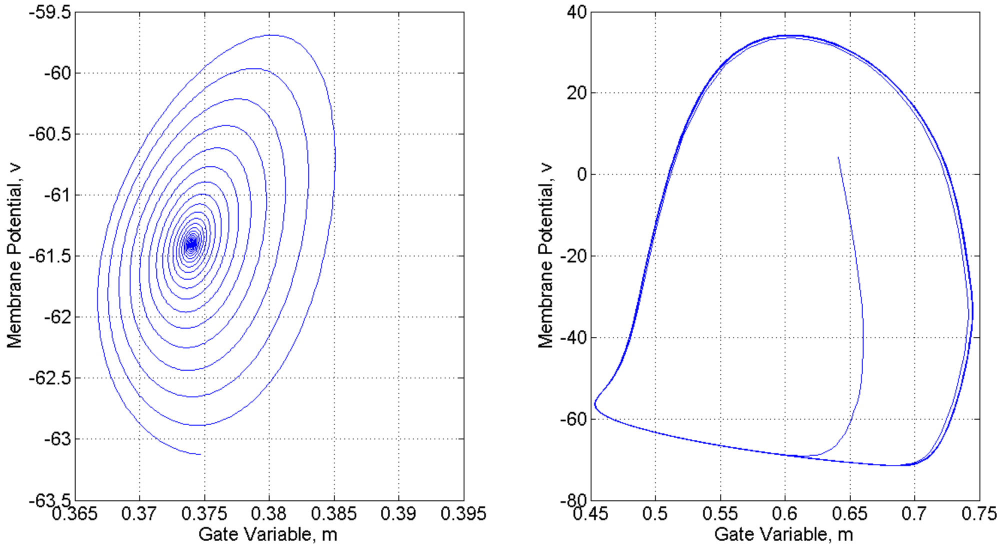

Figure 2.3: a. Phase portrait of an equilibrium point b. Phase portrait of a limit cycle.

while in discrete-time case:

Definition 2.9 $x_0, f(x_0), f^2(x_0), ..., f^{N_0}(x_0) = x_0$ where $f = \varphi^1$ and the period $T_0 = N_0$ ■

2.2 Bifurcation Theory

Bifurcation theory is an approach to investigate the nonlinear dynamical systems in order to explain different types of qualitative changes occuring in the given systems [21]. This strategy allows mathematicians to produce different structural behaviour of the systems by varying the significant parameters. The bifurcations can occur both continuous and discrete time (it is also called iterated maps) dynamical systems. We can consider the dynamical systems showing qualitative change depending on parameter changes in continuous time representation as:

$$\dot{x} = f(x, \alpha) \quad \textbf{(2.8)}$$

while in the discrete-time:

$$x \mapsto f(x, \alpha) \quad \textbf{(2.9)}$$

where the phase variables are denoted by $x \in \mathbb{R}^n$, and the parameters are denoted by $\alpha \in \mathbb{R}^m$ [13]. According to the bifurcation theory, the phase portrait of the system varies with the parameter variations. In this situation, there are two possibilities. One of them is topologically equivalent system to the original one and the other possibility is the change in the topology.

Definition 2.10 If the topological equivalence is broken with parameter change and phase portrait changes than it is called bifurcation [13] ■

The bifurcation point is described as a critical point where the systems topology change. A bifurcation analysis, where the continuous-time dynamical systems is considered, can be found in Appendix A.1.

Even there are different classifications of bifurcation, the most common approach to define the bifurcation classification is based on codimension concept described in geometric topology. On a manifold where codimension 1, then this corresponds to the dimension of topological disconnection by a submanifold [21]. Most of well-known bifurcation types are codimension 1 bifurcations such as Saddle-Node (a.k.a. Fold), Andronov-Hopf, Fold limit cycle, Neimark-Sacker (a.k.a torus) or Flip (a.k.a period doubling) bifurcations.

The numerical methods finding the bifurcation in a given dynamical systems require difficult continuation methods. To deal with issue, software packages (AUTO, MATCONT, XPPAUT, PyDSTool) are used which include implementation of algorithms for bifurcation. These algorithms consist of different bifurcation conditions, differential equation solvers (such as Runga-Kutta, Newton-Raphson Method) and numerical continuation methods. The bifurcation analysis of the given system starts with finding an equilibrium point or periodic orbit. Then, the algorithms in these software packages tries to follow these special attractors by varying significant parameters while using numerical continuation methods. Software programme detects the codimension 1 bifurcations during this branch of solutions [22].

In order to explain the bifurcation theory, we consider two different systems in continuous-time and discrete-time and only some specific bifurcations will be considered, seperately. We will introduce two bifurcations with conditions, one for continuous-time (with Medium Spinny neuron model) and one for discrete-time dynamical systems (with the computational model that we proposed).

2.2.1 Andoronov-Hopf bifurcation

We can explain the Andronov-Hopf bifurcation via the qualitative changes in ordinary differential equations. In Andronov-Hopf bifurcation, a stable equilibrium turns unstable while a limit cycle borns from this bifurcation point [13]. In bifurcation point, stable equilibrium loses its stability via occuring a pair of purely imaginary eigenvalues. There exists two different types of Andoronov-Hopf bifurcation. According to the stability of limit cycle which borns in critical point, the bifurcation is called with supercritical (stable limit cycle) or subcritical (unstable limit cycle) [13]. We handled with modified version of Hodgkin-Huxley neuron model to explain the bifurcation in continuous-time dynamical systems. In order to explain the supercritical Andronov-Hopf bifurcation, we focused on phase portraits in the plane. Subcritical Andronov-Hopf bifurcation can be found in [23]. As you can see in Figure 2.4, system has a stable equilibrium point (red dot) for bifurcation parameter $I_{app} < 5.6518$ (details in Appendix A.1) where the system always converge to this stable equilibrium point.

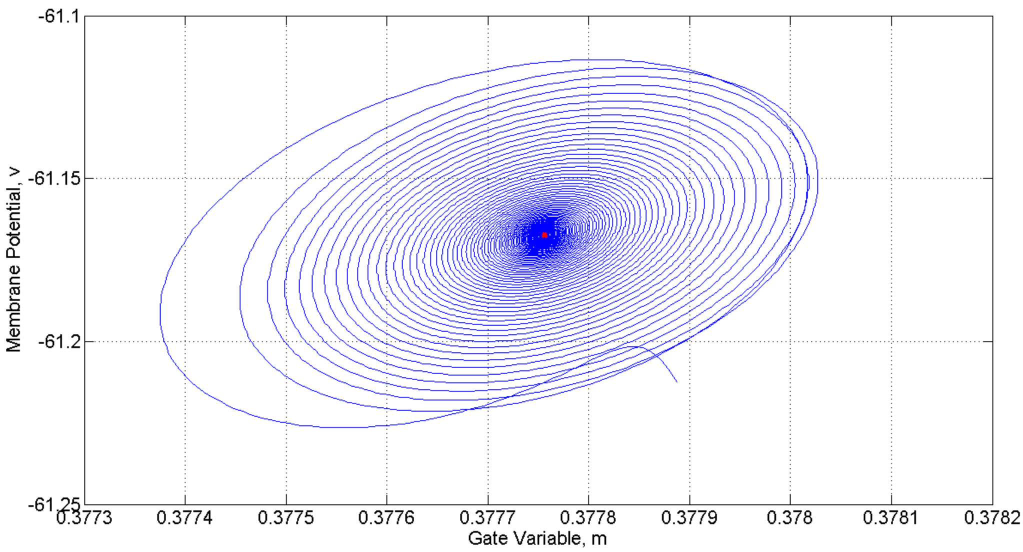

Figure 2.4: The phase portrait of the given system when bifurcation parameter is $I_{app} < 5.6518$. Red dot indicates the equilibrium point.

As you can see in Figure 2.5, when we increased the bifurcation parameter through the critical point, $I_{app} \doteqdot 5.6518$, the stable equilibrium point turns into unstable equilibrium point (which is located in the middle of red cycle) while a stable limit cycle (red cycle) borns which means the system's phase portrait shows the qualitative changes for supercritical Andronov-Hopf bifurcation.

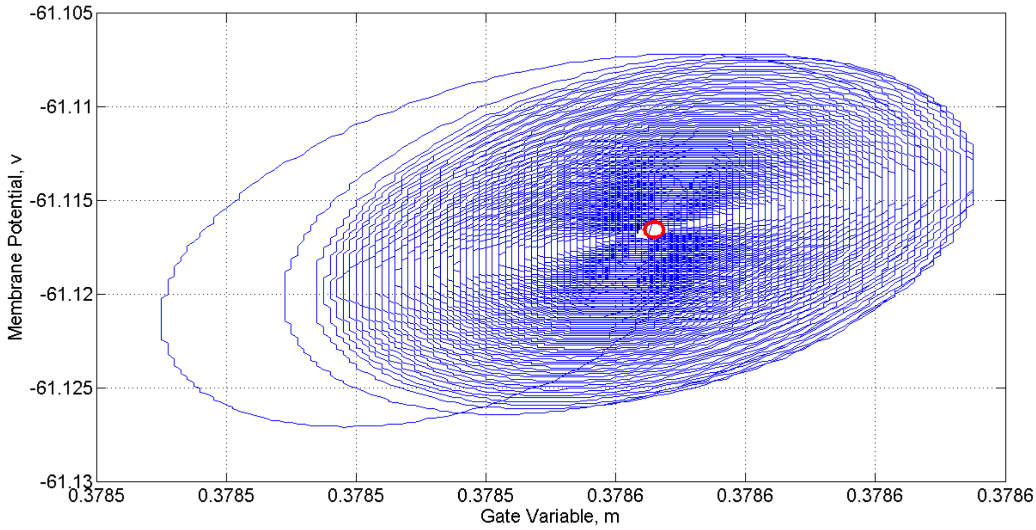

Figure 2.5: The phase portrait of the given system when bifurcation parameter reaches the critical point, $I_{app} \doteqdot 5.6518$. Red cycle indicates the stable limit cycle.

After this critical point, system's phase portrait can be seen in Figure 2.6. As it can be followed, system has one stable limit cycle and all trajectories converge to this limit cycle even it starts with in inside of limit cycle (red line) or outside of it (blue line).

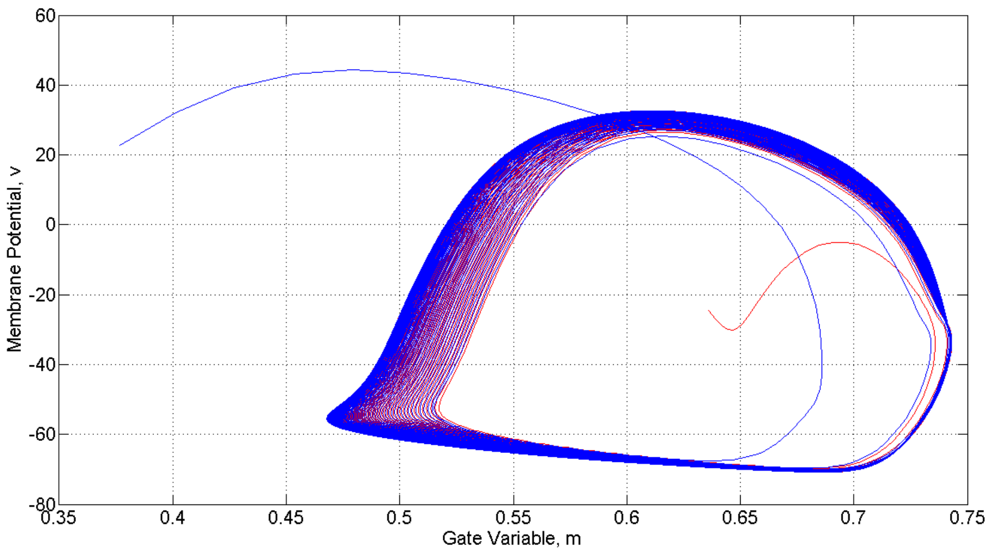

Figure 2.6: The phase portrait of the given system when bifurcation parameter passes the critical point, $I_{app} > 5.6518$.

There some conditions to be satistified for occurence of Andronov-Hopf bifurcation. Now we will explain these conditions.

In continuous-time dynamical system:

$$x^{'} = f(x, \alpha),\ x \in \Re^n \qquad \textbf{(2.10)}$$

where the parameter $\alpha \in \Re$ and f is a smooth function [13]. First, the system has to satisfy the following conditions

- for all $|\alpha|$ within its neigbourhood, there have to be a family of equilibria $x^0(\alpha)$
- Jakobian matrix $A(\alpha) = f_x(x^0(\alpha), \alpha)$

$$\lambda_{1,2}(\alpha) = \mu(\alpha) \pm i\omega(\alpha) \qquad \textbf{(2.11)}$$

This means that system has one pair of complex eigenvalues. If $\alpha = 0$ where $\mu(0) = 0$ and $\omega(0) = \omega_0 > 0$, then the eigenvalue becomes only imaginary. As we mentioned above, if the bifurcation coefficient passes through the critical point, equilibrium loses its stability and a limit cycle bifurcates from this critical point [13]. This is an example of codimension-1 bifurcation under a single bifurcation condition, $\Re\left\{\lambda_{1,2} = 0\right\}$.

Two-Dimensional Case

Let's look mere closely:

$$x_1^{'} = f_1(x_1, x_2, \alpha) \qquad \textbf{(2.12)}$$

$$x_2^{'} = f_2(x_1, x_2, \alpha) \qquad \textbf{(2.13)}$$

If the following conditions (nondegeneracy) are satisfied:

- $l_1(0) \neq 0$. It is worthy to remark that $l_1(\alpha)$ is called first Lyapunov coefficient and we will derive this coefficient below.
- $\mu^{'}(0) \neq 0$.

then, the normal form [24] of this system is as follows:

$$y_1^{'} = \beta y_1 - y_2 + \sigma y_1(y_1^2 + y_2^2) \qquad \textbf{(2.14)}$$

$$y_2^{'} = y_1 + \beta y_2 + \sigma y_2(y_1^2 + y_2^2) \qquad \textbf{(2.15)}$$

where $y = (y_1, y_2)^T, \beta \in \Re$ and $\sigma = sign\ l_1(0) = \pm 1$.

The Andronov-Hopf bifurcation can occur in two different type, supercritical or subcritical, which depend on following conditions (we write these conditions under normal form):

- **Supercritical** $\sigma = -1$, an equilibrium must be located at the origin and if $\beta \leq 0$ it is asymptoticaly [13] stable, otherwise unstable. There is a **stable** limit cycle where $\beta > 0$ with $\sqrt{\beta}$ radius.
- **Subcritical** $\sigma = +1$, an equilibrium must be located at the origin and if $\beta < 0$ it is asymptoticaly [13] stable, otherwise unstable. There is an **unstable** limit cycle where $\beta < 0$ with $\sqrt{\beta}$ radius.

Multi-Dimensional Case

There are additional conditions for n-dimensional case($n > 3$) [13]:

- $\lambda_{1,2} = \pm i\omega_0,\ \omega_0 > 0$
- number of stable eigenvalue n_s, $\Re\lambda_j < 0$
- number of unstable eigenvalue n_u, $\Re\lambda_j > 0$

with $n_s + n_u + 2 = n$.

First Lyapunov Coefficient

We can calculate the *first lyapunov coefficient* by Taylor expension of $f(x,0)$ at $x = 0$ [13]. As we mentioned above, sign of σ defines the which type of Andronov-Hopf bifurcation occurs whether it is supercritical, $(-)$, or subcritical, $(+)$. Indeed, this is also the sign of *first lyapunov coefficient*. Here the derivation for *first lyapunov coefficient* is:

$$f(x,0) = A_0 x + \frac{1}{2}B(x,x) + \frac{1}{6}C(x,x,x) + O(\|x\|^4) \qquad \textbf{(2.16)}$$

where we can define the multilinear functions of $B(x,y)$ and $C(x,y,z)$ as:

$$B_j(x,y) = \sum_{k,l=1}^{n} \frac{\partial^2 f_j(\xi,0)}{\partial \xi_k \partial \xi_l}\bigg|_{\xi=0} x_k y_l, \qquad \textbf{(2.17)}$$

$$C_j(x,y,z) = \sum_{k,l,m=1}^{n} \frac{\partial^3 f_j(\xi,0)}{\partial \xi_k \partial \xi_l \partial \xi_m}\bigg|_{\xi=0} x_k y_l z_m \qquad \textbf{(2.18)}$$

where $j = 1,2...,n$. Let $q \in \mathbb{C}^n$, complex eigenvector of $A_0 \Rightarrow i\omega_0 : A_0 q = i\omega_0 q$ and for adjoint eigenvector $p \in \mathbb{C}^n : A_0^T p = -i\omega_0 p$ with $\langle p, g \rangle = 1$. The *first lyapunov coefficient* is:

$$l_1(0) = \frac{1}{2\omega_0} Re[\langle p, C(q,q,\bar{q}) \rangle - 2\left\langle p, B(q, A_0^{-1} B(q,\bar{q})) \right\rangle + \left\langle p, B(\bar{q}(2i\omega_0 I_n - A_0)^{-1} B(q,\bar{q})) \right\rangle] \qquad \mathbf{(2.19)}$$

Here we will also explain the bifurcation theory with discrete-time dynamical system where we adress the computational model that we suggest.

2.2.2 Fold bifurcation (saddle node for maps)

In discrete-time dynamical systems (iterated maps), we can define the Fold bifurcation (saddle-node for maps, limit point) in the generated map as a birth of two fixed points (one stable the other unstable). Qualitative changes occur when the bifurcation parameter passes through the bifurcation point (critical point). If the bifurcation parameter reaches the critical point, system has an eigenvalue, $\lambda = 1$. This is the general condition for Fold bifurcations in all dimensions [13]. Lets consider a discrete-time dynamical system

$$x \mapsto f(x, \alpha),\ x \in \Re^n \qquad \mathbf{(2.20)}$$

where $\alpha \in \Re$ and f is a smooth function. So,

- For $\alpha = 0$ there exists a fixed point $x^0 = 0$
- Jakobian matrix $A_0 = f_x(0,0)$ with $\mu_1 = 1$

When the bifurcation point reaches the critical point, a fixed point bifurcates to two different fixed points, one of them will be stable and the other one will be unstable. The opposite of this situation can also be happen which means that a stable and unstable fixed point collapses and only one stable fixed point occur. This phenemenon can be seen in Figure 2.7

This is another example of codimension-1 bifurcation ($\mu_1 = 1$) where bifurcation occurs in one-parameter sets of smooth maps. The literal conditions (nondegenerecy

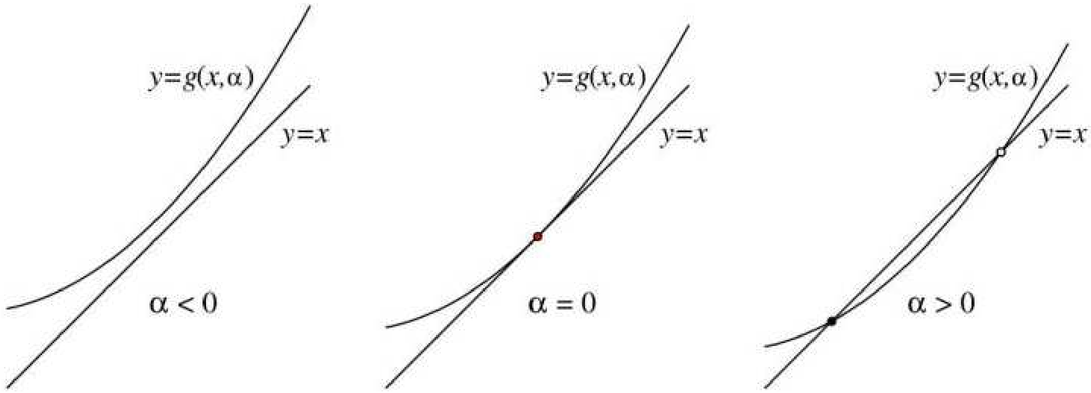

Figure 2.7: Illustration of Fold Bifurcation.

conditions) for 1-dimensional case are as follows:

$$f_x(0,0) = 1 \quad \textbf{(2.21)}$$

$$\alpha(0) = \frac{1}{2} f_{xx}(0,0) \neq 0 \quad \textbf{(2.22)}$$

$$f_\alpha(0,0) \neq 0 \quad \textbf{(2.23)}$$

To explain the quadratic coefficient, we consider the normal form for Fold bifurcation:

$$y \mapsto \beta + y + \sigma y^2 \quad \textbf{(2.24)}$$

where $y \in \Re$, $\beta \in \Re$ and $\sigma = sign\ \alpha(0) = \pm 1$. In the normal form, if $\sigma\beta > 0$ there is not any fixed points in the given system. In contrast, $\sigma\beta < 0$, system has one stable and one unstable fixed point where $y^{1,2} = \pm\sqrt{-\sigma\beta}$. At critical point, $\beta = 0$, there is only one fixed point, $y^0 = 0, \lambda_1 = 1$.

Multi-Dimensional Case

There are additional conditions for n-dimensional case($n \geq 2$):

- $\lambda_1 = 1$
- number of stable eigenvalue n_s, $|\lambda_j| < 1$
- number of unstable eigenvalue n_u, $|\lambda_j| > 1$

with $n_s + n_u + 1 = n$.

Quadratic Coefficients

The derivation of *quadratic coefficient*, α_0, which is involved in the second nondegenerancy conditions, and it can be calculated by Taylor expension as follows:

$$f(x,0) = A_0 x + \frac{1}{2}B(x,x) + O(\|x\|^3) \quad \textbf{(2.25)}$$

where we can define the multilinear functions of $B(x,y)$ as:

$$B_j(x,y) = \sum_{k,l=1}^{n} \frac{\partial^2 f_j(\xi,0)}{\partial \xi_k \partial \xi_l}\bigg|_{\xi=0} x_k y_l, \quad \textbf{(2.26)}$$

where $j = 1,2...,n$. Let $q \in \Re^n$, critical eigenvector of $A_0 : A_0 q = q$, $\langle q,g \rangle = 1$ and standart inner product $\langle p,g \rangle = p^T q$ with $p \in \Re^n : A_0^T p = p, \langle p,g \rangle = 1$. The *quadratic coefficient* is:

$$\alpha_0 = \frac{1}{2}\langle p, B(q,q) \rangle = \frac{1}{2}\frac{d^2}{d\tau^2}\langle p, f(\tau q, 0) \rangle \bigg|_{\tau=0} \quad \textbf{(2.27)}$$

3. SELECTING AN ACTION

When the brain collected input from more than one sensory modalities, it often processes with high and accurate performance in one or more of the neural subsystems. As there exists large sensory input from the environment, it is now accepted that each data possible to distribute to relevant subsystems. In the presence of different input modalities, it is claimed that making a decision is one of the tough processes in the brain. Comprehensive information processing ability in brain provides making a decision for availiable competitive actions by mediating related networks and formations in brain. The studies in neuroscience literature show that the BG circuits play key role in these formations [2, 10, 25].

The BG circuits are situated in the midbrain which means that this group of nuclei are the ancient part of the mammalian brain [26] due to the selecting an appropriate action is necessary for staying alive. In addition, from the point of evolutionary biology, BG circuits can be claimed as a window of cerebral cortex. In BG, unilateral or reciprocal connections exist through the cerebral cortex, cerebellum, thalamus, and other brain regions. BG circuits are associated with in a wide range of brain functions, such as perception, learning, memory forming, besides being effective in motor functions [2, 8, 10, 25, 27–30]. Their functions in cognitive processes as AS, GDB and selective attention have been studied throughly in recent years [2, 3, 27, 31, 32].

Models of AS processes in BG range from conductance-based models, through spiking neural networks, to system-level models [33]. Existence of at least five different loops of cortex-basal ganglia-thalamus has been suggested in [10]. Each loop is assigned different cognitive tasks but they have interconnection between each other, where these connections are claimed to be designed in a spiral frame [34]. The principle substructures of BG circuit are proposed to be comprised of input nuclei, Striatum (Str) and Subthalamic Nucleus (STN), and output nuclei, Substantia Nigra pars Reticulate (SNr) and Globus Pallidus internal (GPi). The major input station in BG is Str and this subcortical area is divided into subsections where two DA receptors are must effective. One of these receptors is D1 subtype and this regions project primarily to the output nuclei of BG to inhibit them. These output nuclei SNr and GPi in turn inhibit the cortex through Thalamus (Thl). This pathway is direct pathway:

$$Cortex\ (Excites)\ \rightarrow\ Striatum-D1\ (Inhibits)\ \rightarrow\ Snr/Gpi\ (Inhibits)\ \rightarrow\ Thalamus$$
$$(Excites)\ \rightarrow\ Cortex\ (Excites)\ \rightarrow\ Brainstem$$

Whereas in the indirect pathway D2 receptors are effective and they have inhibitory effect on Globus Pallidus external (GPe) to disinhibit STN through GPi. The indirect pathway is:

$$Cortex\ (Excites)\ \rightarrow\ Striatum-D2\ (Inhibits)\ \rightarrow\ Gpe\ (Inhibits)\ \rightarrow STN\ (Excites)$$
$$\rightarrow\ Snr/Gpi\ (Inhibits)\ \rightarrow\ Thalamus\ (Excites)\ \rightarrow\ Cortex\ (Excites)\ \rightarrow\ Brainstem$$

Based on these neurophysiological facts, many BG models now exist, all having different approaches especially in explaining the correlation of direct and indirect pathway. One approach considers the direct pathway as a main part of selection mechanism and claims that indirect pathway antagonize the direct pathway by suppressing unwanted movements [32, 35, 36]. To investigate these correlation in the Str, direct pathway is added in the previous model in [7]. Here, to give an insight for BG network, Figure 3.1 is given:

As we mentioned above, there are at least five different loops in BG for assigned different cognitive tasks. Now, we will explain which loop of BG we consider and how we modeled this loop by means of discrete-time dynamical system approach. The main contribution of this work is to extend the model that proposed in [7] including the direct and hyperdirect pathway to explain the DA effect on AS. The extended model helps us to explain the BG circuit's switchboard-like selection mechanism [2, 27] by using nonlinear dynamical approach with the bifurcation analyses which will be discussed in Chapter 4. Here, we will explain the model focusing on DA effect.

3.1 Modelling Action Selection

The sub-regions of neural system communicate with each other by interconnection neurons and realize any process via neurotransmitters along neural pathways. To

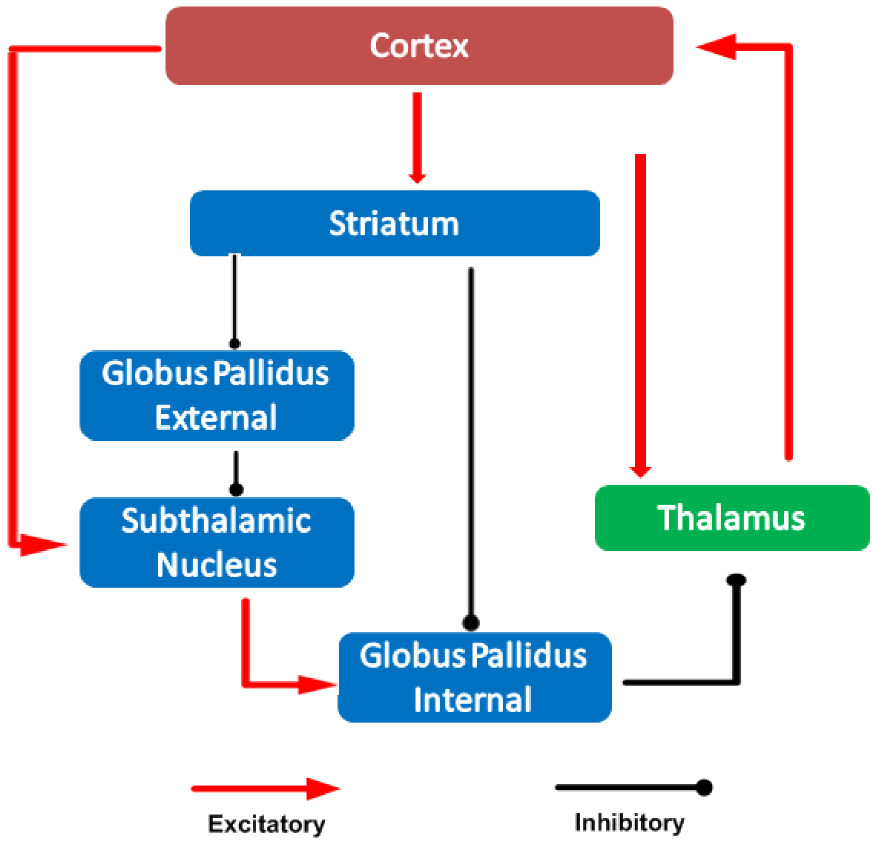

Figure 3.1: Basal Ganglia Circuit.

understand the mechanism giving rise to cognitive processes, we have to consider the neurotransmitter systems. As neurotransmitters have constitutive effect on cognitive processes, DA release has a modification role in BG. Amongst eight different dopaminergic pathway, nigrostriatal pathway modulates AS process in BG [37]. Striatonigral pathway (prefrontal loops in Basal Ganglia) is associated with motor control and related to dopaminergic pathway [28]. Dysfunction of this pathway causes disorders such as PD and HD [10, 38–40]. A proposed corticostriatal neural network model by Amos [41] is achieved to investigate the mental disorders by using Wisconsin Card Sort Test. The perseveration of Schizophrenic and Huntington's patients are demonstrated and suggested that the problem is caused due to unsystematic errors of matching in the Str where the DA modification is well-known today.

Besides different brain areas that produce DA, one of the transmission of DA starts from Substantia Nigra pars Compacta (SNc) through the Str. These regions are the part of the basal ganglia-thalamus-cortex circuits [10]. Retrograde and anterograde tracing studies have shown that the basal ganglia-thalamus-cortex circuits and a.k.a.

striatonigrostrital pathways have important role in AS and learning phenomena, as we mentioned above. There are a number of computational models evaluating the neurophysiologic work on BG and related neural substrates taking part in AS [2, 5, 7–9, 32, 33]. and also some others where cognitive processes are investigated considering the robot models [42, 43]. The cortico-striato-thalamic model considered in this work for implementation of GDB is based on these neurophysiological facts and is capable to explain how primates make appropriate choices and learn associations between environmental stimuli and proper actions [7].

In [10], different regions of BG are considered for different neural circuits, but principle substructures are proposed to be Str, STN, GPi and GPe, SNr and SNc. Due to the reason that the relationship between these substructures, for example cortex and Thl is very complex, the model used in this work consider only a subgroup of these relations which are important for action selection, so it is simpler. The connections considered in the model are illustrated in Figure 3.1. A part of this computational model of AS, where only indirect pathway is considered, has been shown to realize a sequence learning task [7]. Here, we have been extended the model in [7] by including direct and hyperdirect pathway to mimic the modulatory effect of DA release on Str.

We aim to model the modulatory effect of DA on AS with a computational model which can establish sufficient functionality to exhibit relevant behaviour in embodied robotics. The computational model has been developed at the system level [7] as in the well-known work of Prescott et. al. [42]. In [42], saliencies that exist in a tight competition effectively switch the behaviour of the computational BG model realizing the AS mechanism in the embodied architecture. The saliencies which control the behavior of the mobile robot are generated in the motivational and sensory sub-systems as a priori coefficients in [42]. Although the work in [42] is important as it shows that the biologically plausible models of the BG can be used for the control of physical devices, it is only a first step as it lacks the ability of determining saliencies by a learning process. We focus especially on this aspect and proposed a model [7] capable of generating an adaptive process where the parameters of the model modify the behaviour of BG circuit taking part in AS. In order to make the model functional and realize it in embodied robotics' task, the neural substructures, that are related to the

BG circuit we consider, are discussed as a mass model. Furthermore, each neural substructure acts according to the tangent hyperbolic function, $f(.)$:

$$f(x) = 0.5(tanh(3(x-0.45))+1) \quad \textbf{(3.1)}$$

Since the output of the neuronal activity can be defined as an all or none, we used the tangent hyperbolic function to fulfill these phenomenon, see Figure 3.2. Based on the

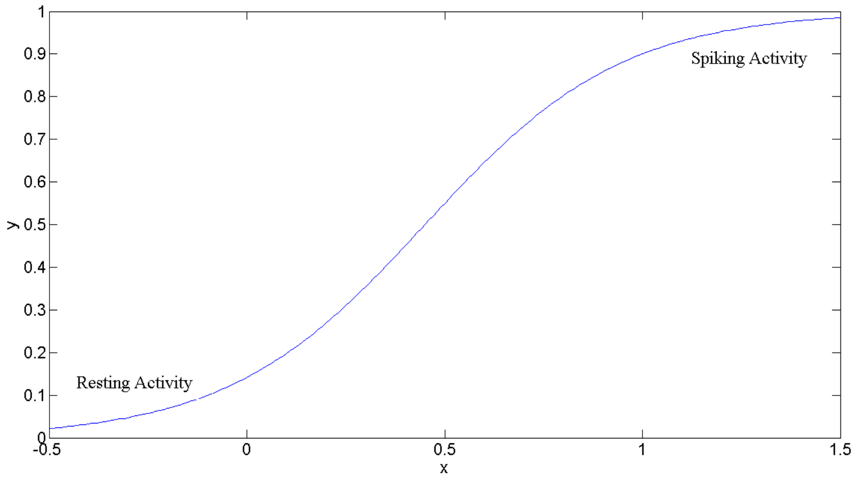

Figure 3.2: Tangent hyperbolic function.

discussion as mentioned above, a representation of BG circuit with its connection to related neural substructures are started with Cortex the following difference equations:

$$S(k) = W_c I(k) \quad \textbf{(3.2)}$$

$$Ctx(k+1) = f(\lambda Ctx(k) + Thl(k) + S(k)) \quad \textbf{(3.3)}$$

The input substructure of the model is Cortex which transmits the sensory data to the Str, Thl and STN. The main effect on cortex is due to excitatory signal from Thl which corresponds to the result of selection process. It is worthy remark that the variables Ctx and Thl stand for vectors corresponding to cortex and Thl constituents of prefrontal loop in BG, respectively. The dimensions of these vectors are determined by the number of actions to be selected, in this case, the vectors belong to $\Re^3$ as there are three actions to be selected that we consider in our robot task. $S(k)$ constitutes the salience matrice, and W_c matrice that we call efficiency of sensory input, $I(k)$, denotes the parameters which impress the environmental senses. W_c is the main bifurcation parameter in our system, we will discuss it in Chapter 4. An action is selected, when the value of one of the cortex variable becomes almost one. This corresponds to firing of related neural structure, as we mentioned above.

The interconnections between substructures Str, GPe, STN and SNr/GPi, that are respectively denoted by Str, GPe, Stn, GPi are modelled with same approach as follows:

$$Str(k+1) = W_r f(Ctx(k)) \quad \textbf{(3.4)}$$

$$GPe(k+1) = f(-Str(k)) \quad \textbf{(3.5)}$$

$$Stn(k+1) = f(Ctx(k) - GPe(k)) \quad \textbf{(3.6)}$$

$$GPi(k+1) = W_d f(Stn(k) - Str(k)) \quad \textbf{(3.7)}$$

$$Thl(k+1) = f(Ctx(k) - Gpi(k)) \quad \textbf{(3.8)}$$

There are two more bifurcation parameters in the model where W_r denotes the effect of DA release on Str and W_d denotes the correlation between the direct and indirect pathways in the circuit. These parameters are effective in selection mechanism, so bifurcation analysis, considering these coefficients, will be given in Chapter 4.

To show the model is capable selecting an action, as give in Figure 3.3. Since the output of the system corresponds to Ctx component, each dots represent one of which cortex variable is firing at the given input . Axises denote the values of three components of salience matrice, respectively. This figure is obtained by solving the Equation 3.2-3.8 in-house MATLAB with initial conditions are generated randomly. When there exist three actions to be selected, three dimensional salience-space comprises different subspaces where W_c, W_r and W_d reshapes this space. As the DA release has a key role in AS, we will show its effect on the computational model of BG. Once we select a priori coefficients which means that a basal DA release exists in the BG, a classification of three subregions are obtained.

In Figure 3.3a there are subregions of salience-space where only one of each salience determine the action to be selected. This means, model selects an appropriate action if the relative salience value is in one of these regions. In Figure 3.3b, two salience combinations out of three determine the action to be selected. The all or none selection regions is given in Figure 3.3c.

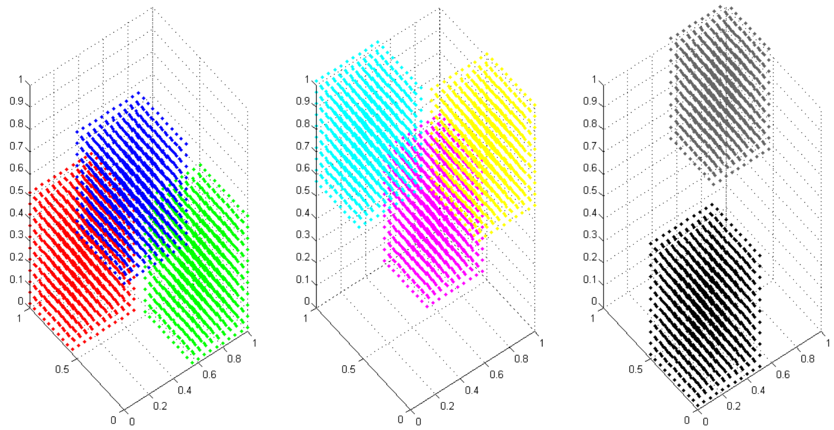

Figure 3.3: $W_c = 0.8$, $W_r = 0.3$ and $W_d = 1$. Legend of the colors; green: first salience, red: second salience, blue: third salience, magenta: first-second salience, yellow: first-third salience, cyan: second-third salience, black: all-selected salience, grey: non-selected salience, respectively. Legend is same as in Figure 3.4 and 3.5.

3.2 Related Diseases

Deficit or excessive level of DA in the D1 and D2 receptors in Str causes several behavioral disorders, such as PD and HD [10, 38–40]. The direct and indirect pathway have an antagonistic functions in BG and DA modulates them. D1 type DA receptors are effective in direct pathway where DA excites the inhibitory effects of Str in Gpi/SNr. This means that Str inhibits Gpi/SNr, so Gpi/SNr activity is decreased and in turn Thl activity is increased which causes a disinhibition in motor activity (one of the possible action selection). On the other hand, D2 type DA receptors are situated in indirect pathway. If the DA release increases then there would be less activity in GPe through Str. Less indirect pathway activity means again increased motor activity. If the balance of the direct and indirect pathway breaks, selecting an action process becomes easy or difficult for body which points out the important behavioural disorders.

The model that we propose can be further used to establish a framework for understanding the cause of physiological diseases related with BG circuits. Based on the bifurcation analysis that we will explain in Chapter 4, the proposed model provides us results to discuss these disorders. With bifurcation analysis, the proposed model shows that the DA effects on the Str causes the transition of the salience-space. The transition of parameter demonstrates that there is an association between subcortical dysfunctions and DA effects on Str in BG circuit. These simulation results are consistent with neurophysiological experiments [10, 38–40].

It is now accepted that loss of dopaminergic neurons in SNc induces the dysfuntion on Str which causes less activity in direct pathway, in contrast more activity in indirect pathway [38]. As the level of the DA has a role in different physiological disorders, it is well-known that loss of DA level in BG circuit causes PD which means difficulty in accomplishing an action as in Figure 3.4. As it is followed from Figure 3.4c, selecting an action became more difficult (increased black dots).

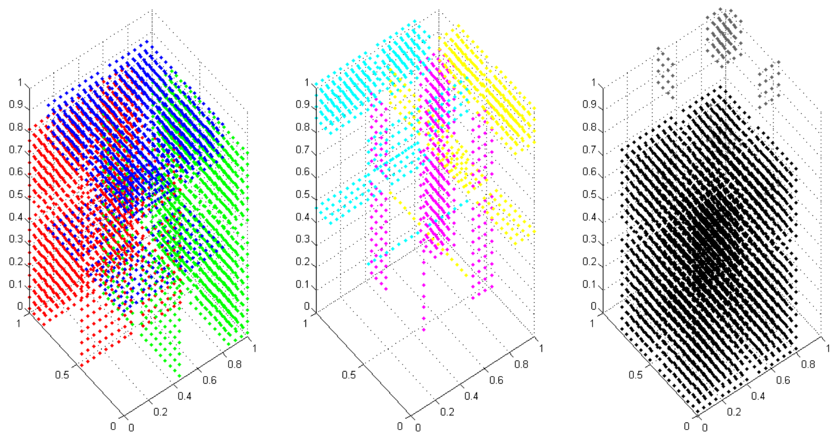

Figure 3.4: $W_c = 0.8$, $W_r = 0.1$ and $W_d = 1$.

HD is another neurophysiological disorder which is related with BG circuit [39, 40]. The key sypmtoms of HD are difficulty in muscle coordination, excessive motor behaviours and cognitive decline with psychiatric issues. HD typically causes an abnormal involuntary movements, difficulty in initiating appropriate actions and inhibiting inappropriate actions, which is the first step of this disorder and is called with chorea [39]. In addition, HD has several impact on cognitive abilities such as planning, cognitive flexibility and rule acquisition [40].

Even how the BG damage causes the HD is not fully understood, it is now accepted that excessive DA level in BG circuit has role in HD which means choosing more than one action at a time, as it can be seen in Figure 3.5. In contrast with Figure 3.4c, selecting an action become very easy (increased grey dots in Figure 3.5c).

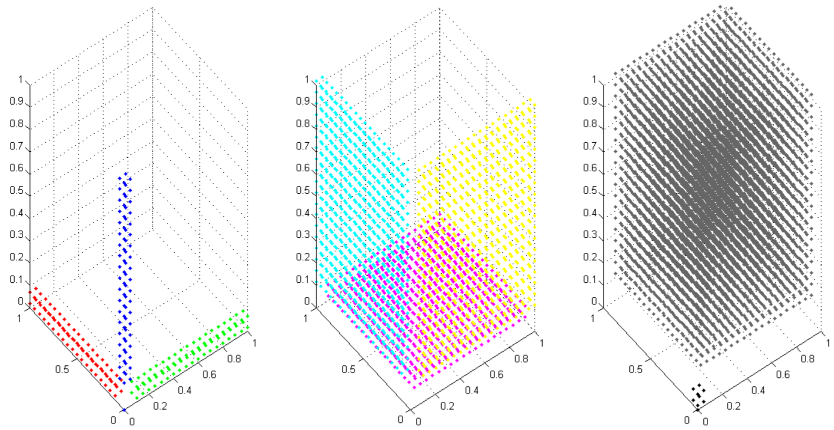

Figure 3.5: $W_c = 0.8$, $W_r = 0.8$ and $W_d = 1$.

So, to avoid this undecisive situations, RL has been implemented to the model and once the learning is completed, system selects the right decision at the right time in all circumstances. We will continue to explain the learning architecture in proposed model next chapter.

4. LEARNING TO SELECT AN ACTION

Goal-Directed behaviour comprises two different concepts; Action Selection and Learning as mentioned above. For many years, neuroscientists were strictly thinking that BG circuits have role in motor behaviour control only. The motivation of this approach was based on the motor control damaged patients who have certain symptoms in their BG substructures [44]. However, recently, there are different works showing the role of BG in different cognitive and emotional functioning such as reward related learning, habit forming, GDB and RL [45, 46]. In this chapter, we will focus on how we constitute the learning and AS together based on the structures of the BG circuits and we will suggest an algorithm to explain the GDB in primates.

The main focus of our approach is to bring the dorsal loop of Cortico-Striato-Thalamic circuits and TDL together. The reason that we use TDL is there are works relating learning process in the brain with temporal difference, especially in BG circuit [1, 4]. Indeed, due to the reason that recent studies show the learning in BG is somehow exactly same as RL, we considered the TDL in constituting the GDB [1]. The interaction between the different loops in BG take part in different cognitive functions beside motor control. Also effect of dopaminergic modifications on learning is well-known in neuroscience literature nowadays [1, 47, 48]. So, we especially focused on the DA effect in Str during the modelling process by means of bifurcation theory. Since the main aim of this work is to adress the relationships between AS and RL, TDL has been implemented into the model considering these bifurcation analysis. As the model is expected to find correct (appropriate) choice after learning is accomplished, we investigate the model by considering the Bifurcation Theory before the implementation of TDL.

For more than fifteen years, nonlinear dynamical systems have been important part of neurocomputational modelling and have consistently became a common tool for the computational neuroscientists to explain the networks in brain. Nonlinear dynamical system approach for modeling the neural structures gives us the possibility to control the system with bifurcation parameters. After obtaining the whole architecture of the model's behaviour by means of the bifurcation analysis, we implemented the RL into the model. Therefore, in this work, we considered the neurocomputational model of BG Circuit with bifurcation theory and RL together, and suggested an algorithm to constitute the learning architecture. We explain the switchboard like mechanism [2, 27] in BG Circuit with Fold Bifurcation (FB). This investigation confirmed that the given model can be modified to obtain appropriate behaviour if it satisfies FB conditions. To realize this modification, RL is utilized and with RL the parameters of the system corresponding to bifurcation parameters are updated to model learning to select appropriate actions in an unfamiliar environment. Thus, first, the bifurcation analysis will be completed based on different bifurcation parameters, then the TDL will be explained with further details. Finally, we will conclude this chapter with how bifurcation analysis and RL are brought the idea of the selection process in BG circuits.

4.1 Bifurcation Analysis

Neurocomputational model for AS circuit given with Equation 3.4-3.8 is investigated using bifurcation analysis and this investigation confirmed that the given model can be modified to obtain appropriate behaviour. Here, we will start to explain the bifurcation analysis with the one parameter bifurcations, then we will use two-parameter bifurcation analysis to explain the structure of the model.

4.1.1 One-Parameter bifurcation analysis

First of all, we used the one parameter bifurcation analysis of the system given in Equation 3.4-3.8 according to parameter W_c in neurocomputational model of BG circuit, here we found FB (the explanation of FB can be found in Section 2.2) in this system. In the context of AS, system has to be able to select an action depending on initial conditions or saliencies. Besides there may exist two or more than two actions to be selected, we use FB to explain the structure of this selection process. Due to two different fixed points existing in the system at the same time, one of them corresponds to the "not to choose" the choice and in contrast, the other one corresponds to "choose".

Starting these analysis with different parameter values have a role to change the bifurcation points and it is obvious that also this continuation allows us to explain the different fixed points locations. We will use this property in the system to explain the AS process in the context of the given task which can be found in Section 5. The model satisfies the fold (saddle-node, tangent) bifurcation conditions (see Section 2.2) where one stable and one unstable fixed points exist in the generated map even may be at the same time or different locations. The location of the different stable fixed points depend on parameter values. In the critical points which are FB points, one of the fixed point disappears and may be at the same time or with some delay another fixed point borns. As we mentioned above, we can reshape the bifurcation diagram with using different parameter values. First, we will investigate that the system given in Equation 1 undergoes one FB, this means the second unstable fixed point borns after the first fixed point disappears and then the unstable one disappears during the second stable one borns, consecutively and this explanation gives rise to the switchboard like mechanism in AS circuit. The bifurcation diagram can be found in Figure 4.1 with eigenvalues during the bifurcation parameter is iterated.

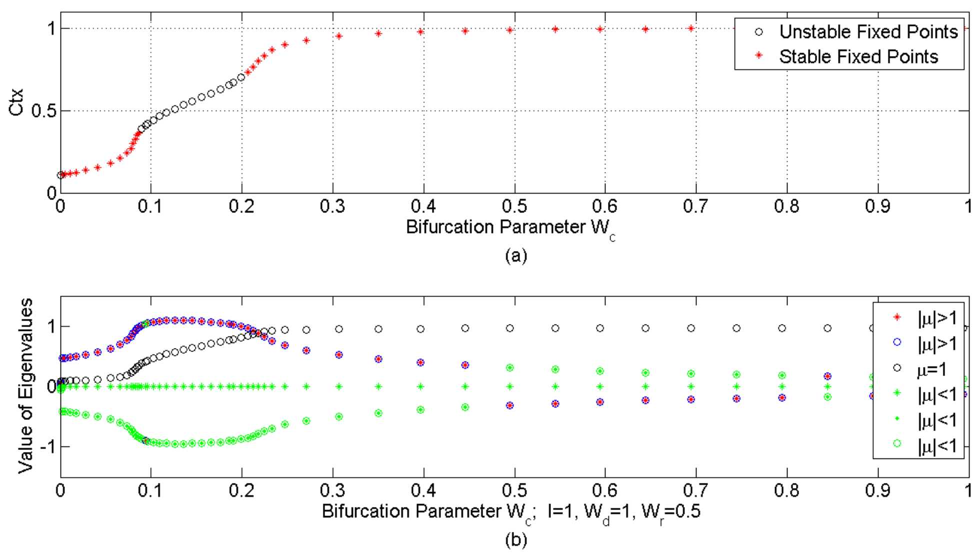

Figure 4.1: a. The circles denote the unstable fixed point, dots denote the stable ones. System has two different stable fixed points with different locations. One of them corresponds to not-selecting an action (near "0"), the other one correponds to selecting an action (near "1"). b. Change of value of the eigenvalues with parameter W_c. There one eigenvalue is fixed at "1", three ones are "<1" and two ones are ">1" in FB point.

As it can be followed in the Figure 4.1, in the generated map, system have different stable fixed points which describe to select and not to select an action, respectively. There exist one stable fixed point which is near “0” and then it disappears and later reappears around “1” while the bifurcation parameter $W_c \in [0,1]$. As the bifurcation parameter reaches FB when $W_c \cong 0.21$, the unstable fixed point collapses with stable one and disappears while another stable fixed point around “0.98” borns. Thus there exist a region where unstable fixed point is observed.

At the critical point where bifurcation parameter $W_c \cong 0.21$, the eigenvalue conditions of being outside the unit circle for unstable and being inside the unit circle for stable case are satisfied as it can be seen in Figure 4.1. The stable fixed point near “0” means that system cannot select an action, on the other hand the stable fixed point at “0.98” means that the system selects an action. Even though we expect to see two stable fixed points, there exists only one.

As we mentioned below, we can reshape this bifurcation diagram using different parameter values. We will use this approach to explain the significant situation in the given task, see Section 5. Here, the system has two different FB at different locations as it can be seen in Figure 4.2. The parameter values can be seen in both Figure 4.1 and 4.2. The difference betweens two maps are; first the FB numbers and then the bi-stable phase portrait occurs in the second situation. We can see the two FB with following the eigenvalues in Figure 4.2. Between the two FB, following system satisfies the eigenvalue conditions for FB. We can give more precise knowledge about this situation; in Figure 4.1b black circles which show the first conditions of FB (see Section 2.2) start to reach “1” during the second stable fixed points birth not the process of the first stable fixed point disappear, and this means system has only one FB during the second stable fixed point birth because the conditions of FB. On the other hand, same as the explanation in first case, one of the eigenvalue, which is denoted by black circles in Figure 4.2b, fixed at “1” while the process of first unstable-stable fixed points collapse and the second stable-unstable fixed points collapse.

It is worthy to remark that between the significant bifurcation parameter values ($W_c \in [-0.22, 0.403]$), there exist two stable fixed points at the same bifurcation parameter value. This situation allows us to use these two different fixed points for different

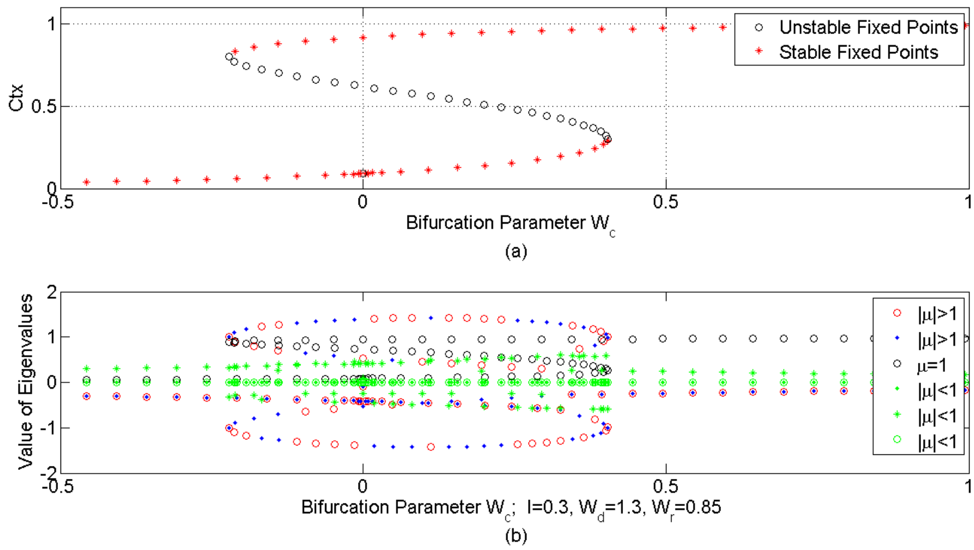

Figure 4.2: Demostrations are same as in Figure 4.1 a. System has two different stable fixed points between the two FB. One of them corresponds to not-selecting an action (near "0"), the other one correponds to selecting an action (near "1"). In this situation, system has a bi-stable phase portrait which is represented by FB. b. Change of value of the eigenvalues with parameter W_c. As it can be followed in Figure 4.1b, FB conditions are satisfied between two Bifurcation points, see Section 2.2.

tasks as it will be explained in Section 5. Around the bifurcation point there are two domains of fixed points between "-0.22" and "0.403". When we fix the parameters in this region, the proposed model decides to select/or not select an action depending on the initial conditions. This explains how the system given by Equation 3.4-3.8 accomplishes AS according to the input value W_c. So to observe the effect of input value W_c while changing W_r, another example can be found in Figure 4.3.

Before start the explanation about two bifurcation analysis, we want to show the significant paramater values effect on system's behaviour as the modification of W_r can also change the dynamic behaviour of the system. As shown in Figure 4.3, in the beginning, system has only one fixed point but there is a region where non-convergent solutions (quasi-periodic) began and changing the value of the parameter further causes system to settle down at another fixed point. The quasi-periodic behavior corresponds to the case where search is carried, so the dynamic systems behavior is not settled to a fixed point. As it can be seen in Figure 4.3, the other bifurcation parameter (in this case, W_r) has a role in changing the dynamic behaviour of the system. So, we will

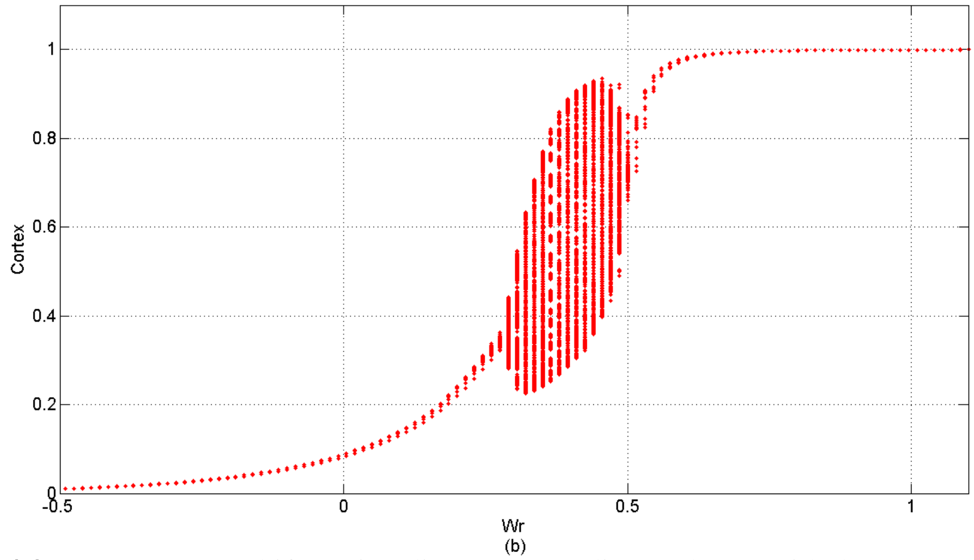

Figure 4.3: One parameter bifurcation diagram according to the W_r with the parameter of $W_c = 0.2$.

continue to explain the dynamic behaviour of the system with using two-parameter bifurcation analysis.

4.1.2 Two-Parameter bifurcation analysis

We aim to model the modulatory effect of DA on AS with a computational model which can establish sufficient functionality to exhibit relevant behaviour in embodied robotics. As it is explained in Section 3.1, W_r corresponds to the DA effect on Str and W_d the correlation between the direct and indirect pathways in the circuit. These parameters are effective in selection mechanism, so bifurcation analysis with considering these coefficients is important to constitute the learning algorithm in the model. To observe the effect of input value W_c while changing W_r, the two parameter bifurcation diagram given in Fig. 4.4 is obtained. Here while $W_r < 0.23$ and $W_c \in [0,1]$ there is one stable fixed point and while keeping $W_c \in [0,1]$ but changing $W_r \in [0.23, 0.51]$ non convergent solutions (quasi-periodic) begin till another stable fixed point appears. This quasi-periodic solutions mean that system cannot decide which action to select. How the system's salience-space is rehaped can be explained with this bifurcation analysis. As it can be seen in Figure 3.3-3.5, with different parameter values, different salience-space configuration is formed. The ability of the system to select an action is increased or decreased with these parameters. Based on

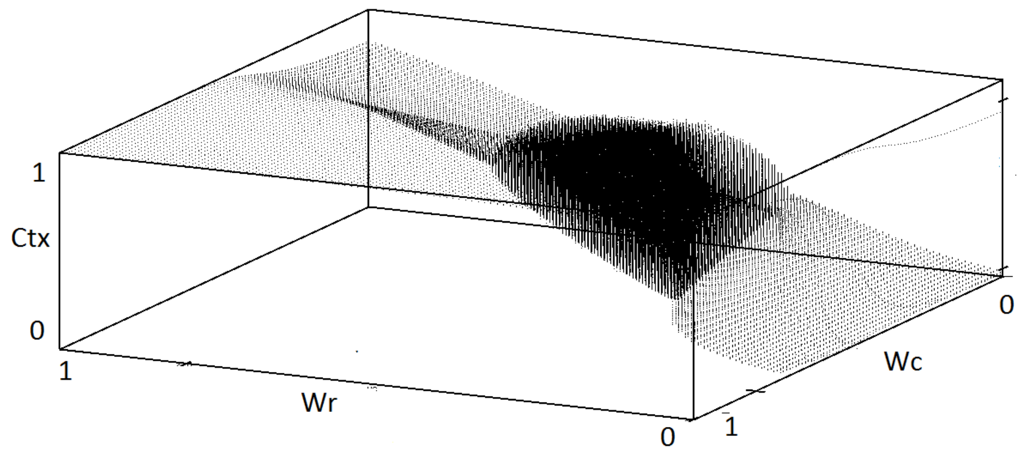

Figure 4.4: Two parameter bifurcation diagram for W_r and W_c.

the discussion in Section 3, DA release on Str has a significant role to explain the selection process in brain. This analysis explains the DA effect on Str: when the stable fixed point is located around "0" even the input value is high enough, system cannot select an action (This is shown in Figure 3.4, where the area of the selection is decreased in Figure 3.4a, and the non-selection area is increased in Figure 3.4c) where the DA level is less to select a desired action. On the other hand when the DA release is increased, system selects the action in all circumstances which corresponds to HD (This is shown in Figure 3.5, where the area of the selection is increased in Figure 3.5c, and the non-selection area is dicreased in Figure 3.5c).

There exists another bifurcation parameter, W_d, which corresponds to correlation of the direct and indirect pathways in the BG circuit. The bifurcation analysis for W_d shows that the model's behaviour becomes unstable when $W_d \in [0.67, 0.75]$ (Figure 4.5). If this connection increases constantly, system possibly selects an undesired action and it means that indirect pathway cannot antogonizes the direct pathway to select an appropriate action.

4.2 Reinforcement Learning

Reinforcement learning is a learning formation with respect to the interaction between agent and environment [49]. Based on this interaction, agents make the appropriate decisions by evaluating the consequences of these decisions. These evaluations depend on different perspectives such as; past experiences, new conditions or

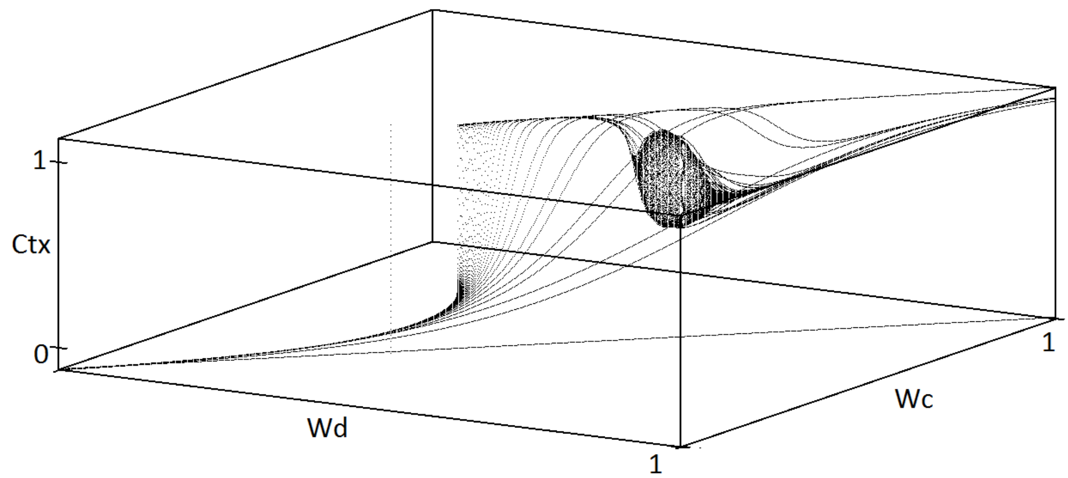

Figure 4.5: Two parameter bifurcation diagram for W_d and W_c.

reward/punishment expectations. The main architecture of RL depends on trial/error learning with essential reward predictions [4, 31]. In RL, the ultimate goal is to reach the minimization of the error in expectation through the maximization of value for reward which are calculated over time. The evaluation to find the appropriate output of network in RL depends on value assignment about given situation with reward/punishment formation rather than consider by any teacher's assessments. The basic definition of RL can be given with reward/punishment formation in which the network is matched the association between reward/punishment and stimuli/actions. Each action, agents decide, creates a reinforcement by means of particular situation and this is explicitly a result of unsupervised learning: not from a teacher's assessments, from the basis of reinforcement they receive themselves.

The experimental studies on animals, psychologists have focused on more than a century have an important impact in constitution of TDL framework. The studies on the TDL theory have been started with Rescorla-Wagner where the Pavlovian conditioning is considered adressing the correlation of learning theme and associative stimulus [50]. Most of studies starting with this theory point the explanation of how human brain evaluates the error prediction. In comtemporary neuroscience, as many studies have suggested [31,51], the midbrain, significantly nigrostriatal DA pathways, is the adress of these learning phenomena. We can claim that TDL is a learning framework where the Pavlovian conditioning experiments is the base of it. There exist different studies both in experimental and computational neuroscience to understand the DA release and its effect on significant brain networks and these studies point out the reward related learning and also drug addiction [1, 11]. This is the reason why we have been used TDL within computational model of AS circuit. Besides different disciplines considering RL for the problems they deal with, the principle case that we are interested in RL is Machine-Learning. We can summarize the framework of TDL in Machine-Learning with two discriptions: Actor-Critics and Q-Learning [49].

We have been focused on neural circuits with the aspect of mathematical and computational models, where we deal with how we can explain the learning mechanism with unsupervised learning approaches. The main focus of our work is inherented in RL, particularly in TDL. We have defined the AS as making an appropriate decision at the right time, see Section 3.1. Facing with rewards or punishment can cause significant effects on decision-making which means the appropriate actions depend on the adaptation of the conditional changes in existing environment. The ultimate goal of our work by which the reward or punishment has a role to make an appropriate decision is to represent the neural formation of learning. The TDL we use to handle with finding the appropriate decision is based on neurobiological and psychological experiments, as mentioned above. In addition, this approach can give an insight of modification process in synapsis [52].

We will discuss the process of TDL within our study. So, we will start to explain the learning process of the salience differences which depends on the adaptation of the efficiency of sensory input matrix, W_c. Then, we will continue with sequential learning where we thereafter consider the DA effect on Str, W_r. Before concluding this chapter, the whole architecture of the learning process in our work will be given with pseudocode that we have been suggested. We have been simulated the model with TDL in MATLAB before implementing on the robot.

4.2.1 Adaptation of the efficiency of sensory input

The parameters of the dynamical system corresponding to neurotransmitters are modified with TDL. Here, we will explain the efficiency of sensory input (W_c) adaptation through TDL to determine the proper action. To modulate the AS circuit (see Section 3.1), the adaptation equations start with as follows:

$$W_c = W_c + \eta_c \delta(k)_c f(ctx(k))S(k) \tag{4.1}$$

Here, the learning rate $\eta_c = 0.2$, δ denotes the error in expectation, $f(ctx(k))$ represents the output value of the AS circuit and $S(k)$ represents the multiplication of input value, $I(k)$, and efficiency of this input value, W_c.

The reward function is determined by using AS circuit's output. If the system find the desired action, this function is set to "1" for reward otherwise "-1" for punishment. Then, system evaluates the error in expectation with following equations same as traditional TDL:

$$\delta_c(k) = r_c + \mu V_c(k+1) - V_c(k) \tag{4.2}$$

Error in expectation δ depends on the value assigner which is attained to the selected action and future expectation with reward or punishment. The step size of learning depends on discount factor, μ. The generated expectation value is based on value assigner of the action which is also updated as follows:

$$V_c(k) = W_{v_c}(k)S(k) \tag{4.3}$$

$$V_c(k+1) = V_c(k) + \mu_c \delta_c(k)S(k) \tag{4.4}$$

Once the saliencies are established, the AS circuit determines an action. This selected action and the reward obtained for it interacts with the learning and the selection process in AS restarts at each time step. The simulation results of the adaptation process for efficiency of sensory input, W_c, can be seen in Figure 4.6. As it can be seen in Figure 4.6, error in expectation converges the zero when the reward function is "1" , otherwise the error in expectation diverges the zero. Once the error in expectation reaches approximately "0", W_c (efficiency of sensory input) is fixed at the ($W_c \cong 0.5$). This final W_c value corresponds to the significant meaning for our system's behaviour. As it can be seen in Figure 4.1.a, if the W_c value is bigger than the bifurcation point, system has only one fixed point to determine the action to be selected.

After this learning process, system is able to distinguish the differences between the actions to be selected. In order to compare the learning effect on AS process, the

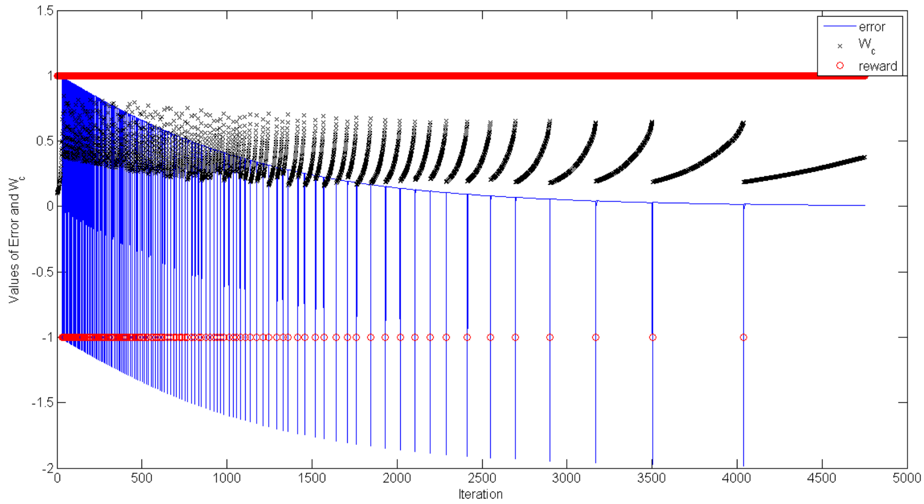

Figure 4.6: Temporal difference learning process for W_c.

efficiency of sensory input matrices for beginning and for after learning process is given as following:

$$W_c^{beginning} = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}, W_c^{after} = \begin{bmatrix} 0.5235 & 0.1314 & 0.0225 \\ 0.0233 & 0.5311 & 0.0223 \\ 0.0236 & 0.1316 & 0.5226 \end{bmatrix} \quad \textbf{(4.5)}$$

The differences between each matrices show that the value on the diagonal of W_c are set up the latter of FB point (Figure 4.1). There less significant changes occur rest of the matrices' components but these modificition can be omitted due to the system's architecture as mentioned in Section 2.1.

As it can be followed in Figure 4.6, system evaluates an error in expectation every step time which depends on reward/punishment value. The difference between the error in expectation in each step time is the result of the differences between the reward/punishment and expectations. If the expectation of the value assigner is not similar to the given reward or punishment, system evaluates significantly bigger error.

4.2.2 Adaptation of dopamine effect on striatum

We have extended the learning process in AS circuit by adapting parameter W_r which corresponds to DA effect on Str. The main reason of this learning process is to mimic the modulatory effect of this neurotransmitter in BG circuit. Based on the bifurcation diagrams for the W_c and W_r parameters (see Section 4.1), we are able to explain not only the importance of DA on determining the saliencies but also differences between each salience circumstances. The ability of the system to decide the action's importance can be controlled with DA parameter, W_r. The bifurcation analysis and learning idea depend on the DA effect on learning process, whereas the psychological experimental results show the importance of this neurotransmitter in primate's brain. The modification of the W_r parameter starts once the TDL for W_c parameter is accomplished. The TDL equations for W_r parameter is similar to the modulation of W_c parameter with certain differences. The equation for the modification of W_r parameter as follows:

$$W_r = W_r + \eta_r \delta_r(k) f(Ctx(k)) Str(k) \tag{4.6}$$

where η_r denotes discount factor, $Str(k)$ denotes Str as mentioned in Section 2.1. The error in expectation:

$$\delta_r(k) = r_r + \mu_r V_r(k+1) - V_r(k) \tag{4.7}$$

where $\mu_r = 0.2$ in this case. The value assigner equations are:

$$V_r(k) = W_{v_r}(k) Str(k) \tag{4.8}$$

$$V_r(k+1) = V_r(k) + \mu_r \delta_r(k) f(Ctx(k)) \tag{4.9}$$

Once the TDL process for W_r is accomplished, system is able to decide the importance of each action. It is worthy to remark that learning process of W_c is important for the competition of saliencies between each other, in contrast after the learning process of W_r, system focuses on each saliencies respectively and understand the importance of them. The differences between the beginning and after learning for W_r for each salincies can be found as follows:

$$W_r^{beginning} = \begin{bmatrix} 0.0560 & 0.0613 & 0.0301 \end{bmatrix}, W_r^{after} = \begin{bmatrix} 0.8559 & 0.7641 & 0.7540 \end{bmatrix} \tag{4.10}$$

The simulation result of TDL for W_r parameter can be seen in Figure 4.7.

The concepts of this learning process is same as the learning for W_c. Once the learning is accomplished, the error in expectation converges "0" where these errors depend on the reward/punishment value. In Figure 4.7, it can be seen that the $W_r \cong 0.8$ which is the desired value due to the our main aim as discussed in Section 4.1. In this case,

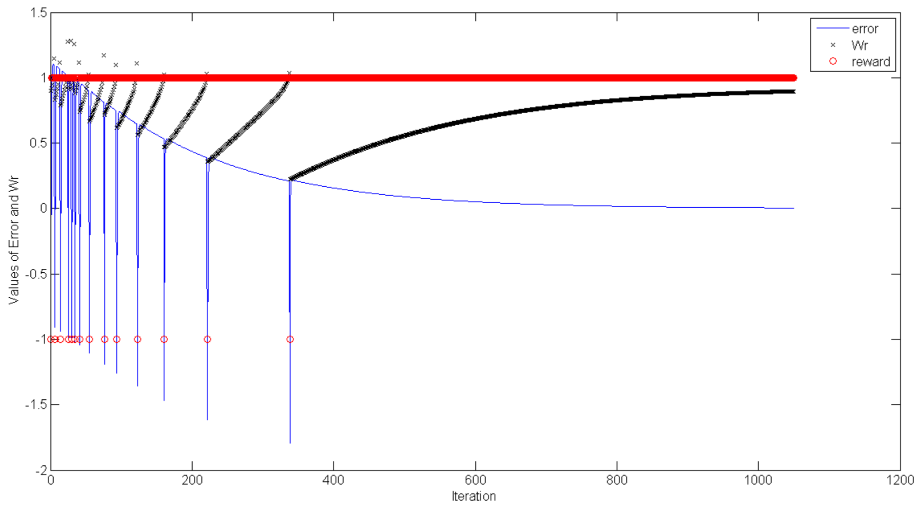

Figure 4.7: Temporal difference learning process for W_r.

system has two different fixed points at the same time, so system is able to decide to which fixed points based on the importances of occuring case (see Figure 4.2).

4.2.3 Proposed method

The algorithm we have been proposed for learning is constituted from bifurcation analysis both in one-parameter and two-parameter through TDL for W_r and W_c parameter, as explained above. Considering these two concepts together helps us to propose an approach to understand how primates make appropriate choices under different circumstances. It seems intuitively obvious that exploration of a new maze would lead to be a first step of our method to be tested.

Algorithm 1 Calculate W_c and W_r

Require: These codes work on Webots 6.0.0 and all of the pseudecodes are given with C Programming Synopsis.
Ensure: You should include the related library at the beginning of your codes, such as if you want to use the distance sensors you should call the distance sensors library.
Require: $0 < W_c < 1 \wedge W_d < 0.67$
Ensure: $\mu = f_x(0,0) = 1$
$\forall |\mu_{stable}| < 1 \wedge \forall |\mu_{unstable}| > 1$
$n_{stable} + n_{unstable} + 1 = n$
$s \Leftarrow W_c I$
while $Error > |0.04|$ **do**
  for (**do** i=0;i<51;i++)
p $(MotorCortex)$
  $p1..3[i+1] = L\_func(lambda2 * p1..3[i] + m1..3[i] + s1..3)$
m $(Thalamus)$
  $m1..3[i+1] = L\_func(p1..3[i] - d1..3[i])$
r $(striatum)$
  $r1..3[i+1] = Wr1..3 * L\_func(p1..3[i])$
e $(GlobusPallidusExternal)$
  $e1..3[i+1] = L\_func(-r1..3[i])$
n $(SubthalamicNuclues)$
  $n1..3[i+1] = L\_func(p1..3[i] - e1..3[i])$
d $(GlobusPallidusInternal)$
  $d1..3[i+1] = L\_func(Wd * n1..3[i] - r1..3[i])$

  end for
  $N \Leftarrow Ctx(k)$
  if $N < 0.85$ **then**
    $Reward_c \Leftarrow 1$
  else
    $Reward_c \Leftarrow -1$
  end if
  $Value_c(:,kk+1) \Leftarrow val_c * I'$
  $Error_c(:,kk) \Leftarrow Reward_c + \gamma_c * Value_c(:,kk+1) - Value_c(:,kk)$
  $val_c \Leftarrow val_c + nu_c * Error_c(:,kk) * S(k)$
  $W_c \Leftarrow W_c + nu_c * [Error_c(:,kk)]' * W_c. * f(N). * S(k)$
  if $W_c > 0.403$ **then**
    $Reward_r \Leftarrow -1$
  else
    $Reward_r \Leftarrow 1$
  end if
  $Value_r(:,kk+1) \Leftarrow val_r * Str(:,k)$
  $Error_r(:,kk) \Leftarrow Reward_r + \gamma_r * Value_r(:,kk+1) - Value_r(:,kk)$
  $val_r \Leftarrow val_r + nu_r * Error_r(:,kk) * f(N)$
  $W_r \Leftarrow W_r + nu_r * [Error_r(:,kk)]' * W_r. * f(N) * Str(:,k)$
  $kk \Leftarrow kk + 1$
end while

5. AN APPLICATION: FORAGING TASK

In order to explain the efficiency of brain-inspired computational model summarized in Chapter 4 and investigated in Chapter 5, we have been tested our model on robotic application which is considered in cognitive robotics' topics. In the framework of cognitive robotics, the principle aim is to model the animal's cognition while information processing in robot mimics the formations of brain as opposed to Artificial Intelligence approaches [53, 54]. The cognitive robotics applications have capability to demostrate a wide range of brain formations, such as processing of perception, attention, planning, motor behaviours and perhaps even about their instrinsic states [42, 43, 53]. Recently, the applications of these cognitive processes have been increased throughly in computational neuroscience since the autonomous devices became important for medicine devices or elder services.

In cognitive robotics, the interaction between the agents and environment is the challenging step that has to be solved. On the contrary of usual artificial intelligence algorithms, specifically supervised learning, brain-inspired mathematical models embody the unsupervised learning while mimicking the information processing in brain. Since the ultimate goal of cognitive robotics is creating an intelligent agent which act in the real world, brain-inspired computational models have a capability to demostrate the sensation, action selection and motor behaviour at the same time. The computational model that we suggested above serves to demonstrate these abilities under competitive situations.

For more than ten years, cognitive robotics applications have been studied to explain the cognition processes in brain where the neural substructures and the connections between them are considered [42, 53, 54]. In the well-known work of Prescott et.al. [42] a robot model for action selection is given in the embodied architecture. This robot model mimics the behaviour of a rat in an unfamiliar environment and it is based on mathematical model of basal ganglia which is inspired by neurophysiologic studies [2, 8]. In [42], saliencies that exist in a tight competition effectively switch the behaviour of the computational BG model realizing the AS mechanism in the embodied architecture. The saliencies which control the behavior of the mobile robot are generated in the motivational and sensory sub-systems as a priori cofficients in [42]. Although the work in [42] is important as it shows that the biologically plausible models of the BG can be used for the control of physical devices, it is only a first step as it lacks the ability of determining saliencies by a learning process. In this thesis, we focused especially on this aspect and extended the proposed model [7] which capable of generating an adaptive process where the parameters of the model modify the behaviour of BG circuit taking part in AS.

In this thesis, the idea is to accustom the work in [42], and develop it further by implementing reinforcement learning to determine the saliencies which influence the choice of the rat. Thus, it is shown that robot implementation of neural circuits which are capable of realizing reinforcement learning is possible. It has to be emphasized that a more complex cognitive process than action selection, i.e., goal-directed behavior is implemented on a mobile robot, Khepera II. It is shown that basal ganglia take part in selecting an action amongst different choices based on the saliencies of each possibility. It is worthy to remark that again, the process of learning has not been considered in [42], where the choices depend only on a priori saliencies where they need busy signals to avoid the inconvenience situations. So the saliencies are reconsidered and priority of one over the other is determined according to the environmental conditions with reinforcement learning. It is shown that a simpler model of the cortico-striato-thalamic circuit considered for action selection can fulfill the expected behaviour based on these saliencies. Thus, the improvement of this work over [42], is the utilization of reinforcement learning to determine the choices and this is provided by using a simpler model of cortico- striato-thalamic circuit for action selection [7].

In this chapter, we present the fundamental concepts about the task that we solved. Then, the description of the whole architecture of the model and the experiments that show the differences between the a priori situation and learning process. After that, we will continue to explain the simulation results for the implementation of the proposed model on the mobile robot Kheperea II. Before concluding, the dopamine effect on the learning process will be given.

5.1 The Task and Environment

The task of the robot is inspired from [42] where the robot mimics the behaviour of a rat's search for food in an unfamiliar environment. The action selection task that we considered in this thesis based on this scenario depends on three saliencies which are search-detection, pick up-carry the food and return to the nest. We have been extended the task suggested in [42]. The task is planned such that as the feeling of fear ceases, the rat begins to search for food and during random search, learns to recognize food supplies and picks up the food and carries it to the nest. In an unfamiliar environment the rat's first feeling is to fear and action choice is not to move, but sometime later their need for food becomes irresistible and they begin to feel hunger. As time goes by, they start to feel more confident and the feeling of fear ceases. These two feelings, fear and hunger are intrinsic processes for the rats. Hunger depends on time and fear depends on environmental conditions which may be the existence of lights, foreign animals or noises. The differences between the sensory systems of the mobile robots and the rat have to be resolved in order to mimic the behavior of a rat. In rat's, the central nervous system is responsible in recognizing the food and differentiating it from other objects. They use their eyes and/or olfactories while mobile robots use the distance or light sensors to recognize the objects. Khepera II has eight different light and distance sensors, the location of these sensors and also the obstacle, light location and potential food can be seen in Figure 5.1.

In our context, at the beginning of the experiment reinforcement learning forces the robot to move by giving reward. If the robot starts to move, which means the first salience recognition is accomplished, a randomly seeking a food process starts. Robot evaluates the situation by means of reinforcement learning. If robot finds a food in front of it but not pick it up, reinforcement learning reinforces the robot to pick it up. The idea is same for the third salience which is called deposit the food into the nest. Each situation that robot meet causes a new calculation for the salience matrice which is comprised with random numbers in the beginning of the experiment. The inputs

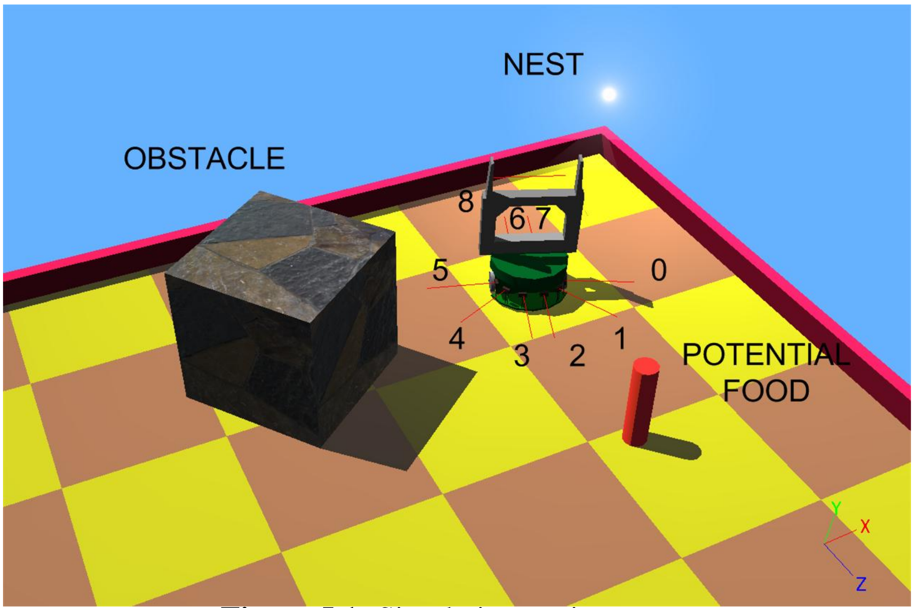

Figure 5.1: Simulation environment.

are defined by multiplying salience matrice and sensory informations. There are three different sensory informations; distance sensors' value, p_{cyl}, gripper sensors' value, p_{grip}, and light sensors' value, p_{nest}. Robot distance, gripper and light sensors give natural numbers between "0" to "1050" which means absence or presence of related object. These numbers are scaled from "0" to "1" so the variables corresponding to distance of cylinder or obstacle, gripper position and distance of the nest are denoted by rational numbers. You can see the input matrices, for at the beginning of the experiments as follows:

$$S = W_c^{beginning} I = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} p_{cyl} \\ p_{grip} \\ p_{nest} \end{bmatrix} = \begin{bmatrix} 0.0023 & 0.0021 & 0.0045 \\ 0.0034 & 0.0091 & 0.0012 \\ 0.0056 & 0.0019 & 0.0088 \end{bmatrix} \begin{bmatrix} p_{cyl} \\ p_{grip} \\ p_{nest} \end{bmatrix} \quad (5.1)$$

The complete architecture of the model can be found in Figure 5.2.

Based on the structure of Khepera II mobile robot, the distance and light sensors are used to collect data from environment. As it can be followed from Figure 5.1, Khepera II mobile robot has 8 distance and light sensors, respectively. These data collected from sensors form the model input vector, $I = [p_{cyl}\ p_{grip}\ p_{nest}]^T$, which is weighted by coefficient matrice, W_c, to define the saliencies $S = W_c I$, as mentioned above. The dimension of this matrix is determined by the number of action choices and the saliencies build up the perceptual system. This matrix is modified through

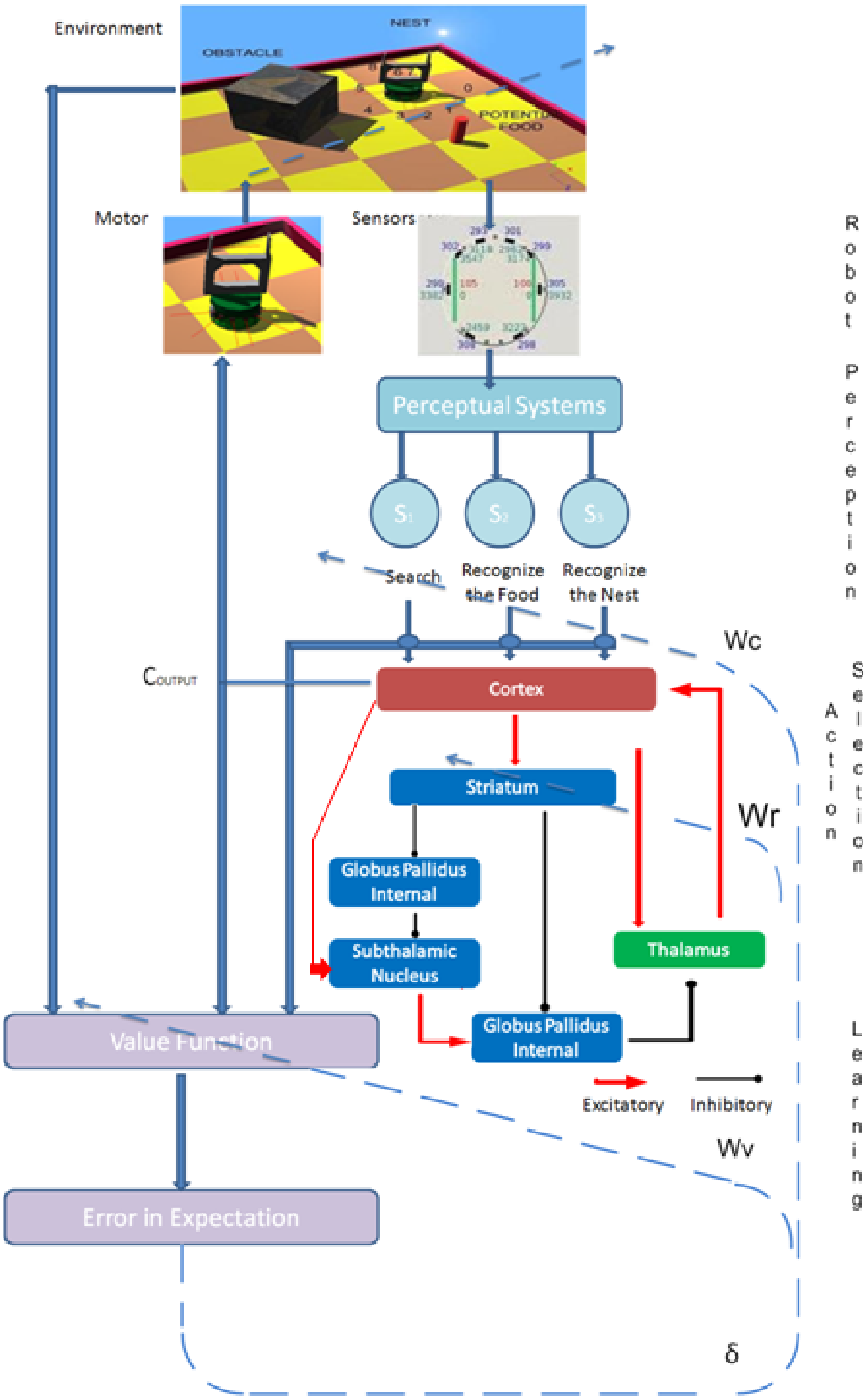

Figure 5.2: The architecture of the model realizing goal- directed behaviour.

reinforcement learning process. In the problem considered, there are three saliencies corresponding to search, recognizing the food and recognizing the nest and they are formed with the data collected from sensors. Since there are three saliencies, S, built by weighting three sensor information, I, the dimension of vectors and the matrices are $S, I \in \Re^3$ and $W_c \in \Re^{3x3}$.

Once the saliencies are established, the cortico- striato-thalamic circuit determines an action. This selected action and the reward obtained for it interacts with the learning block and the selection process in action selection block restarts at each time step (see Figure 5.2). During this process, W_c and W_r are adapted according to Equation 3.4-3.8 continuously, till the robot comes across heavy foods, potential foods, corners or nest (which is assigned by light). During the learning process, saliencies determine the action selected, and selected action is used to control the behaviour of robot. Unlike the work of Prescott et al. [42], there is no need for a busy signal as the sensor data is considered constantly.

The action selection adapted by reinforcement learning block is the contribution of this work. In [42], the idea of behavioural selection is based only on certain targets and the sensor data and it is designed on a rule based algorithm. Early studies of action selection and reinforcement learning phenomena are proposed in different contexts for separate tasks. However, both tasks are considered together in this thesis.

5.2 Simulation Results

Khepera II mobile robot is used to simulate the task of a rat searching for food in an unfamiliar environment, recognizing the nest and carrying food there. We simulated rats' intrinsic feelings in a simple learning task. The robot is placed in any starting point from which the rat could do any one of the three goal actions. Only one of the goal actions is chosen on each trial, and the chosen action is searching in the beginning of the experiment. The expectation is the robot will be able to make appropriate actions due to the reinforcement learning. So by determining the priority of sensory stimulus and effect of dopamine release on the striatum, the robot is able to give the right decision at the right time.

Experiments such as those illustrated in Figure 5.3, 5.4 and 5.5 would clarify the difference between each processes. The case in Figure 5.3 corresponds to the work in [42] where the saliencies are determined a priori.

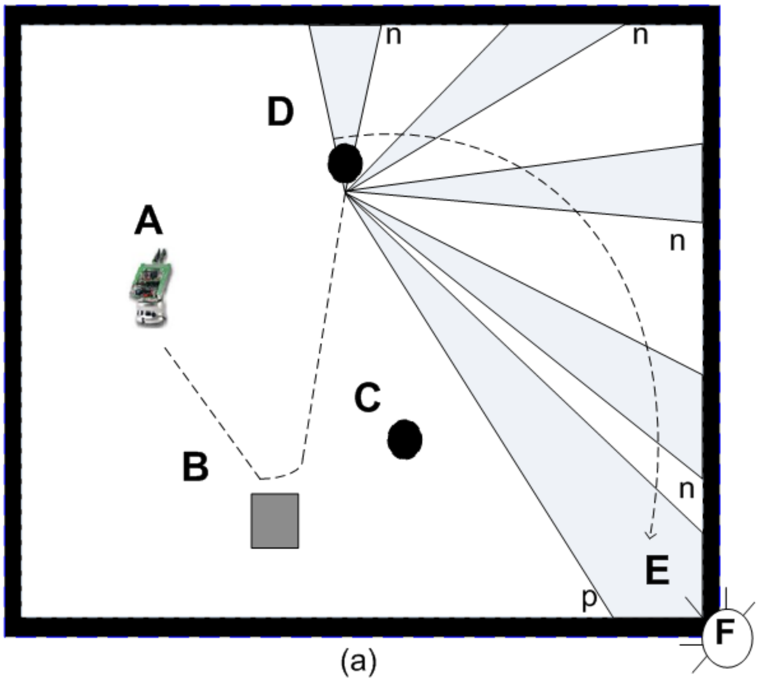

Figure 5.3: Robot foraging on an unfamiliar environment is illustrated with a priori saliencies, n: negative for light, p: positive for light, A: Khepera II, B: Obstacle, C, D: Potential food, E: Nest, F: Light.

So robot begins to search instantly with the correct choice, it recognizes the obstacle and food without mistake. When food is picked up by the robot, the light sensor begins to search light source which is the indicator of the nest. Notice that for this process the search continues until light sensors recognize the nest.

In Figure 5.4, there are no a priori determined saliencies, the robot learns the environment with the choices it makes and the rewards it obtains. Thus it begins with random search and it takes some trial and error steps till it finds food, picks it up and carries to the nest. As it can be followed in Figure 5.4, robot starts to seek the potential food in random place after the learning process for the first salience is accomplished. In the sixth trial, robot recognized the food and in the fourth trials robot recognized the nest. The trial number depends on the discount factor in temporal difference learning, γ, which is explained in Chapter 4. It means that we can control the learning speed.

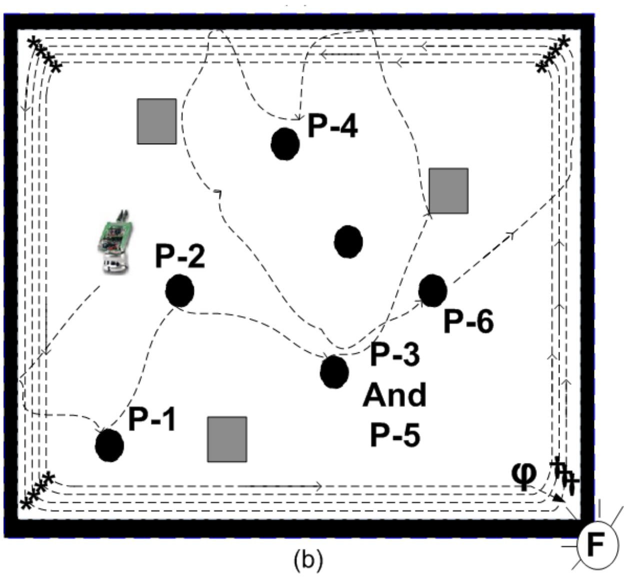

Figure 5.4: Robot foraging on an unfamiliar environment is illustrated with during learning, P: Process of Learning, ∗: negative for nest, †: nest but not enough for deposit, φ: deposit it to the nest.

Once the learning process is completed, it can immediately pick up the food and carry it to nest as shown in Figure 5.5.

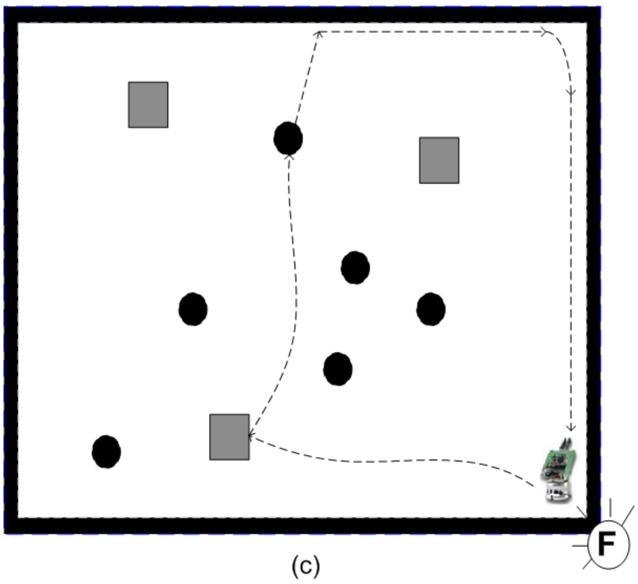

Figure 5.5: Robot foraging on an unfamiliar environment is illustrated with after learning, P: Process of Learning, ∗: negative for nest, †: nest but not enough for deposit, φ: deposit it to the nest.

As the robot is not moving at the beginning of the experiment depicted in Figure 5.4, the reinforcement learning block force it to move and begin to search. This is provided by increasing coefficient "a_{11}" through reinforcement learning. In Figure 5.6, the adaptation of coefficient "a_{11}", change in expectation error and reward are given, respectively. Once “a_{11}” is large enough and "search" salience is selected, robot begins to move.

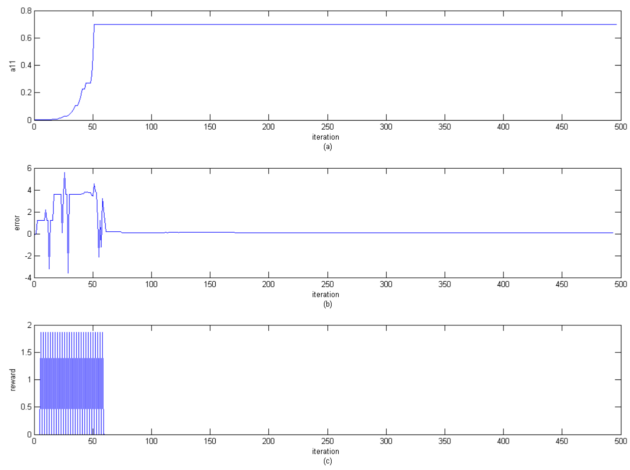

Figure 5.6: Simulation results for searching phase during learning process. After 80^{th} iteration, learning process ends for the search salience but the given illustration is continued till the 500^{th} iteration. Reward is 1.8 for this coefficient.

If Khepera II robot comes across to any one of the potential food, coefficient of “a_{22}” begins to increase. This is the learning phase of recognizing food. Results of this phase are given in Figure 5.7. Once the robot learns to recognize food and picks it up, it has

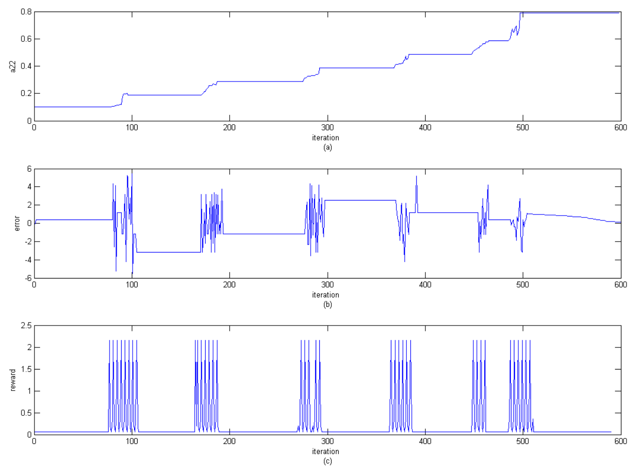

Figure 5.7: Simulation results for pick up and carrying phase. After 550^{th} iteration, learning process ends for this phase. Reward is 2.2 for this coefficient.

to begin searching the nest. In order to reach the nest the robot moves along the wall and seeks for the light source when it finds it, reward is given. This reward modifies the value of “a_{33}” and when “a_{33}” reaches a certain value, the robot learns the place of the nest. There are six coefficients besides "a_{11}", "a_{22}", "a_{33}" but these are not necessary for the determination of saliencies, so they are kept constant. Learning to deposit of food in the nest is same as the other learning routines and the results are given in Figure 5.8.

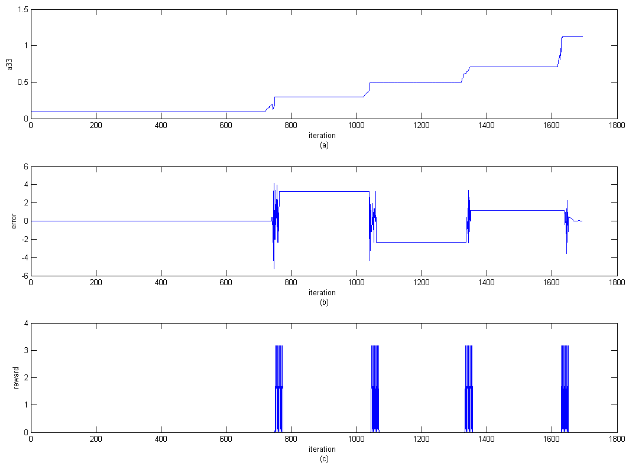

Figure 5.8: Simulation results for learning to find the nest and deposit phase. After 1700^{th} iteration, learning process ends for finding the salience so in figure 1800^{th} iteration are shown. Reward is 3 for this coefficient.

After this learning process, the efficiency of input matrice as follows:

$$W_c^{after} = \begin{bmatrix} 0.7201 & 0.0021 & 0.0045 \\ 0.0034 & 0.7975 & 0.0012 \\ 0.0056 & 0.0019 & 1.2764 \end{bmatrix} \begin{bmatrix} p_{cyl} \\ p_{grip} \\ p_{nest} \end{bmatrix} \tag{5.2}$$

Once, the robot places the food to the nest, the search for food begins again, but as it learned the food and the nest it picks up the first food it comes across and carries it to the nest directly (see Figure 5.5). As the robot does not learn the coordinates of the latest food it picked up, the searching process is made randomly.

In the beginning of these experiments, the differences between the corner and the heavy food and potential food are not considered. We added obstacles and heavy foods after the learning of the salience recognition is completed. First, system has to understand the differences between the saliencies and then be able to recognize the importance of individial circumstances. In other words, the robot learns when exactly it has to move, recognize the cylinders, i.e., it has to stop and try to pick it up whatever the weight is and finally recognizes the nest to deposit the food in any corner. After this learning process, we considered the bifurcation analysis to learn to track the differences between the potential food/heavy food and nest position, till the robot is able to recognize the food and is able to match the light to the nest position. We solved this problem by means of DA effect on Str, so in the beginning of the experiment the W_r is fixed at "0.5" but at the end of this learning process it reaches "0.85". When robot finds a cylinder in front of it, it picks up and calculates the weight by evaluating gripper aperture range. If it finds a heavy cylinder, robot immediately dismisses the cylinder. In this case, the nonlinear system determining the robot actions is at the fixed point near "0". On the other hand, when the fixed point reaches "1" robot finds one of the potential foods. The same idea is used for recognizing the nest position. If robot finds a light source in any corner, system's fixed point reaches "1" otherwise it is "0". In Figure 5.2, the illustration of this process can be found:

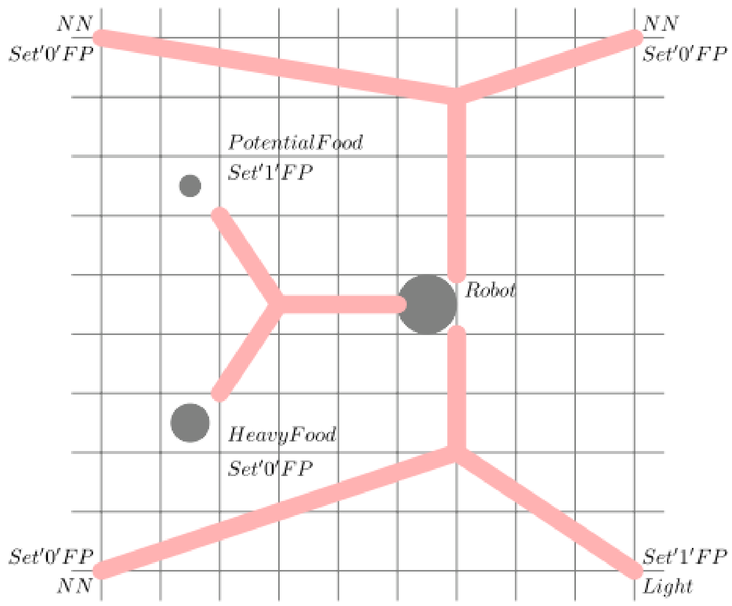

Figure 5.9: Circles illustrate the robot, heavy food and potential food, respectively in descending order. NN means 'Not a Nest' and the light illustrates the nest position. 'Set "0" FP' and 'Set "1" FP' show the which fixed point is occured in corresponding situation.

6. CONCLUSIONS AND DISCUSSIONS

Even the role of the basal ganglia circuits in action selection have been extensively studied, there is a very active research on basal ganglia circuits role in learning as the neurophysiological experimental results indicate that these networks are related to reward-related learning. Especially neuroscientists are focusing on striatal medium spinny neurons and dopamine transmitters in the context of goal-directed or incentive learning [55–57]. These experimental results show that the dopaminergic neurons' activity in SNc becomes more active during the reward related situations [4, 58]. In [4,59], the authors of these papers claim that the idea of the error signal in temporal difference learning and the training signal in striatal neurons have a significant theoretical connection. The motivation of this thesis is based on these discussions.

In this thesis, the neurocomputational model for the basal ganglia circuit proposed in [7] is improved to include both direct, indirect and hyperdirect pathways of basal ganglia circuit. Furthermore the effect of dopamine release on the striatum is shown to be captured with this model by bifurcation analysis. The bifurcation analysis are obtained by XPPAUT and in house built MATLAB codes. Nonlinear dynamical system approach to modeling the neural structures gives us the possibility of understanding the phenomenon modeled by bifurcation analysis. So based on these bifurcation analysis, the learning method for basal ganglia circuits is proposed. It is shown that robot implementation of neural circuits which are capable of realizing reinforcement learning is possible. Here, the model proposed in [7] is reconsidered and implemented on mobile robot Khepera II to mimic the behaviour of a rat searching for food in an unfamiliar environment. It has to be emphasized that a more complex cognitive process than action selection, i.e., goal-directed behaviour is implemented on a mobile robot. So, the work considered here improves [42], in two aspects, reinforcement learning process is implemented on Khepera II with considering bifurcation analysis in Str and goal-directed behaviour is realized. The task considered could be easily upgraded for more complex scenarios. Here the choices of the robot are determined by saliencies depending on sensor data and dopamine effect on Str.

The work presented here mainly depends on previous works presented in [7]. The improvement of this work is implementing reinforcement learning in a robot model which is not considered in [42]. Also, here the bifurcation analysis of the system corresponding to BG circuit is given and learning is explained with these results. First a model for the task of a rat learning to find food, recognize the nest and carry it to nest in an unfamiliar environment given and realized in MATLAB and these results are given in Chapter 4, then details of implementing the task on Khepera II are presented in Chapter 5. Though the task considered is simple, the results are encouraging for further trails, especially with more complicated scenarios. An ultimate goal would be to develop the computational model of cortico-striato-thalamic circuit further where the intrinsic variables of hunger and fear do change according to the on going process in the limbic system. A simpler and easy to realize goal would be to adapt more parameters in matrix with reinforcement learning.

This approach can be further used to establish a framework for understanding the cause of physiological diseases related with BG circuits. The level of the dopamine has a role in different physiological disorders as it is well-known that loss of dopamine level in BG circuit causes Parkinson's disease which means difficulty in accomplishing an action as in Figure 3.4, on the other hand excessive dopamine level in BG circuit causes Huntington's disease which means choosing more than one action at a time, see Figure 3.5.

REFERENCES

[1] **Dayan, P.**, 2009. Dopamine, Reinforcement Learning, and Addiction, *Pharmacopsychiatry*, **42**, 56–65.

[2] **Gurney, K.**, **Prescott, T.J. and Redgrave, P.**, 2001. Computational Model of Action Selection in the Basal Ganglia I: A New Functional Anatomy, *Biological Cybernetics*, **84(6)**, 401–410.

[3] **Taylor, J.G. and Taylor, N.R.**, 2000. Analysis of Recurrent Cortico-Basal Ganglia-Thalamic Loops for Working Memory, *Biological Cybernetics*, **82(5)**, 415–432.

[4] **Schultz, W.**, **Dayan, P. and Montague, P.**, 1997. A Neural Substrate of Prediction and Reward, *Science*, **275**, 1593–1599.

[5] **Karabacak, O. and Sengor, N.**, 2006. A computational model for the effect of dopamine on action selection during Stroop test, Proceedings of International Conference of Artificial Neural Networks (ICANN'06), volume4131, Lecture Notes in Computer Science, pp.485–494.

[6] **Metin, S. and Sengor, N.**, 2010. A neurocomputational model of nicotine addiction based on reinforcement learning, Proceedings of International Conference of Artificial Neural Networks (ICANN'10), Lecture Notes in Computer Science, pp.228–233.

[7] **Sengor, N.S.**, **Karabacak, O. and Steinmetz, U.**, 2008. A Computational Model of Cortico-Striato-Thalamic Circuits in Goal-Directed Behavior, Proceedings of International Conference of Artificial Neural Networks (ICANN'08), Springer- Verlag, pp.328–337.

[8] **Gurney, K.**, **Prescott, T.J. and Redgrave, P.**, 2001. Computational Model of Action Selection in the Basal Ganglia II. Analysis and simulation of behaviour, *Biological Cybernetics*, **84(6)**, 411–423.

[9] **Gillies, A. and Arbuthnott, G.**, 2000. Computational Models of the Basal Ganglia., *Movement Disorders*, **15(5)**, 762–770.

[10] **Alexander, G.E.**, **Cruther, M.D. and DeLong, M.R.**, 1990. Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions, *Progress in Brain Research*, **85**, 119–146.

[11] **Gutkin, B.S.**, **Dehaene, S. and Changeux, J.P.**, 2000. A Neurocomputational Hypothesis for Nicotine Addiction, *PNAS*, **103(4)**, 1106–1111.

[12] **Izhikevich, E.M.**, 2007. Dynamical Systems in Neuroscience, The Geometry of Excitability and Bursting, The MIT Press.

[13] **Kuznetsov, Y.A.**, 1998. Elements of Applied Bifurcation Theory, Second Edition, Springer-Verlag New York, Inc.

[14] **Strogatz, S.H.**, 1994. Nonlinear Dynamics and Chaos, With Applications to Physics, Biology, Chemistry and Engineering, Perseus Books Publishing, L.L.C.

[15] **Khalil, H.K.**, 2002. Nonlinear Systems, Third Edition, Prentice-Hall, Inc.

[16] **Hodgkin, A. and Huxley, A.**, 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve, *The Journal of Physiology*, **117**, 500–544.

[17] **Aulbach, B. and B., K.**, 2001. On Three Definitions of Chaos, *Nonlinear Dynamics and Systems Theory*, **1(1)**, 23–37.

[18] **Lorenz, E.**, 1963. Deterministic nonperiodic flow, *Journal of the Atmospheric Sciences*, **20**, 130–141.

[19] **Ermentrout, B.**, 2006, XPPAUT, `http://www.scholarpedia.org/article/XPPAUT`, scholarpedia.

[20] **Burke, J.**, **Desroches, M.**, **Barry, A.M.**, **Kaper, T.J. and Kramer, M.A.**, 2012. A showcase of torus canards in neuronal bursters, *The Journal of Mathematical Neuroscience*, **2(3)**.

[21] **Guckenheimer, J. and Holmes, P.**, 1983. Nonlinear Oscillations, Dynamical systems and Bifurcations of Vector Fields, Springer.

[22] **Ermentrout, B.**, 2002. Simulating, Analyzing, and Animating Dynamical Systems: A Guide to Xppaut for Researchers and Students.

[23] **Izhikevich, E.M.**, 2000. Subcritical Elliptic Bursting of Bautin Type, *SIAM J. Appl. Math.*, **60(2)**, 503–535.

[24] **Murdock, J.**, 2006, Normal Forms, `http://www.scholarpedia.org/article/Normal_forms`, scholarpedia.

[25] **Graybiel, A.M.**, 1998. The basal ganglia and chunking of action repertoires, *Neurobiology of Learning and Memory*, **70**, 119–136.

[26] **Cajal, R.y.**, 1899. Comparative Study of the Sensory Areas of the Human Cortex, *Santiago*.

[27] **Redgrave, P.**, **Prescott, T. and Gurney, K.**, 1999. The basal ganglia: A vertebrate solution to the selection problem?, *Neuroscience*, **89**, 1009–1023.

[28] **Haber, S.N.**, 2010. The reward circuit: Linking primate anatomy and human imaging, *Neuropsychopharmacology Reviews*, **35**, 4–26.

[29] **Saeb, S.**, **Weber, C. and Triesh, J.**, 2009. Goal-directed learning of features and forward models, *Neural Networks*, **22**, 586–592.

[30] **Houk, J.**, **Bastianen, C.**, **Fansler, D.**, **Fishbach, A.**, **Fraser, D. and Reber, P.J., R.S.S.L.**, 2007. Action selection and refinement in subcortical loops through basal ganglia and cerebellum, *Phil. Trans. R. Soc. B*, **362(1485)**, 1573–1583.

[31] **Dayan, P. and Daw, N.**, 2008. Decision theory, reinforcement learning and brain, *Cognitive, Affective and Behavioral Neuroscience*, **8(4)**, 429–453.

[32] **Frank, M.J.**, 2011. Computational models of motivated action selection in corticostriatal circuits, *Current opinion in Neurobiology*, **21**, 381–386.

[33] **Gurney, K.N.**, 2009. Computational Models in Neuroscience, From Membranes to Robots, *Computational Modelling in Behavioural Neuroscience*, 107–136.

[34] **Haber, S.N.**, **Fudge, J.L. and McFarland, N.R.**, 2000. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum, *Journal of Neuroscience*, **20(6)**, 2369–2382.

[35] **Mink, J.W.**, 1996. The basal ganglia: Focused selection and inhibition of competing motor programs, *Progress in Neurobiology*, **50**, 381.

[36] **Albin, R.L.**, **Young, A.B. and Penney, J.B.**, 1989. The functional anatomy of basal ganglia disorders, *Trends in Neurosciences*, **12**, 366.

[37] **Byrne, J.H.**, 1997. Neuroscience Online: An Electronic Textbook for the Neurosciences, Department of Neurobiology and Anatomy, The University of Texas Medical School at Houston.

[38] **Wiecki, T.V. and Frank, M.J.**, 2010. Neurocomputational models of motor and cognitive deficits in Parkinson's disease, *Progress in Brain Research*, **183**, 275–297.

[39] **Montoya, A.**, **Price, B.H.**, **Menear, M. and Lepage, M.**, 2006. Brain imaging and cognitive dysfunctions in Huntington's disease, *Journal of Psychiatry Neuroscience*, **31(1)**, 21–29.

[40] **Crossman, A.R.**, 2000. Functional anatomy of movement disorders, *Journal of Anatomy*, **196**, 519–525.

[41] **Amos, A.**, 2000. A Computational Model of Information Processing in the Frontal Cortex and Basal Ganglia, *Journal of Cognitive Neuroscience*, **12(3)**, 505–519.

[42] **Prescott, T.J.**, **Gonzalez, F.M.M.**, **Gurney, K.**, **Humphries, M.D. and Redgrave, P.**, 2006. A robot model of the basal ganglia: Behavior and intrinsic processing, *Neural Networks*, **19**, 31–61.

[43] **Girard, B.**, **Tabareau, N.**, **Pham, Q.C.**, **Berthoz, A. and Slotine, J.J.**, 2008. Where neuroscience and dynamic system theory meet autonomous robotics: A contracting basal ganglia model for action selection, *Neural Networks*, **21(4)**, 628–641.

[44] **Seger, C.A.**, 2006. The Basal Ganglia in Human Learning, *Neuroscientist*, **12(4)**, 285–290.

[45] **Humphrys, M.**, 1997. Action Selection Methods Using Reinforcement Learning, Ph.D. thesis, Cambridge, Trinity Hall.

[46] **Cohen, M.X. and Frank, M.J.**, 2009. Neurocomputational models of basal ganglia function in learning, memory and choice, *Behavioural Brain Research*, **199**, 141.

[47] **Frank, M.J.**, 2005. Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism., *Journal of Cognitive Neuroscience*, **17**, 51.

[48] **Smith-Roe, S.L. and Kelley, A.E.**, 2000. Coincident activation of nmda and dopamine d1 receptors within the nucleus accumbens core is required for appetitive instrumental learning, *Journal of Neuroscience*, **20**, 7737.

[49] **Dayan, P. and Watkins, C.J.C.H.**, 2006. Reinforcement Learning: A Computational Perspective, John Wiley and Sons, Ltd.

[50] **Rescorla, R.A. and Wagner, A.R.**, 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement, *Classical Conditioning II*, 64–99.

[51] **Bullock, D.**, **Tan, C.O. and John, Y.J.**, 2009. Computational perspectives on forebrain microcircuits implicated in reinforcement learning, action selection, and cognitive control, *Neural Network*, **22**, 757–765.

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

[53] **Webb, B.**, 2000. What does robotics offer animal behavior?, *Animal Behavior*, **60**, 545–558.

[54] **Fleischer, J.G. and Edelman, G.M.**, 2009. Brain-based devices, *IEEE Robotics and Automation Magazine*, 33–41.

[55] **Dayan, P. and Balleine, B.W.**, 2002. Reward, motivation, and reinforcement learning, *Neuron*, **35(2)**, 285–298.

[56] **Hollerman, J.R.**, **Tremblay, L. and Schultz, W.**, 2000. Involvement of basal ganglia and orbitofrontal cortex in goal-directed behavior, *Progress in Brain Research*, **126**, 193–215.

[57] **Kimura, M.**, 1995. Role of basal ganglia in behavioral learning, *Neuroscience Research*, **22(4)**.

[58] **Schultz, W.**, **Apicella, P. and 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(3)**.

[59] **Montague, P.R.**, **Dayan, P. and Sejnowski, T.J.**, 1996. A framework for mesencephalic dopamine systems based on predictive hebbian learning, *Journal Of Neuroscience*, **16(5)**.

APPENDICES

APPENDIX A.1 : Medium Spinny Neuron Model

APPENDIX A.1

Table A.1: Parameters for Medium Spinny neuron

Column A	Column B	Column C	Column D
$I_{app} = 0$	$gL_{str} = 0.3$	$gK_{str} = 36$	$gNa_{str} = 120$
$gL_CaStr = 0.001$	$vL_str = -54.6$	$vK_str = -77$	$vNa_str = 55$
$Cm = 1$	$vCa_str = 140$	$gCa_str = 0.4$	$gK_strv1 = 2.8$
$gAHP_str = 11$	$eps_str = 0.0001$	$kCa_str = 20$	$kL_str = 5$
$z_Ca = 2$	$F = 96485$	$R = 8314$	$q = 310$
$Ca_Me = 2.0$	$vSp_Sp = -60$	$tetag1 = 30$	$tetag = 15$
$alfa_str = 0.01$	$beta_str = 1.43$	$w_D1 = 1$	$w_D2 = 1$
$sigmas = 8$	$tetas = -42.97$	$gD1_Ca = 1$	$gD1_Na = 3$

Current equations for Medium Spinny Neuron as follows:

$$IL_{strD} = gL_{str}(vc - vL_{str}) \quad \textbf{(A.1)}$$

$$IK_{strD} = gK_{str}mc^4(vc - vK_{str}) \quad \textbf{(A.2)}$$

$$INa_{strD} = gNa_{str}nc^3kc(vc - vNa_{str}) \quad \textbf{(A.3)}$$

$$IL_{CastrD} = ((gL_{Castr}lc^2vcz_{Ca}^2F^2)/(Rq))((Ca_{Me}exp(-vcz_{Ca}F/(Rq)) - tc) \quad \textbf{(A.4)}$$
$$/(1 - exp(-vcz_{Ca}F/(Rq))))$$

$$IAHP_{strD} = gAHP_{str}(vc - vK_{str})(tc/(tc + kL_{str})) \quad \textbf{(A.5)}$$

$$ICa_{strD} = gCa_{str}ssnsz_{str}vc^2(vc - vCa_{str}) \quad \textbf{(A.6)}$$

$$IKv1_{strD} = gK_{strv1}nc^2kc(vc - vK_{str}) \quad \textbf{(A.7)}$$

$$ID1 = gD1_{Ca}msnsz_{strL}vc^2(vc - vCa_{str}) \quad \textbf{(A.8)}$$
$$+gD1_{Na}msnsz_{str}vc^3(vc - vNa_{str})$$

State variables for Medium Spinny Neuron as follows:

$$vc' = (-IL_{strD} - IK_{strD} - INa_{strD} - ICa_{strD} - IL_{CastrD} \quad \textbf{(A.9)}$$
$$-IAHP_{strD} - IKv1_{strD} + I - app - ID1)/Cm$$

$$nc' = (msnsz_{str}(vc) - nc)/taum_{str}(vc) \quad \textbf{(A.10)}$$

$$mc' = (nsnsz_{str}(vc) - mc)/taun_{str}(vc) \quad \textbf{(A.11)}$$

$$kc' = (hsnsz_{str}(vc) - kc)/tauh_{str}(vc) \quad \textbf{(A.12)}$$

$$lc' = (msnsz_{strL}(vc) - lc)/taum_{strL}(vc) \quad \textbf{(A.13)}$$

$$tc' = eps_{str}(IL_{CastrD} - kCa_{str}tc) \quad \textbf{(A.14)}$$

Gate equations for Medium Spinny Neuron as follows:

$$msnsz_{str}(x) = (-40-x)/(-40-x+1.1exp(-(14x+360)/90)-1.1exp((-x)/18)) \quad \textbf{(A.15)}$$

$$nsnsz_{str}(x) = (-55-x)/((-55-x)+5.545exp(-(9x+440)/80)-5.545exp((-x)/80)) \quad \textbf{(A.16)}$$

$$hsnsz_{str}(x) = (0.0027exp(-x/20+(-35-x)/10)+0.0027exp(-x/20)) \\ /(0.0027exp(-x/20+(-35-x)/10)+0.0027exp(-x/20)+1) \quad \textbf{(A.17)}$$

$$ssnsz_{str}(x) = 1/(1+exp(-(x-tetas)/sigmas)) \quad \textbf{(A.18)}$$

$$tauh_{str}(x) = (exp((-35-x)/10)+1)/(0.0027exp((-x)/20+(-35-x)/10) \\ +0.0027exp((-x)/20)+1) \quad \textbf{(A.19)}$$

$$taum_{str}(x) = (exp((-40-x)/10)-1)/(0.1(-40-x)+0.108exp((-40-x)/10 \\ -x/18)-0.108exp(-x/18)) \quad \textbf{(A.20)}$$

$$taum_{strL}(x) = (exp((-10-x)/10)-1)/(0.1(-10-x)+1.2exp((-10-x)/10 \\ -x/50)-1.2exp(-x/50)) \quad \textbf{(A.21)}$$

$$taun_{str}(x) = (exp((-55-x)/10)-1)/(0.01(-55-x) \\ +0.0549exp(-(9x+440)/80)-0.0549exp((-x)/80)) \quad \textbf{(A.22)}$$

$$msnsz_{strL}(x) = (-28-x)/(-28-x+10exp(-(14x+252)/70)-10exp((-x)/14)) \quad \textbf{(A.23)}$$

Bifurcation diagram for Medium Spinny Neuron is given in Figure A.1.

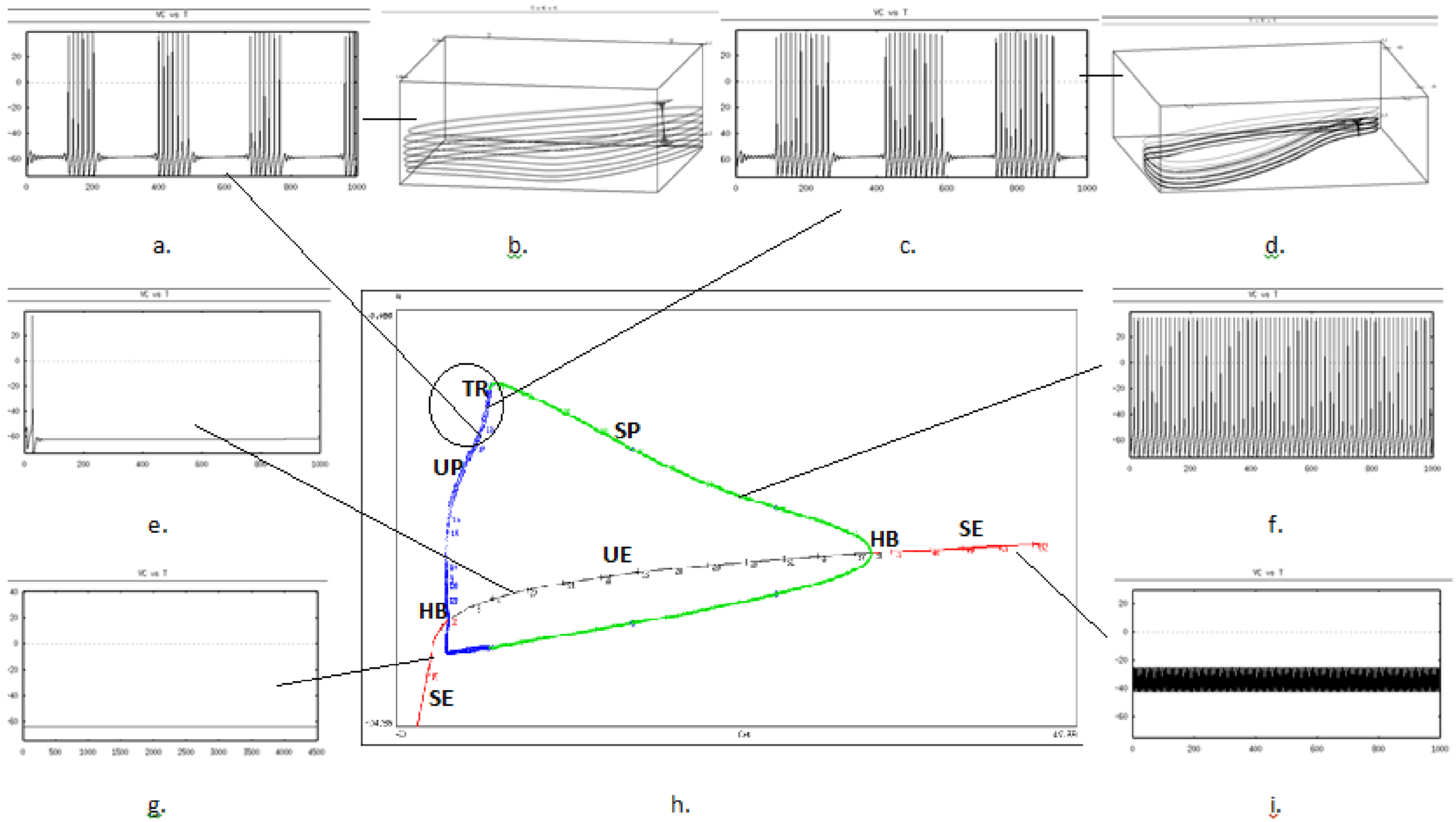

Figure A.1: Bifurcation diagram with significant phase portraits. (SE: Stable Equilibrium, UE: Unstable Equilibrium, HB: Hopf Bifurcation, SP: Stable Periods, UP: Unstable Periods, TR: Torus Bifurcation) a.I_{app} =17.3, system shows bursting generation near Torus attractor which can be shown in b., c. I_{app} =21.6, system shows bursting generation near Torus attractor which can be shown in d., e. System generates only one spike in unstable equilibrium point, f. The behaviour of the neuron is unstable, it creates regular spikes without frequency adaptation, g. I_{app} =0.6, it is too small to generate any spikesI_{app} =346.7, same as in f. but small amplitude.

PHOTO

CURRICULUM VITAE

Name Surname: Berat Denizdurduran

Adress: ITU Electronics & Communications Dep. Neuroscience Research and Modelling Group, Room No: 1117, 34469, Maslak, Istanbul, Turkey

E-Mail: denizdurdu@itu.edu.tr

B.Sc.: Electronics Eng., Istanbul Technical University

M.Sc.: Electronics Eng., Istanbul Technical University

Professional Experience and Rewards:

▪ Barcelona Cognition, Brain and Technology Summer School 2011 Travel Grant

▪ Invited research assistant in Blue Brain Project at Ecole Polytechnique Federal de Lausanne (EPFL).

List of Publications and Patents:

▪ **Denizdurduran, B.**: "Analyzing Different Neuron Models with XPPAUT," The 4^{th} International Student Conference on Advanced Science and Technology, 2010, Izmir, Turkey.

▪ **Denizdurduran, B.**, Sengor, N.S.: "A Biophysical Network Model for Action Selection," Bernstein Conference Computational Neuroscience/Neurotechnology and Neurex Annual Meeting, 2011, Freiburg, Germany.

PUBLICATIONS/PRESENTATIONS ON THE THESIS

▪ **Denizdurduran B.**, Sengor N. S., 2012: A Realization of Goal-Directed Behavior, Implementing a Robot Model Based on Cortico-Striato-Thalamic Circuits. *Proceedings of the* 4^{th} *International Conference on Agents and Artificial Intelligence*, February 6-8, 2012 Algarve, Portugal.

▪ **Denizdurduran B.**, Sengor N. S., 2011: Implementation of Cortico-Striato-Thalamic Circuits on Robot. 10^{th} *National Neuroscience Meetings*, February 9-12, 2011 Istanbul, Turkey.