Universiti Teknologi MARA

Reinforcement Learning for Autonomous Online Swing-Up and Balancing Agent

Muhammad'Umar Rasydan Bin Romli

Thesis submitted in fulfilment of the requirements for
Bachelor of Science (Hons.) Intelligent System
Faculty of Computer and Mathematical Sciences

August 2014

ACKNOWLEDGEMENT

"In the name of Allah, the most Gracious and the most Merciful"

Alhamdulillah, praise be to *Allah* for providing me His almighty blessings to complete the project. Without His help, mercy and facility it would be impossible for me to complete this project.

Special appreciation goes to my mother, for the continuous moral support and Du'a throughout all my studies.

I'm deeply grateful to my supervisor, Dr Shuzlina Abdul Rahman for her concern, guidance, patience and relentless support even with all the trouble I faced throughout the two semester she supervised me. I consider myself very fortunate for being able to work with a very encouraging supervisor such as her from the beginning of the journey of this final year project up to the completion of this report. Without her support, completion of this report would be close to impossible.

I'm indebted to my CSP600 lecturer, Prof Dr Azlinah Mohamed for teaching and her tutelage in improving the proposal of this project and to Mr. Zaki Zakaria as the examiner of the proposal in approving the proposal to continue developing this project and also all of the lecturers that nurture and helped me throughout my study in UiTM, Shah Alam.

I'm thankful and honoured for the interest, guidance and monetary support given by GranuLab (M) Sdn. Bhd. towards the completion of this project.

Finally, I am thankful to have CS2236A as my classmates who supported me in any respect during the completion of this project.

Thank you and may Allah bless all of them.

ABSTRACT

Reinforcement Learning (RL) is a form of machine learning that learns through trial-and-error method. Uniqueness of RL is in the way problems can be modelled by specifying the goal, reward when that goal is achieved and punishment when it is not achieved which is a more natural representation of learning based problems. There is a potential for RL to be the main learning algorithm used in various fields such as, robotics, economics, and game theory which is a major prospect of this project. The problem with the current balancing agent implementation of RL is either to balance or to swing-up but it never occur simultaneously. The implementation of RL is always based upon the capability of either autonomous or online but not both. This project is designed with the testing of the RL algorithms in a simulated environment for the domain of a swing-up and balance. The design also requires the capability of autonomy and running in real time. The Q-Learning algorithm with improvements tested from the simulation are implemented in a physical agent to for a real world testing. The core approach of this project is to bring Q-Learning algorithms to the test with few modification to the algorithm in order to know the capability of RL in a real world environment. The project completed successfully where the modified algorithms shows convergence even though the pole is not balanced. The convergence of the modified RL and the implementation of the modified RL into a physical agent shows an unexpected intelligence by the agent.

TABLE OF CONTENTS

CONTENTS **PAGE**

SUPERVISOR'S APPROVAL i

STUDENT'S DECLARATION ii

ACKNOWLEDGEMENT iii

ABSTRACT iv

TABLE OF CONTENTS v

LIST OF FIGURES ix

LIST OF TABLES x

CHAPTER ONE: INTRODUCTION 1

1.1 Background of Study 1

1.2 Problem Statement 2

1.3 Research Questions 3

1.4 Research Objectives 3

1.5 Project Scopes 4

1.6 Significance of the Project 5

1.7 Summary 6

CHAPTER TWO: LITERATURE REVIEW 7

2.1 Agent 7

CHAPTER 1

INTRODUCTION

This chapter provides the background and foundation of the project, problem statement, research questions, objectives, scope of project, significance in the field of reinforcement learning and finally the summary of the project.

1.1 Background of Study

Reinforcement Learning (RL) is a type of machine learning in the area of Artificial Intelligence (AI) that learns through trial-and-error or rather the carrot-and-stick method which is rooted in the field of psychology. Contrary to supervised and unsupervised learning, RL focuses on sequential decision making and learning from delayed reward making it intuitive when defining it for a problem as it learns lifelike. Hence, the main motivation for this project is in the potential of RL to be the go to learning algorithm for anyone wanting to incorporate AI into their venture.

This project focuses on the development of an algorithm for autonomous online swing-up and balancing agent. An algorithm is autonomous when its parameters are adjusted automatically without any human intervention. Online refers to learning algorithms that learn in real-time without stopping in between. In reinforcement learning however the term real-time and online are sometimes used interchangeably but have different meaning, continuous and iterative. Continuous are as defined previously but iterative is when it is learning in real-time but in steps or rather the agent needs to reset to learn in each step.

To develop the algorithm, there must be a platform in which it can be tested. Hence, a modified variant of the classic cart-pole problem is chosen as a testing