Electronics and Computer Science

Faculty of Physical and Applied Sciences

University of Southampton

Daniel Lee

03/04/2017

Linear Quadratic Regulation using Reinforcement Learning

Project supervisor: Dr Bing Chu

Second examiner: Dr Mohammed El-Hajjar

Abstract

The aim of this project is to demonstrate that the *optimal control* for a *Linear Quadratic Regulation* (LQR) problem can be obtained by the use of *Reinforcement Learning* methods, without prior knowledge of the system’s dynamics. For a better understanding of the core topics of reinforcement learning, simulations and demonstrations of core reinforcement learning topics, including *Bellman equations*, *greedy* and *-greedy* algorithms, *policy evaluation*, *policy iteration*, *policy improvement*, *temporal difference learning*, *exploration* and *exploitation,* *Markov Decision Process* and *Q-Learning,* were performed. Related to control systems, research on *optimal control theory*, *state-space modelling*, *cost function* and *Algebraic Ricatti Equation* was also made, as it was an essential requirement for this project. Ultimately, this background theory was applied to solve the inverted pendulum problem, which consists of balancing an inverted pendulum attached to a cart that can only move sideways, through the use of the adaptive policy iteration with Q-Learning.

**1. Introduction**

If an accurate description of the dynamics of a system is available, the solution for the optimal control, within the constraints of LQR, can be trivially obtained by solving the Algebraic Ricatti equation [1]. However, in many cases, an exact representation of a dynamic system’s model cannot be easily found. This is where reinforcement learning can fill in the gaps. A very powerful feature of reinforcement learning is that it can *learn* the system dynamics through experience. This means that by analysing several aspects of the behaviour of a system throughout many finite iterations, or *episodes*, a close estimate of the system dynamics can be computed, therefore allowing us to find the optimal policy.

To solve this problem, the most fundamental reinforcement topics were first studied and demonstrated in simulations, to build the required basic understanding, as shown in the following chapters.

**1.1 Background Reading**

Background reading was mainly based on the topics of reinforcement learning and control theory. Many of the concepts and equations used throughout this project are based on the book ‘Reinforcement Learning: An Introduction” by Sutton and Barto.

Reinforcement learning is defined as